Proceedings of International Conference on Computational Statistics & Mathematical Applications (ICCSMA 2K22)

30th-31st March, 2022

Organized by

APOLLO

ARTS AND SCIENCE COLLEGE CHENNAI

(Approved by Govt. of Tamil Nadu & Affiliated to University of Madras)

Kalvoy Village, Guduvanchery - 603 108.

Department of Statistics

in association with NORM D FEST

&

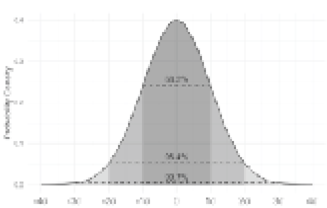
ARYABHATTA JOURNAL OF MATHEMATICS & INFORMATICS

ISSN (Print) : 0975-7139 ISSN (Online) : 2394-9309

VOL. 14 **Special Conference Issue**

Indexed in Copernicus, Google Scholars, CNKI scholar, EBSCO Discover,
Indian Science Abstract (ISA) Indian Citation Index **(ICI)** Cite factor & indianjournals.com

Double-Blind Peer Reviewed Refereed International Journal



Editor-in-Chief :
Dr. T.P. Singh

Editors :
Mrs. M. Aniji
Dr. G. Manimannan

Website : www.aryanseducationtrust.com (www.abjmi.com)

(ICCSMA 2K22)

Resource Persons

Name & Profile	Photo
Dr. Jothi Kumar Joint Director of Collegiate Education (P and D), (Retired) Dean, Faculty of Science and Humanities, SRM Institute of Science and Technology Kattangulathur (Retired), Chennai.	
Dr. Vasna Josua Scientist C Division of Epidemiology and Biostatistics Indian Council of Medical Research, Chennai.	
Dr. Anbupalam Thalamuthu Senior Research Fellow and Statistical Genetics at the Center for Health Brain Ageing (CHeBA), University of New South wales,, Australia.	
Dr. Senthil Kumar Dean-SSIT and Associate Professor, Skyline University, Kano, Nygeria	
Dr. T.P. SINGH M.Sc. (Maths & Chem.), Ph.D. (Maths) PGDCA Professor-PG & Research Dept. of Mathematics & OR, Baba Mast Nath University, Hrayana. Visiting Professor, MM University, Mullana, Ambala	
Dr. Kirthiga Professor, Department of Statistics, Pondichery University, Puduchery	
Dr. R. Kannan, Professor, Department of Statistics,Annamalai University Gurumoorthy Pattabiraman AI & ML Practitioner & Leader, Corporate Trainer & Adjunct Faculty, Pune.	
Gurumoorthy Pattabiraman AI & ML Practitioner & Leader, Corporate Trainer & Adjunct Faculty, Pune.	
Dr. M. Ramakrishnan Assistant Professor, RMK.Vivekanandha College (Autonomous). Mylapore, Chennai-600004	

(ICCSMA 2K22)

Organizing Members

Name & Profile	Photo
Mrs S. Saraswathy M.Sc.,M.Phil.,D.T.Ed., Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.	
Mrs. A. M. Priyanka M Sc.,M Phil., Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.	
Mr. R. Rajesh MSc., BEd., MPhil., Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.	
Mr. P. Murugaian M Sc.,M Phil., Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.	
Mr. S. Balamurugan, MBA (Gen.), MBA (EM). MMV, DSA, Assistant Professor, Department of Commerce, Apollo Arts and Science College Chennai.	

Editors:

Name & Profile	Photo
Mrs. M.Aniji MSc.,MPhil.,ME., Head & Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.	
Dr. G. Manimannan MSc.,MPhil.,MCA.,PhD Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai	

PRINCIPAL'S MESSAGE

"An investment in knowledge pays the best interest" said Benjamin Franklin. I am glad that the Department of Statistics, Apollo Arts and Science College Chennai invests and decipher knowledge for the wellbeing of all knowledge seekers through the international conference on Computational Statistics and Mathematical Applications on 30th March 2022. I believe that this conference not only helps to enhance the knowledge on computational Statistics but also the contemporary Mathematical applications to reach the pinnacle of productive success in teaching.

I congratulate the Head of the Department and all the faculty members of Statistics, for their challenging endeavorto conduct it in an effective manner and wish the conference a great success.

Dr. R. MUGUNDAN,

M Com.,MPhil.,MBA.,LLB.,PhD.

PRINCIPAL

ISSN (Print) : 0975-7139 , ISSN (Online) : 2394-9309

ARYABHATTA JOURNAL OF MATHEMATICS & INFORMATICS

Volume 14 | Special Conference Issue | 30th-31st March, 2022

CONTENTS (ICCSMA 2K22)

DOI No. : 10.5958/2394-9309.2022.00001.4
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A PERFORMANCE OUTLOOK OF GROUPING COMPANIES THROUGH DATA MINING APPROACH

Krithika Subbiah*, N. Manjula Devi and Manimannan.G*****
*M.Sc. Public health, University of Waterloo, Ontario,Canada
**Biostatistician, KarpagaVinayaga institute of medical science and research center, Chengalpattu
***Assistant Professor, Department of statistics, Apollo Arts and Science College Chennai, Guduvanchery
E-mail : manimannang@gmaill.com

ABSTRACT

Three different methods of rating and projecting the performances of the top ranking companies on the basis of certain financial ratios based on data mining techniques are proposed in this paper. It is well known that statistical information on financial ratios are being extensively used by researchers for many purposes. To facilitate the present study, the financial information of public and private sector companies rated as the best with reference to net sales, published by kaggle.com were considered for the period from 2015 to 2020. Out of numerous ratios, twenty financial ratios were sieved carefully that had different notions of the objectives and significant meaning in the literature. To exploit the hidden structure present in the data, data mining tools such as factor, k-means clustering and discriminant analyses are applied in succession. Factor analysis is initiated first to uncover the structural patterns underlying financial ratios. The scores from extracted factors were then used to find initial groups by K-means clustering algorithm to prune the data.

In Method I, clusters obtained by extracted factor are followed by iterative discriminant procedure with original ratios until cent percent classification was achieved for the year 2015. Data in the following years are trained using the previous year group means to obtain initial groups which, then iterated until cent percent classification is reached. In Method II, it is assumed that the rates of increase in the group means are constants from one year to the next, and the group means are increased by the corresponding rates, to mine the data for the years from 2016 onwards to inherent the patterns from the previous year. Method III mines the data using the estimated means for the clusters, assuming that the performance of the companies followed linear trend over the years. From the present study it is observed that the three different approaches on the classifications are equivalently good. Method I have slightly higher aggregate mean of 93.5% of correct classification than the other two approaches. It is also interesting to note that the clusters obtained by all the three methods could be arranged according to the magnitude of their group means on the ratios, thus permitting the groups to be identified on the basis of their performance. Finally, the groups were identified as companies belonging to Grade A, Grade B and Grade C in that order, which exhibit the behavior of High performance, Moderate performance and Low performance.

Keywords: Data Mining, Financial Ratios, Factor Analysis, K-mean Clustering, Discriminant Analysis, Linear Trend.

1. INTRODUCTION

In the past two decades numerous statistical classification models have been constructed for financial related purposes. Statistical information on financial ratios are being used by the managers to place their companies performances in perspective. It is also well known that ratio analysis is the most powerful tool for financial statement analysis that are being extensively used by many researchers to meet their objective. These included models that have been developed to predict impending of corporate bankruptcy (Beaver, 1966; Altman, 1968); financial failure (Deakin, 1972; Booth, 1983); failing company (Blum, Marc, 1974); firm's performance (Bayldon, Woods and Zafiris, 1984) and forewarning indicators of corporate health (Prasant, Mishra and Satpathy, 1996). All these models typically link a set of "explanatory" variables to a "predictor" variable that can take two or more discrete values. Usually only two groups are considered in financial problem where the group information are known aprior. In all these models the operational objective is to assign the firm or company to one of the group after data analysis (for example bankruptcy Vs. non bankruptcy, etc,). With respect to grading companies (Manimannan G and R. Lakshmi Priya, 2020) where no assumptions are made with regard to the number of group or any other structural patterns in advance, to inherent groups or classes that reflected the performance of companies based on certain financial ratios.

The objectives of the present study are as follows:

1. To present three different methods of rating the top ranking companies on the basis of their performance using the concepts of data mining.
2. To identify the most appropriate proposed method of rating.

2. RESEARCH METHOD

This section is devoted to a discussion of the database, the ratios selected for the analysis and the Data Mining Techniques.

2.1. Database

The financial data considered for this study are published by *kaggle.com*, which covers the period 2015-2020 and combines public and private sector companies rated as best on their net sales in India. As it is either difficult to get the annual reports or to compare the financial, banking sectors and state sector corporations are meaningless, the publisher excluded them from the data. However, only top 500 companies are considered out of 1000 companies for the analysis for each study period. Among the listed companies, number of companies varied over the study period (*Table 1*) owing to amputation of those companies for which the required data are not available although they are published.

2.2. The Ratios

As ratios are simple and easy to understand, many researchers used them to analyse some of the aspects of the firms' financial condition and performances. However, the number of ratios that can be calculated from a typical set of financial statements is much large to in incorporate in this study. Moreover, due to availability of limited financial statement for company in the study only twenty financial ratios are sieved carefully out of numerous ratios that had different notions and meaningful interpretation. The different ratios computed are given in *Appendix*.

2.3. Data Mining Techniques

Data mining has been popularly treated as synonym to Knowledge Discovery in Databases (KDD). Although data mining is a new term, the technology is not. Researcher view data mining as the process of discovering previously unknown and potentially useful information such as patterns, associations and other significant structures from the data in databases. In general, a knowledge discovery process mainly consist of an iterative sequence of the following steps:

Step 1: Data Cleaning and Integration

Step 2: Data selection and transformation

Step 3: Data Mining

Step 4: Knowledge presentation

Mining also enables the company owners to determine the impacts of sales, customer's satisfaction and corporate profits to place their company performance in perspective. The data mining, and knowledge presentation processes are most important steps in mining process, which reveal new structural patterns present in the data. In the present context data mining exhibits the structural patterns by applying few techniques namely, factor analysis, **k**-means clustering and discriminant rule in Step 3. This structure discovers knowledge that is presented visually to the user, which is the final phase of data mining.

Table 1
Number of Companies In The Analysis Before And After Data Cleaning

Year	NUMBER OF COMPANIES	
	Before	After
2015	458	438
2016	491	466
2017	496	473
2018	399	349
2019	402	367
2020	414	383

2.3.1. Factor Analysis

Different factor analysis methods are used to test the stability of financial patterns over time. Although there are several techniques of data reduction, factor analysis is by far the most frequently used method in financial researchers (Mahmoud, Judith and Cecilio, 1987). Like all data reduction methods, factor analysis reduces the variable space under consideration to a smaller number of patterns that retain most of the information contained in the original data matrix. In the present context, principle component analysis is first initiated to ascertain the structural patterns through a linear combination of the financial ratios of companies. However, in factor extraction method the first *m* number of factors that explained 90% of variance are considered as knowledgeable. Both orthogonal rotations such as Varimax and Quartimax rotations are used to measure the similarity of a variable with a factor by its factor loading. In factor analysis, the interest is centered on the parameter in the factor model that estimated values of the common factor, called *factor scores.*These scores are subjected to further analysis to mine the data.

2.3.2. k-Means Clustering Algorithm

Many data mining applications make use of clustering techniques in classifications problems. In this study, a nonhierarchical clustering algorithm suggested by MacQueen (1967) also known as ***unsupervised classification***is well thought-of, as no presumption are made regarding the group structures present in the database. This process partition or group the data set into mutually exclusive group such that the members of each groups are as close as possible to one another and different groups are as far as possible from another. Generally this technique uses Euclidean distances measures computed by variables. The **k**-means clustering is one such technique in applied statistics that discovers acceptable classes. Thus forming the nuclei of clusters or groups as seed points exhibited in factor analysis. The number of cluster **k** is determined to be **3** as part of the clustering procedure, which had meaningful interpretations.

2.3.3. Discriminant Analysis

Multivariate Discriminant Analysis (MDA) have been extensively used by many researchers in financial problems with a prior group information for classification and model building. In the present study, iterative discriminant analysis is used to exhibit groups graphically and judge the nature of overall performance of the companies. This process re-allocated the companies that were assigned a group label by **k**-means clusters as a seed point. Re-allocation is subjected until cent percent classification is attained, by considering the classification of group obtained in iteration **t** as the input into the next iteration **t+1**. It is to be noted that the concept of performing repetitive DA is new in accessing the performance of the top rated companies in terms of net sales.

3. ALGORITHMS

A brief algorithm to prune the data during each of the study period to remove the outliers that could not be assigned to any of larger group is described below:

Step 1: Factor analysis is initiated to find the structural pattern underlying the data set and scores were extracted.

Step 2: **K** –means analysis partitioned the data set into **k**-clusters using factor scores as input matrix.

Step 3: Repeat Steps 1 and 2 until meaningful groups are obtained, by removing outliers in each cycle.

3. 1 METHOD – I

For the pruned data set (*Table 1*) the following algorithm is proposed under this method to grade the companies during the study period 1995-1999 based on their overall performances is described below:

Step 1: Discriminant analysis is performed with original ratios by considering the groups formed by **k** –means algorithm.

Step 2: Groups means are extracted for the year (**i + 2015)** by repeating **Step 1** from iteration **t** to the next iteration **(t+1)** for some **t** until cent percent classification is achieved. (**i = 0)**

Step 3:**K** –means analysis assigns initial labels to (**i + 2016)$^{\text{th}}$** years using group mean obtained from Step **2** as initial cluster centers.

Step 4: Repeat **Step 1** to **Step 3** for next **i**

3. 2 METHOD – II

Assuming that the rates of increase from one year to next in the financial ratios of companies are constant. A linear regression line is fitted to the final sorted group means of the pruned data sets, which are assigned initial classes by conventional **k** –mean analysis and then followed by iterative discriminant analysis. Hence, the proposed algorithm to grade the companies on the basis of overall performance from 1995 – 1999 in this method is describe below:

Step 1: Increment the corresponding ratios by a constant for (**i+2016)$^{\text{th}}$** group means respectively. (**i=0)**

Step 2:Mac Queen's **K** –means analysis assigns initial labels to (**i + 2016)$^{\text{th}}$** years using the group mean from Step **1** as initial cluster centers.

Step 3: Discriminant analysis in then performed for (**i+1995**) th year by considering the groups formed by **Step 2**

Step 4: Repeat **Step 3** until cent percent classification is achieved from iteration **t** to the next iteration (**t+1**) for some **t** and extract the group means.

Step 5: Repeat **Step 1** to **Step 3** for next **i**.

3. 3 METHOD – III

In this method, we assume that the performances of the companies over the years followed linear trend, to mine the data from 2015-2020in order to gauge the companies according to their performances. However, in this approach Step **1** and Step **5** are excluded from method **II**, Step **2** Step **3** and Step **4** are identical for mining. The initial cluster centers are estimated from the so fitted model.

4. RESULTS AND DISCUSSION

As discussed in Section 2.3.1 both Varimax and Quartimax criterion of orthogonal rotation have been used for the pruned data consolidated by general algorithm. The results obtained under both the methods are very similar but the varimax rotation provided relatively better clustering of financial ratios. Factor analysis revealed consistently five factors each year that explained 90 percent of total variation in the data with eigen values little less than or equal to unity. From this analysis we observed that the clustering of financial variables is stable during the study period. Although slight changes are encountered which are due to statistical variations in the original data.

After performing factor analysis, the next stage in data mining process is to assign initial group labels to each company followed by iterative discriminant analysis in succession by the three different suggested methods. Having decided to consider only 3 clusters as mentioned in section 2.3.2 with regard to previous research affirmed in section 1. Inspite of incorporating the results for each method for the study periods processed through the proposed algorithms, only the summary statistics are reported in *Table 2*. The first column in *Table 2* provides the groupings done by **k-**mean cluster analysis by three different methods except the year 2015. The second column gives the groupings after the application of discriminant analysis until 100 percent classification is achieved. Column three indicates the number of cycles required for convergence using different approaches.

Table 2
Number of companies in the clusters

Years	**Initial Cluster**			**Converged Discriminant**			**Number of Cycles**
	1	**2**	**3**	**1**	**2**	**3**	
2015	150	241	47	87	329	22	14
Method – I							
2016	24	76	336	23	104	339	10
2017	17	71	385	32	123	318	20
2018	15	49	285	14	51	283	3
2019	12	45	310	12	49	306	3
2020	8	12	338	10	65	308	8
Method –II							
2016	21	68	377	24	134	308	17
2017	7	66	400	22	110	341	13
2018	11	48	290	11	40	298	6
2019	10	47	310	13	47	307	6
2020	8	37	338	10	65	308	8
Method –III							
2016	10	53	403	24	144	298	22
2017	17	71	385	32	123	318	20
2018	21	60	268	19	62	268	5
2019	12	45	310	12	47	308	3
2020	8	37	338	10	65	308	8

1 – Grade **A** 2– Grade **B** 3 – Grade **C**

Table 2 indicates that majority of companies are in the low performance category. *Figure1* through *6*shows the groupings of companies into 3 clusters for each year of the study span by method- **I**. Similarly, *Figure7* through *12* and *Figure 13* through *18*depict the grouping of companies into 3 clusters by method- **II** and **III** respectively. From *Table 2* it is to be noted that all the three different methods classified the companies equivalently good. However, on comparing the performances of the different approach in terms of clustering the companies as data mining process, Method **-I** had slightly higher correct classification of 93.5% than the other two methods with 92.3 % and 93.4% respectively.

From the present study we also observed that the mean vectors of these clusters can be arranged in the increasing order of magnitude as show in *Table 3-7* for each of the study periods by all the approaches. Thus, permitting is to rate the members in the first cluster as Grade **A**, and the second as Grade **B** and the third as Grade **C**. Companies belonging to Grade **A** category are the ones that performs better than those of Grade **B** and Grade **C**. Similarly the companies belonging to Grade **B** category are superior to those of Grade **C**, indicating the members in the category Grade **C** are at a low profile in terms of the ratios considered in the present analysis. Also, Method **III** performancesand **I** are almost the same, therefore with any of these two methods could be used to gauge the companies performances in near future.

5. CONCLUSION

The purpose of this paper was to propose three different methods to identify the meaningful groups of companies that are rated as best with respect to their performance in terms of net sales using data mining techniques. We attempted to analysis the financial data relating to public and private sector companies over a period of six years from 2015 to 2020. Initially, factor analysis is used to identify the underlying structure in the 20 financial ratios. The factor scores are used to partition the companies into different clusters by using **k-**means clustering algorithm to prune the data.

In Method **I** successive year pruned data are mined using the previous year mean vectors as initial cluster center, keeping means vectors of 2015 as the base and so on. But the rest of the two methods mine the data by the suggested algorithm as in section 3.2 and 3.3. The unique feature in all these approach is the application of k-means as data mining tool to assign initial nuclei to cluster only once. Then they are followed by iterative discriminant analysis to re-allocate the members from iteration **t** to the next iteration **(t+1)**, until the process converges, that is, a member belonging to a cluster is assigned to itself.

The companies could be grouped only to 3 clusters for each year. The members of Cluster 1 are found to have high values for the financial ratios and hence they performed well. Thus, the members of Cluster 1 are labeled as Grade **A** companies. Similarly, the Cluster 2 included companies, which performed moderately well and the Cluster 3 with low-profile companies.

The present analysis has shown that only 3 groups could be meaningfully formed for each year. This indicates that only 3 types of companies existed over a period of six years. Further, the companies find themselves classified into *High* (Grade **A**), *Medium* (Grade **B**)and *Low* (Grade **C**) categories depending on the financial ratios. A generalization of the results is under investigation to obtain a unified class of 3 groups of companies for any given year.

Appendix

1.	Gross Profit / Net Sales	PBDT/NS
2.	Net Profit / Net Sales	PAT/NS
3.	Earning Before Interest and Tax /Total Assets	EBIT/NS
4.	Net Profit/Total Assets	PAT/A
5.	Net Profit before tax/Net Sales	PBT/NS
6.	Net Profit/Net Worth	PAT/NW
7.	Operating Profit/Net Sales	PBDIT/NS
8.	Operating Profit /Gross Sales	PBDIT/GS
9.	Gross Profit/Gross Sales	PBDT/GS
10.	Operating Profit/Total Assets	PBDIT/A
11.	Net Sales / Total Assets	NS/A
12.	Gross Profit / Total Assets	PBDT/A
13.	Cost of Sales/Net Sales	COGS/NS
14.	Cash Flow/Net Sales	CF/NS
15.	Cash Flow/Net Worth	CF/NW
16.	Net Worth/Net Sales	NW/NS
17.	Retained Earning/Total Assets	RP/A
18.	(Net profit/Net Worth) * (1 - Payout)	SGR
19.	Earning Before Interest and Tax/Interest	TIER
20.	Pay out ratio	PAY_OUT

Table 3: Centroids of Final Groups (2016)

Ratios	METHOD –I			METHOD -II			METHOD -III		
	A	B	C	A	B	C	A	B	C
PAY_OUT	.7564	.7482	.7146	.7609	.7609	.7074	.7475	.7262	.7127
PBDIT/GS	.1774	.1704	.1474	.1895	.1541	.1530	.1942	.1684	.1457
PBDT/GS	.1619	.1530	.0946	.1776	.1282	.0982	.1796	.1381	.0922
PAT/NS	.1214	.1067	.0684	.1185	.0891	.0724	.1327	.108	.0650
PBDIT/NS	.1922	.1887	.1656	.2153	.1714	.1698	.2391	.1889	.1594
NW/NS	.6083	.5397	.5252	.5942	.4807	.4758	.6049	.5576	.5391
CF/NS	.1120	.1086	.0822	.1079	.0937	.0863	.1167	.1040	.0803
COGS/NS	.8935	.8341	.8207	.8910	.8577	.7980	.8995	.8454	.7804
PBT/NS	.1628	.1389	.0751	.1839	.1180	.0762	.2032	.1269	.0687
PBDT/NS	.1793	.1659	.1065	.2019	.1422	.1089	.2195	.1545	.1005
PBDIT/A	.2249	.1808	.1346	.2913	.1793	.1253	.2794	.1855	.1215
PBDT/A	.2141	.1566	.0855	.2738	.1482	.0772	.2593	.1513	.0745
NS/A	1.166	1.126	1.039	1.502	1.247	.9727	1.338	1.336	.9337
PAT/A	.1439	.0942	.0539	.1627	.0898	.0502	.1594	.0927	.0477
PBIT/A	.2089	.1556	.1106	.2671	.1544	.1018	.2623	.1586	.0985
RP/A	.1118	.0715	.0406	.1223	.0691	.0375	.1186	.0707	.0360
PAT/NW	.3015	.2113	.1559	.3185	.2182	.1457	.2918	.2293	.1401
CF/NW	.2765	.2193	.1903	.2912	.2300	.1814	.2464	.2448	.1762
SGR	.0592	.0507	.0373	.0763	.0486	.0355	.0757	.0517	.0336
TIER	20.104	7.082	2.418	17.874	6.328	2.408	.0757	.0517	.0336

Table 4: Centroids of Final Groups (2017)

Ratios	METHOD –I			METHOD -II			METHOD -III		
	A	B	C	A	B	C	A	B	C
PAY_OUT	.2855	.2755	.2645	.2912	.2691	.2632	.2855	.2755	.2645
PBDIT/GS	.1563	.1519	.1442	.1862	.1578	.1438	.1563	.1519	.1442
PBDT/GS	.1306	.1258	.0982	.1577	.1517	.0889	.1306	.1258	.0982
PAT/NS	.0826	.0720	.0692	.1093	.0995	.0619	.0826	.0721	.0692
PBDIT/NS	.1714	.1656	.1540	.2017	.1684	.1581	.1713	.1657	.1540
NW/NS	.5849	.4351	.2685	.5405	.5278	.3940	.5849	.4351	.2684
CF/NS	.0869	.0867	.0687	.1909	.0980	.0772	.0869	.0867	.0687
COGS/NS	.8926	.8629	.8605	.9028	.8382	.8294	.8926	.8629	.8605
PBT/NS	.1202	.1147	.0726	.1471	.1399	.0657	.1202	.1147	.0726
PBDT/NS	.1395	.1370	.1074	.1706	.1617	.0972	.1395	.1370	.1074
PBDIT/A	.2792	.1832	.1240	.2517	.2085	.1244	.2892	.1832	.1241
PBDT/A	.2513	.1470	.0752	.2416	.1759	.0744	.2513	.1470	.0752
NS/A	2.024	1.571	.8605	1.694	1.303	1.029	2.024	1.571	.8605
PAT/A	.1287	.0905	.0481	.1560	.1023	.0465	.1287	.0905	.0481
PBIT/A	.2436	.1592	.0998	.2278	.1800	.1006	.2436	.1592	.0998
RP/A	.0871	.0688	.0366	.1147	.0773	.0347	.0871	.0688	.0366
PAT/NW	.2914	.2137	.1267	.2995	.2261	.1303	.2914	.2137	.1267
CF/NW	.2784	.2158	.1675	.2579	.2305	.1692	.2784	.2157	.1675
SGR	.2024	.1537	.0951	.2094	.1669	.0957	.2024	.1537	.0951
TIER	13.610	6.809	2.629	25.842	6.658	2.370	13.606	6.809	2.629

Table 5: Centroids of Final Groups (2018)

Ratios	METHOD –I			METHOD -II			METHOD -III		
	A	B	C	A	B	C	A	B	C
PAY_OUT	.2921	.2696	.2688	.2777	.2722	.2567	.2917	.2699	.2435
PBDIT/GS	.2195	.1529	.1528	.2171	.1663	.1518	.2167	.1661	.1487
PBDT/GS	.2073	.1335	.0885	.2066	.1482	.0895	.2032	.1402	.0833
PAT/NS	.1379	.0838	.0504	.1331	.0915	.05175	.1388	.0899	.0460
PBDIT/NS	.2325	.1674	.1672	.2266	.1768	.1669	.2307	.1805	.1632
NW/NS	.7701	.5733	.5558	.7155	.5745	.5713	.7154	.7009	.5406
CF/NS	.1465	.0886	.0715	.1463	.0974	.0718	.1460	.0946	.0682
COGS/NS	.9034	.8542	.7804	.9016	.8423	.7843	.9087	.8478	.7837
PBT/NS	.1810	.1182	.0593	.1719	.1297	.0617	.1817	.1210	.0542
PBDT/NS	.2196	.1457	.0966	.2157	.1577	.0984	.2163	.1521	.0912
PBDIT/A	.2298	.1907	.1320	.2333	.2040	.1335	.2406	.1796	.1298
PBDT/A	.2171	.1667	.0748	.2217	.1818	.0778	.2255	.1526	.0714
NS/A	1.304	1.271	.9903	1.315	1.264	1.000	1.327	1.163	.9974
PAT/A	.1288	.0923	.0391	.1305	.1003	.0410	.1362	.0864	.0363
PBIT/A	.1963	.1590	.1038	.1708	.1956	.1055	.2075	.1487	.1016
RP/A	.0932	.0673	.0276	.0966	.0744	.0287	.1034	.0624	.0252
PAT/NW	.2203	.1939	.0974	.2133	.2117	.1004	.2260	.1824	.0938
CF/NW	.2161	.2106	.1446	.2293	.2238	.1452	.2269	.1985	.1428
SGR	.1564	.1410	.0682	.1583	.1555	.0700	.1682	.1323	.0650
TIER	15.952	6.895	1.987	17.103	7.996	2.136	14.612	5.876	1.874

Table 6: Centroids of Final Groups (2019)

Ratios	METHOD –I			METHOD -II			METHOD -III		
	A	B	C	A	B	C	A	B	C
PAY_OUT	.2522	.2274	.2156	.2528	.2245	.2109	.2524	.2253	.2156
PBDIT/GS	.1940	.1697	.1505	.1976	.1659	.1504	.1928	.1697	.1509
PBDT/GS	.1738	.1622	.0803	.1772	.1584	.0803	.1722	.1622	.0811
PAT/NS	.1181	.1164	.0386	.1203	.1130	.0387	.1168	.1164	.0393
PBDIT/NS	.2107	.1849	.1649	.2142	.1816	.1648	.2098	.1849	.1654
NW/NS	.6691	.6685	.6037	.6858	.6388	.6027	.6773	.6691	.6029
CF/NS	.1274	.1184	.0666	.1300	.1159	.0667	.1255	.1184	.0673
COGS/NS	.9123	.8234	.8113	.9123	.8267	.8081	.9115	.8234	.8126
PBT/NS	.1551	.1495	.0457	.1576	.1462	.0459	.1545	.1495	.0466
PBDT/NS	.1886	.1766	.0876	.1918	.1732	.0877	.1874	.1766	.0885
PBDIT/A	.2178	.2074	.1175	.2167	.2134	.1178	.2128	.2074	.1189
PBDT/A	.1974	.1954	.0639	.2026	.1946	.0643	.1974	.1903	.0656
NS/A	1.418	1.206	.8785	1.464	1.177	.8814	1.419	1.191	.8829
PAT/A	.1268	.1192	.0297	.1281	.1190	.0300	.1268	.1153	.0309
PBIT/A	.1820	.1807	.0889	.1847	.1810	.0892	.1808	.1785	.0900
RP/A	.0974	.0934	.0199	.0993	.0936	.0201	.0974	.0902	.0209
PAT/NW	.2399	.2172	.0767	.2416	.2161	.0772	.2399	.2143	.0781
CF/NW	.2360	.2288	.1322	.2358	.2355	.1324	.2332	.2288	.1333
SGR	.1806	.1687	.0516	.1837	.1686	.0518	.1806	.1667	.0526
TIER	19.420	8.071	1.841	18.502	8.064	1.849	19.042	8.319	1.844

Table 7: Centroids of Final Groups (2020)

Ratios	METHOD –I, II & III		
	A	B	C
PAY_OUT	.4815	.2634	.2056
PBDIT/GS	.2402	.1508	.1368
PBDT/GS	.2147	.1435	.0667
PAT/NS	.1276	.0999	.0230
PBDIT/NS	.2507	.1644	.1500
NW/NS	.7468	.5794	.3329
CF/NS	.1540	.0743	.0569
COGS/NS	9273	.8432	.7759
PBT/NS	.1720	.1361	.0266
PBDT/NS	.2240	.1567	.0726
PBDIT/A	.3076	.2141	.1114
PBDT/A	.2930	.1917	.0581
NS/A	1.928	1.141	.9141
PAT/A	.1886	.1088	.0228
PBIT/A	.2702	.1692	.0793
RP/A	.1023	.0870	.0134
PAT/NW	.3118	.2011	.0431
CF/NW	.2502	.2274	.1181
SGR	.1619	.1604	.0194
TIER	27.758	12.410	1.934

Clustered Groups (METHOD – I)

Figure: 1 Year 2015

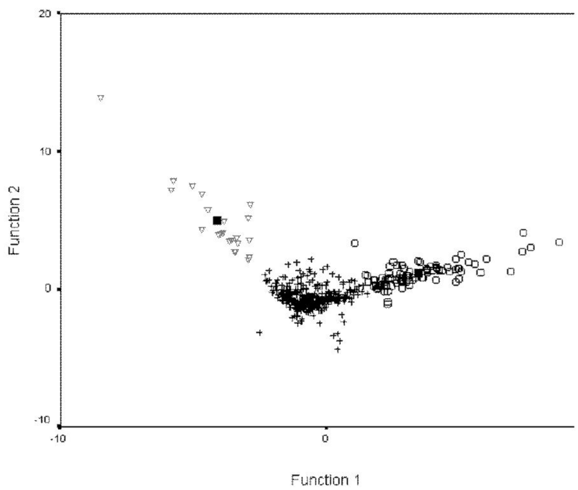


Figure: 2 Year 2016

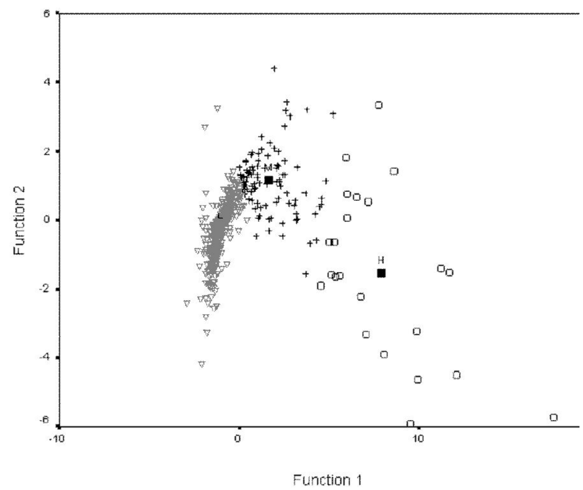


Figure: 3 Year 2017

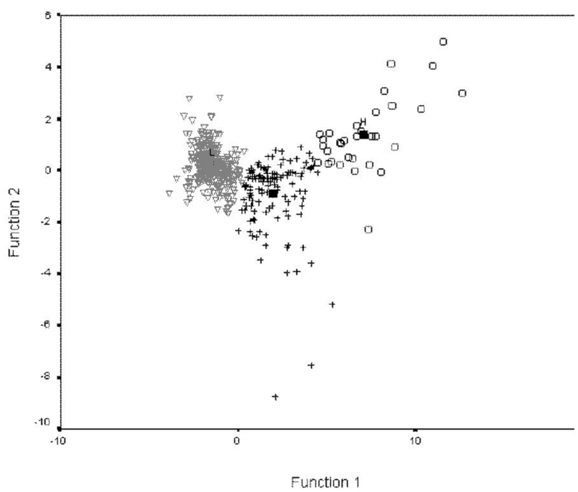


Figure: 4 Year 2018

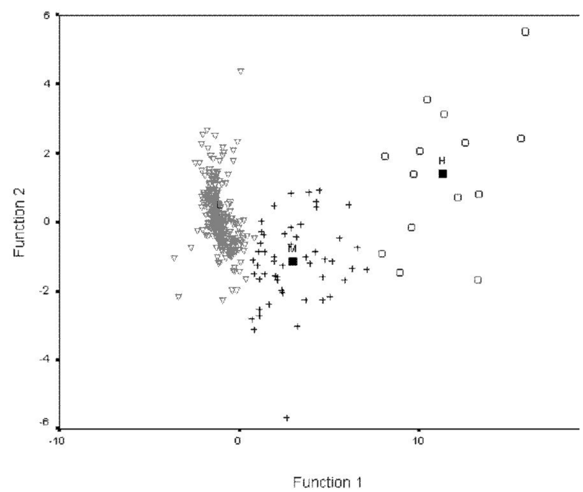


Figure: 5 Year 2019

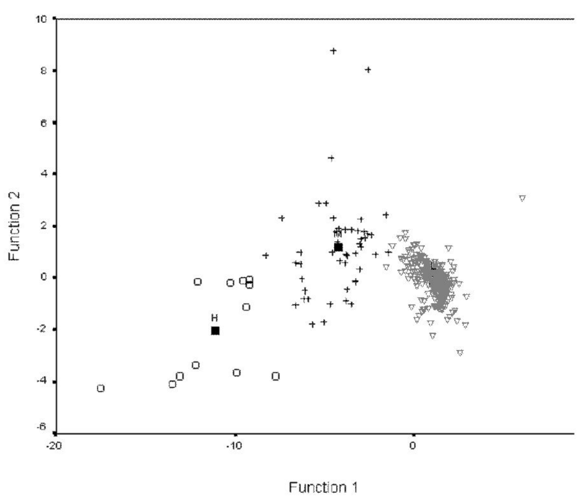
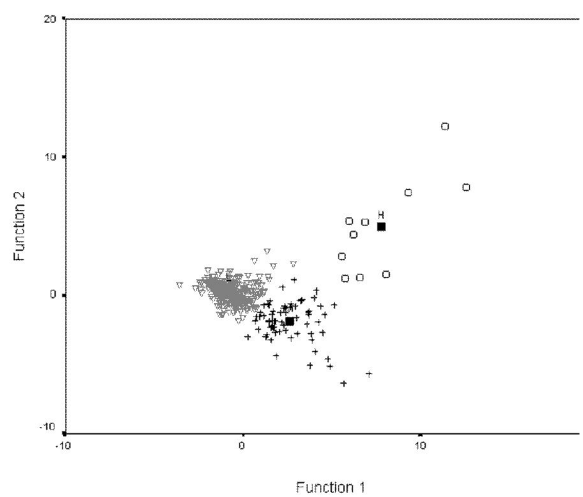


Figure: 6 Year 2020



Cluster 1 (Grade **A**) + Cluster 2 (Grade **B**) Cluster 3 (Grade **C**) Cluster centroids

Clustered Groups (METHOD – III)

Figure: 13 Year 2015

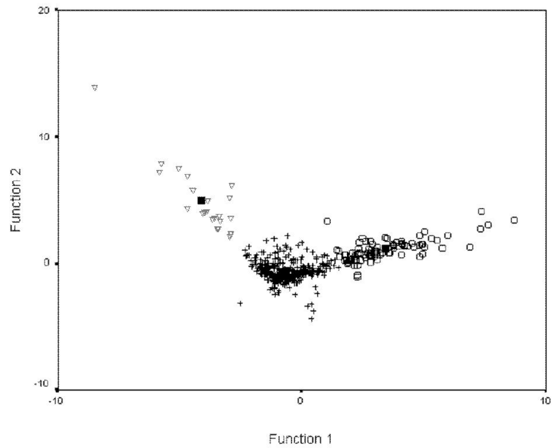


Figure: 14 Year 2016

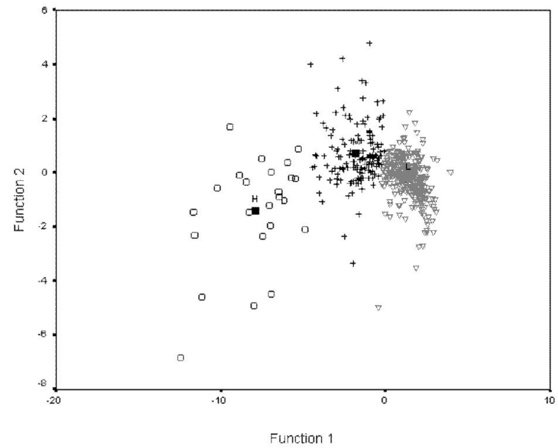


Figure: 15 Year 2017

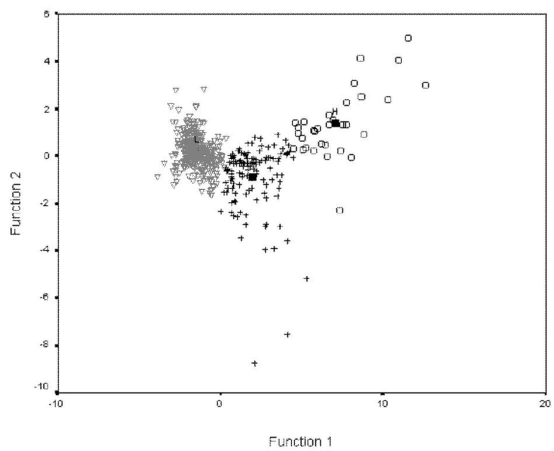


Figure: 16 Year 2018

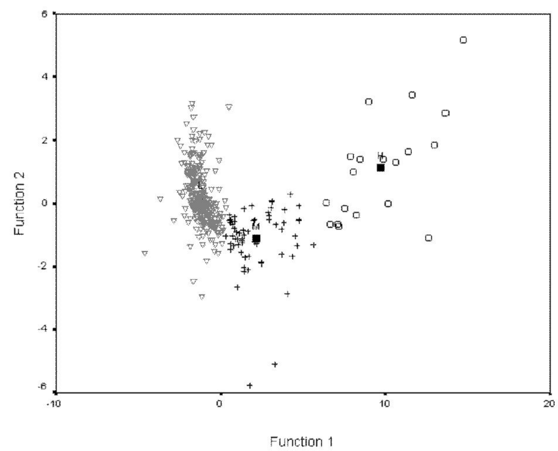


Figure: 17 Year 2019

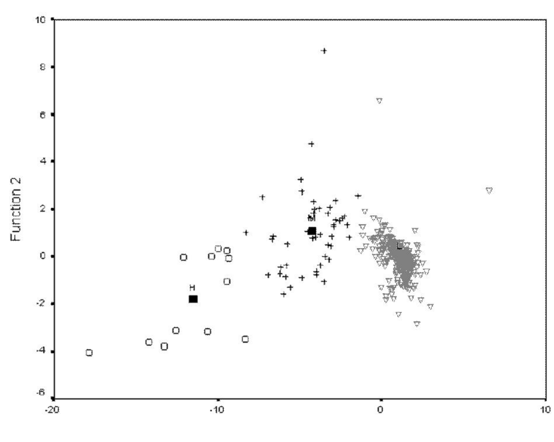
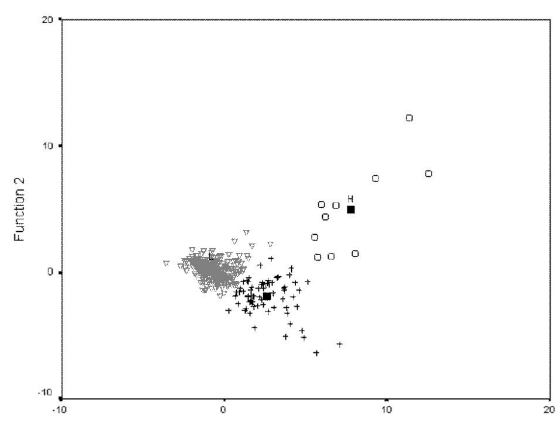


Figure: 18 Year 2020



Cluster 1 (Grade **A**) + Cluster 2 (Grade **B**) Cluster 3 (Grade **C**) Cluster centroids

REFERENCES

1. *Altman E I (1968), Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, The Journal of Finance, vol.23, pp. 589-609*
2. *Anderson T W (1984), An Introduction to Multivariate Statistical Analyis, 2/e, John Wiley and Sons, Inc., New York*
3. *Breaver W H (1966), Financial Ratios as Predictors of Failure, Journal of Accounting Research, pp. 179-192*
4. *Blum M. (1974), Failing Company Discriminant Analysis, Journal of Accounting Research, pp. 179-190.*
5. *Manimannan G. and R. Lakshmi Priya (2020), Assessing Indian Industries On The Basis Of Financial Ratios, Lambert Academic Publishing .Republic Of Moldova.*
6. *David Wishart (2001), K-Mean Clustering with Outlier Detection Mixed Variables and Misssing Values, www. clustan.com*
7. *David Wishart (1998), Efficiency hierarchical cluster analysis for data mining and knowledge discovery, Computing Science and Statistics, pp. 257-263*
8. *Deakin E. B (1972), A Discriminant Analysis of Predictor of Financial Failure, Journal of Accounting Research, pp. 167 - 179.*
9. *Everitt (1980), Cluster Analysis, Halsted Press, Division of John Wiley and Sons, New York*
10. *Pieter Adriaans and DolfZantinge (1996), Data Mining, Addison-Wesly Longman Limited, England*
11. *Prasanna Chandra (1997), Financial Management Theory and Practice, 4/ed, Tata McGraw-Hill Publication Company Limited, New Delhi*

DOI No. : 10.5958/2394-9309.2022.00002.6
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

PREDICTING HIV-INFECTED PATIENTS' LIFE TIMES THROUGH A STOCHASTIC APPROACH

Subash Chandrabose. G*, Vinoth Raman, Vignesh. T*** Sudha.V*****

*Lecture in statistics, Department of Community Medicine, AVMC&H, Pondicherry.
**Assistant Professor, Deanship of Quality and Academic Accreditation,
Imam Abdulrahman Bin Faisal University, Kingdom of Saudi Arabia
***Assistant Professor, Department of Community Medicine, AVMC&H, Pondicherry.
E-mail: subashstat@gmail.com

ABSTRACT

In modern culture, public health issues are extremely essential, particularly for any virus that plays a significant role in the spread of infectious diseases. HIV is another cause of death for which scientists are currently searching for a cure. Because of the unique nature of HIV infection, a stochastic model was derived from a theoretical framework of sound epidemiology with specific assumptions. The shock model concept is used in this research to analyze the threshold level using a stochastic technique based on a three-parameter Exponentiated Weibull distribution. Secondary data was used to observe the model fit, which will be used to support the model's development in a real-world scenario.

1. INTRODUCTION

The current researchers aim to determine how the external transmission of the semen of HIV-infected patients differs at different stages of the disease. The actual infection varies from place to place. Therefore, current research plans to calculate the mean. In addition, the extent of infection depends on the gender type. Some are higher compared to others.

A mathematical model was obtained for the expected breakdown time to the servo conversion threshold level. In the HIV/AIDS servo transformation, the hypothesis was that the time between contact intervals randomly distributed the same independent random variables.

distribution family Exponentiated Weibull (EWD) was first proposed by Mudholkar and Srivastava (1993) to cover a family exponential one parameter. The Family Exponential is a sub-family and the Weibull two-parameter family is best used as a special sub-family.

The EWD family allows for non-repetitive hazard functions. which is one particular example In order to provide more detail about the time expected to cross the threshold level of the time interval. seroconversion We can refer to Esary et.al., (1973), Elangovan et.al., (2009), Pandiyan et.al., (2015).

Factors responsible for damage after the immune system are affected include ELISA- HBE + VE, ELISA- HBE –VE, WESTERN TEST, CD4 + T CELL TEST, etc. Current research attempts to fit the original data into the resulting model. Received as expected

2. ASSUMPTIONS OF THE MODEL

- Sexual contacts are the only source of HIV infection person.
- The threshold of any individual is a random variable.
- If the total damage crosses a threshold level Y which itself is a random variable, the seroconversion occurs and a person is recognized as an infected.
- The interarrival time between the successive contacts is at random variable which are identically.

3. MODEL DESCRIPTION

The Cumulative density function (CDF) of the three parameter Exponentiated Weibull distribution

$$F(x,\lambda) = \left[1 - e^{-\left(\frac{x}{\lambda}\right)^k}\right]^\alpha \quad ; \quad x > 0$$

The corresponding survival function is

$$\overline{\mathrm{H}}(\mathrm{x}) = \mathrm{e}^{-\left(\frac{\mathrm{x}}{\lambda}\right)^{\mathrm{k}}} \tag{1}$$

Y: Continuous random variable denoting the threshold level of three parameter Exponentiated Weibull distribution.

$$P(X_i < Y) = \int_0^\infty g^*(x)\overline{H}(x)dx \tag{2}$$

S(t) : the survivor function i.e $P(T > t)$

$$P(T > t) = \sum_{k=0}^{\infty} V_k(t)P(X_i < Y) \tag{3}$$

Taking Laplace Transformation of the life time $= 1 - S(t)$ we get Let the random variable U denoting inter arrival time which follows exponential with parameter c. Now $f^*(s) = \left(\frac{c}{c+s}\right)$, substituting in the above equation (4) we get

$$l^*(s) = \frac{\left[1 - g^*\left(\frac{1}{\sigma}\right)^\beta\right] f^*(s)}{\left[1 - g^*\left(\frac{1}{\sigma}\right)^\beta f^*(s)\right]} \tag{4}$$

Let the random variable U denoting inter arrival time which follows exponential with parameter. Now $f^*(s) = \left(\frac{c}{c+s}\right)$, substituting in the above equation we get

$$E(T) = -\frac{d}{ds} l^*(s) \qquad given \quad s = 0$$

$$= \frac{1}{c\left[1 - g^*\left(\frac{1}{\lambda}\right)^{\beta}\right]} \tag{6}$$

Then

$$E(T) = \frac{\mu\lambda^k + 1}{c} \tag{7}$$

Where

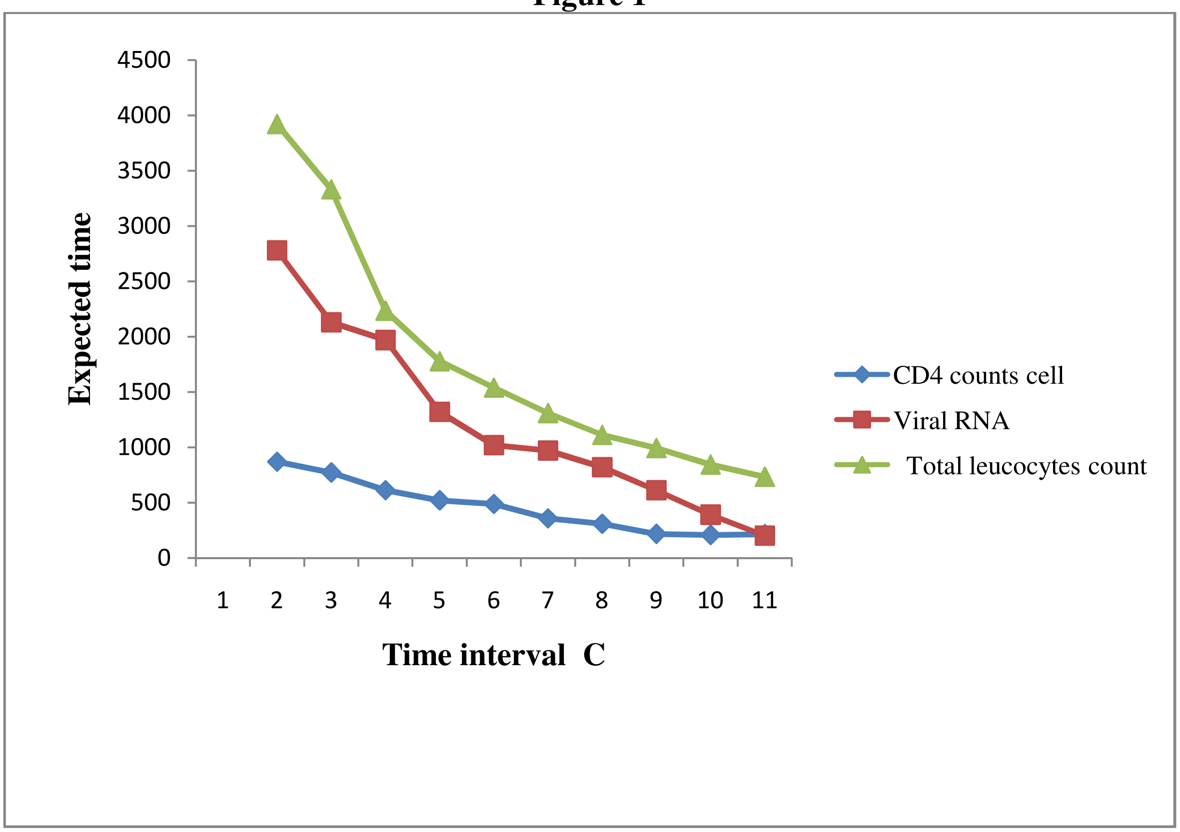
c=Time interval

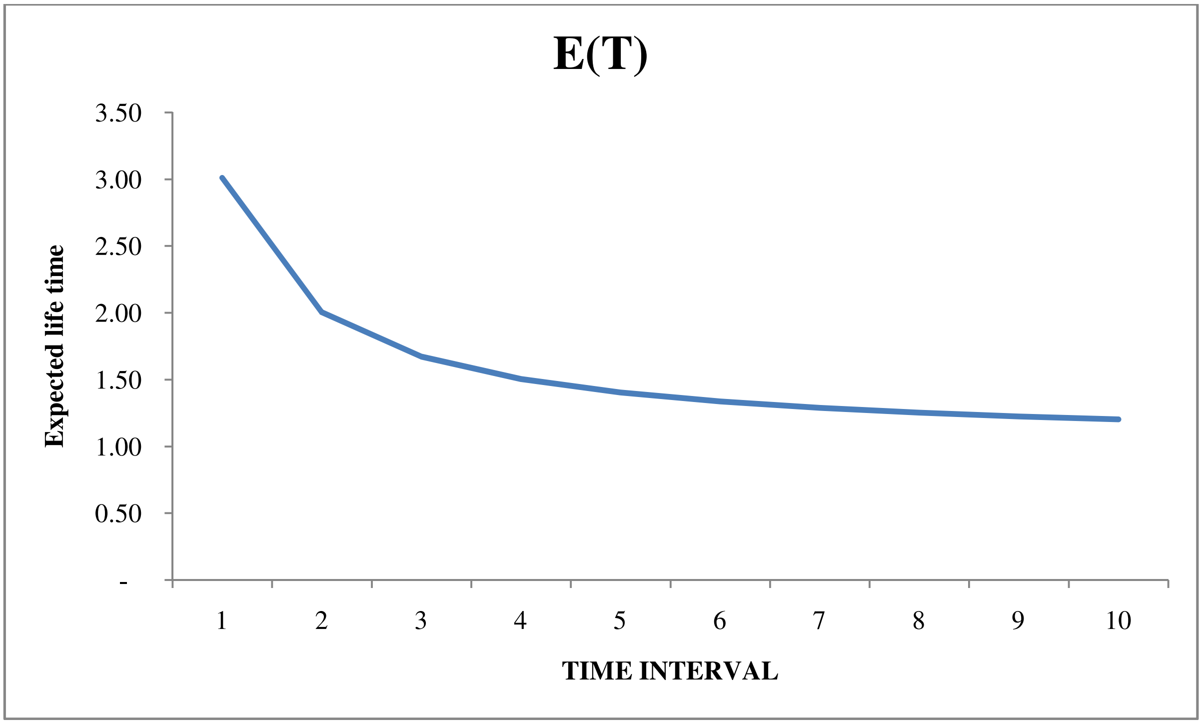
μ = CD4 counts cell

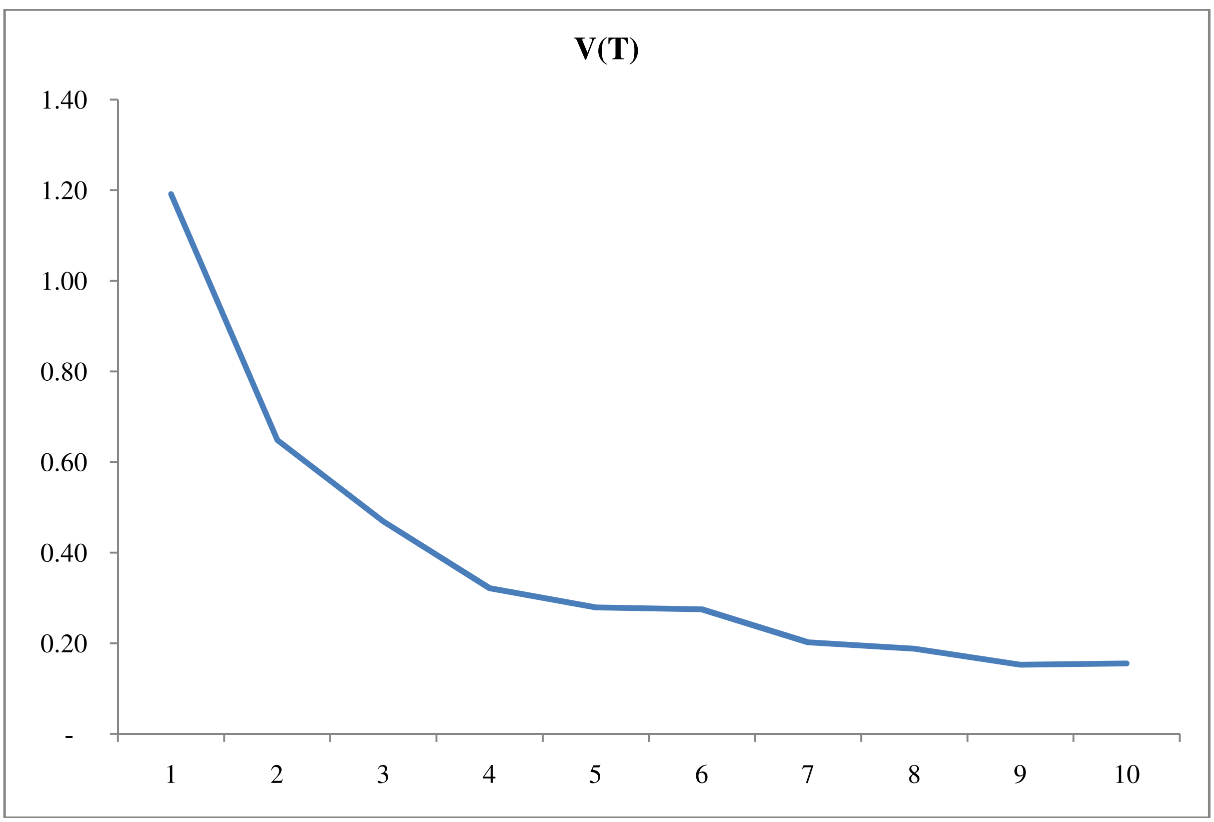
λ= Viral RNA

k= Total leucocyted count

Figure 1




E(T)
Expected life time
3.50
3.00
2.50
2.00
1.50
1.00
0.50
-
1
2
3
4
5
6
7
8
9
10
TIME INTERVAL


V(T)
1.40
1.20
1.00
0.80
0.60
0.40
0.20
-
1
2
3
4
5
6
7
8
9
10

5. CONCLUSION

The current researchers from the above study concluded that human immunodeficiency virus infections crossed the threshold more quickly. When a person is infected, immune cells are damaged and are more likely to become infected when infected with an immune virus.

Note that if parameter c of the exponential distribution represents an increase in arrival time between contacts. The expected time is reduced if the CD4 Count cell, RNA, Total leucocytes count parameters of the random variable indicate damage to seroconversion increased then the expected time was also found to decrease.

The time interval for infection depends on the duration of the sexual contact of the infected person. There is a sudden drop in the spam threshold level in the person's expected life. This model was clearly observed when a person was infected with HIV. the immune system is damaged The model also shows that in the case of infected individuals The immune system will begin to break down. If the number of sexual contacts increases after the threshold level The expected time of HIV patients are reduced As observed in the figure above. The life expectancy of the affected person can be extended with proper medical advice and through regular use of medications.

REFERENCE

1. ***Esary, J.D, A.W.Marshall and F. Proschan,****(1973). "Shock models and wear processes". Ann. Probability, 1(4), pp.627-649.*
2. ***Wodarz, D and M.A., Nowak, (2002).*** *Mathematical models of HIV pathogenesis and treatment. Bio Essays, Vol. 24, pp.1178-1187.*
3. ***Purushottam A Giri, Jayant D Deshpande, and Deepak B Phalke (2013).*** *Prevalence of Pulmonary Tuberculosis Among HIV Positive Patients Attending Antiretroviral Therapy Clinic, N Am J Med Sci. 2013 Jun; 5(6): 367–370.*

4. ***Pandiyan.P, A. Loganathan, Vinoth. R and Kannadasan.K, (2013),*** *Calculating the survival time of HIV patient through three parameter exponentiatedweibull distribution, Antarctica Journal of Mathematics, Vol.10, No.5, pp.479-487.*

5. ***G.S.Mudholkar and D.K.Srivastava, (1993),*** *"ExponentiatedWeibull family for analyzing bath-tab failure-rare data", IEEE Trans, Reliability, Vol.42, pp.299-302.*

DOI No. : 10.5958/2394-9309.2022.00003.8
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ON A UNIFIED SUBCLASS OF ANALYTIC UNIVALENT FUNCTIONS

S. Sunil Varma* and Thomas Rosy**
Department of Mathematics, Madras Christian College (Autonomous)
Tambaram, Chennai – 600059, Tamil Nadu, India

ABSTRACT

In this paper, we introduce a new unified subclass of analytic univalent functions defined in the open unit disk in the complex plane and having Taylor's series expansion about the origin. Necessary and sufficient conditions for this subclass is proved. Geometric properties like bounds for coefficients, growth and distortion theorems for functions in this subclass are obtained. As an application, sufficient conditions for Hypergeometric function with various normalization to be in this subclass are derived.

Mathematics Subject Classifcation 2020 : 30C45

Keywords: Analytic functions, Univalent functions, Geometric properties, Hypergeometric functions.

1. INTRODUCTION

Let $\Delta = \{z \in \mathbb{C} : |z| < 1\}$ denote the open unit disk in the complex plane $\mathbb{C}$. Denote by $\mathcal{A}$ the class of all analytic functions defined on the open unit disk Δ with the normalization $f(0) = 0 = f'(0) - 1$. Such functions will have Taylor's series expansion of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n \ . \qquad (1)$$

Let $\mathcal{S}$ denote the subclass of $\mathcal{A}$ consisting of univalent functions. Denote by $\mathcal{T}$ the subclass of $\mathcal{S}$ consisting of functions of the form

$$f(z) = z - \sum_{n=2}^{\infty} |a_n| z^n \ . \qquad (2)$$

the subclass of $\mathcal{S}$ containing functions with negative coefficients **[4].**

A new subclass of starlike and convex subclasses of univalent functions were introduced and studied by Rosy *et.al.* in **[2,3].**Motivated by their works, we introduce and study the following subclass.

Definition 1. A function $f \in S$ is said to be in the subclass $USCD(\alpha, \beta, \gamma)$ if

$$\mathcal{R}e\left(\gamma f'(z) + (1-\gamma)\frac{f(z)}{z}\right) - \alpha\left|\gamma z f''(z) + (1-\gamma)(f'(z) - \frac{f(z)}{z}\right| > \beta,$$

where $0 \le \gamma \le 1, 0 \le \alpha \le 1, 0 \le \beta < 1$.

The class $USCD(\alpha, \beta, \gamma)$ reduces to various subclasses of $\mathcal{S}$ for different values of the parameters α, β, γ.

Remarks:

(i) When $\gamma = 1$, the above subclass reduces to $UCD(\alpha, \beta)$ studied by T.Rosy *et.al.* **[2].**

(ii) When $\gamma = 0$, the above subclass reduces to $SD(\alpha, \beta)$ studied by T.Rosy *et.al.* **[2].**

We let $\mathcal{T}USCD(\alpha, \beta, \gamma) = \mathcal{T} \cap USCD(\alpha, \beta, \gamma)$.

2. PROPERTIES OF THE SUBCLASS $\mathcal{T}USCD(\alpha, \beta, \gamma)$

Theorem 1: Let $f$be a function of the form (1) in the class$S$. If

$$\sum_{n=2}^{\infty}[\alpha\gamma n^2 + (\alpha + (1 - 2\alpha)\gamma)n + (1 - \alpha)(1 - \gamma)]|a_n| \le 1 - \beta$$

then $f \in USCD(\alpha, \beta, \gamma)$.

Proof:$\mathcal{R}e\left(\gamma f'(z) + (1 - \gamma)\frac{f(z)}{z}\right) - \alpha\left|\gamma z f''(z) + (1 - \gamma)(f'(z) - \frac{f(z)}{z}\right| - \beta$

$$\ge 1 - \beta - \gamma|f'(z) - 1| - \alpha\gamma|zf''(z)| - (1 - \gamma)\left|\frac{f(z)}{z} - 1\right| - \alpha(1 - \gamma)\left|f'(z) - \frac{f(z)}{z}\right|$$

$$= 1 - \beta - \gamma\sum_{n=2}^{\infty} n|a_n| - \alpha\gamma\sum_{n=2}^{\infty} n(n - 1)\,|a_n| - (1 - \gamma)\sum_{n=2}^{\infty}|a_n| - \alpha(1 - \gamma)\sum_{n=2}^{\infty}(n - 1)|a_n|$$

$$= 1 - \beta - \sum_{n=2}^{\infty}[\alpha\gamma n^2 + (\alpha + (1 - 2\alpha)\gamma)n + (1 - \alpha)(1 - \gamma)]|a_n| \ge 0.$$

Hence, $f \in USCD(\alpha, \beta, \gamma)$.

Theorem 2:A function f of the form (2) is in the class $\mathcal{T}USCD(\alpha, \beta, \gamma)$ if and only if

$$\sum_{n=2}^{\infty}[\alpha\gamma n^2 + (\alpha + (1 - 2\alpha)\gamma)n + (1 - \alpha)(1 - \gamma)]|a_n| \le 1 - \beta. \tag{3}$$

Proof: By the virtue of aboveresult, it is enough to show that if $f \in \mathcal{T}USCD(\alpha, \beta, \gamma)$, then (3) holds.

For $f \in \mathcal{T}USCD(\alpha, \beta, \gamma)$,

$\mathcal{R}e\left(\gamma f'(z) + (1 - \gamma)\frac{f(z)}{z}\right) - \alpha\left|\gamma z f''(z) + (1 - \gamma)(f'(z) - \frac{f(z)}{z}\right| > \beta,$

$\Longrightarrow \mathcal{R}e(1-\sum_{n=2}^{\infty}[(n-1)\gamma+1]|a_n|)-\alpha\sum_{n=2}^{\infty}[-\gamma n^2+(2\gamma-1)n+1-\gamma]|a_n|>\beta.$

Allowing $z \longrightarrow 1$ through the real values of z, we obtain

$$\sum_{n=2}^{\infty}[\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)]|a_n| \leq 1-\beta.$$

Theorem 3: Let a function f of the form (2) be in the class $\mathcal{T}USCD(\alpha,\beta,\gamma)$. Then

$$|\mathrm{a_n}| \leq \frac{1-\beta}{\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)}\ , n \geq 2.$$

Theorem 4: Let a function f of the form (2) be in the subclass $\mathcal{T}USCD(\alpha,\beta,\gamma)$. Then

$$|\mathrm{z}|-\frac{1-\beta}{(1+\gamma)(1+\alpha)}|z|^2 \leq |f(z)| \leq |\mathrm{z}|+\frac{1-\beta}{(1+\gamma)(1+\alpha)}|z|^2,\ |z|<1.$$

Proof: If f of the form (2) is in the subclass $\mathcal{T}USCD(\alpha,\beta,\gamma)$, then

$$\sum_{n=2}^{\infty}|a_n| \leq \frac{1-\beta}{(1+\alpha)(1+\gamma)}. \quad (4)$$

Also, if f is of the form (2), then

$$|z|-|z|^2\sum_{n=2}^{\infty}|a_n| \leq |f(z)| \leq |z|+|z|-|z|^2\sum_{n=2}^{\infty}|a_n| \quad (5)$$

Using (4) in(5), we obtain the result.

Theorem 5: Let a function f of the form (2) be in the subclass $\mathcal{T}USCD(\alpha,\beta,\gamma)$. Then

$$1-\frac{2(1-\beta)}{(1+\gamma)(1+\alpha)}|z| \leq |f'(z)| \leq 1+\frac{2(1-\beta)}{(1+\gamma)(1+\alpha)}|z|,\ |z|<1.$$

Proof: If f of the form (2) is in the subclass $\mathcal{T}USCD(\alpha,\beta,\gamma)$, then

$$\sum_{n=2}^{\infty}n|a_n| \leq \frac{2(1-\beta)}{(1+\alpha)(1+\gamma)}. \quad (6)$$

Also, if f is of the form (2), then

$$1-|z|\sum_{n=2}^{\infty}n|a_n| \leq |f'(z)| \leq 1+|z|\sum_{n=2}^{\infty}n|a_n| \quad (7)$$

Using (6) in (7), we obtain the result.

3. APPLICATIONS

Motivated by the works in **[1,5],** we now obtain certain sufficiency condition for certain modified form of hypergeometric function and Hovlov operator to be in the subclass $USCD(\alpha,\beta,\gamma)$.

Theorem 6: Let $a,b \in \mathbb{C}-\{0\}$ and $c>|a|+|b|+2$. Then a sufficient condition for the function $zF(a,b;c;z)$ to be in the subclass $USCD(\alpha,\beta,\gamma)$ is that

$$\frac{\Gamma(c-|a|-|b|)\Gamma(c)}{\Gamma(c-|a|)\Gamma(c-|b|)}\left[\frac{(|a|)_2(|b|)_2}{c-2-|a|-|b|}\alpha\gamma+\frac{(\alpha+\gamma+\alpha\gamma)|a||b|}{c-1-|a|-|b|}+1\right]\leq 2-\beta. \tag{8}$$

Proof: Consider $zF(a,b;c;z)=z+\sum_{n=2}^{\infty}\frac{(|a|)_{n-1}(|b|)_{n-1}}{(|c|)_{n-1}(1)_{n-1}}z^n$.

Here$a_n=\frac{(|a|)_{n-1}(|b|)_{n-1}}{(|c|)_{n-1}(1)_{n-1}}$. Let

$$T:=\sum_{n=2}^{\infty}[\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)]|a_n|$$

Then,

$$T:=\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]|a_{n+1}|$$

$T{:=}\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]\left|\frac{(a)_n(b)_n}{(c)_n(1)_n}\right|$

It is enough to show that $T\leq 1-\beta$.

Using $|(a)_n|\leq(|a|)_n$, we have

$$T\leq\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]\frac{(|a|)_n(|b|)_n}{(|c|)_n(1)_n}$$

$=\alpha\gamma\sum_{n=1}^{\infty}\frac{(n+1)^2(|a|)_n(|b|)_n}{(|c|)_n(1)_n}+(\alpha+(1-2\alpha)\gamma)\sum_{n=1}^{\infty}\frac{(n+1)(|a|)_n(|b|)_n}{(|c|)_n(1)_n}+(1-\alpha)(1-\gamma)\sum_{n=1}^{\infty}\frac{(|a|)_n(|b|)_n}{(|c|)_n(1)_n}$

$=\alpha\gamma\sum_{n=0}^{\infty}\frac{(n+1)^2(|a|)_n(|b|)_n}{(|c|)_n(1)_n}+(\alpha+(1-2\alpha)\gamma)\sum_{n=0}^{\infty}\frac{(n+1)(|a|)_n(|b|)_n}{(|c|)_n(1)_n}+$

$$(1-\alpha)(1-\gamma)\sum_{n=0}^{\infty}\frac{(|a|)_n(|b|)_n}{(|c|)_n(1)_n}-1$$

$=\frac{\Gamma(c-|a|-|b|)\Gamma(c)}{\Gamma(c-|a|)\Gamma(c-|b|)}\left[\frac{(|a|)_2(|b|)_2}{c-2-|a|-|b|}\alpha\gamma+\frac{(\alpha+\gamma+\alpha\gamma)|a||b|}{c-1-|a|-|b|}+1\right]$-1

$<1-\beta$by(8).

This completes the proof.

Theorem 7: Let $a,b\in\mathbb{C}-\{0\}$ and $c>|a|+|b|+3$ and $f\in S$. Then a sufficient condition for the function $I_{a,b,c}(f)$= $zF(a,b;c;z)*f(z)$to be in the subclass $USCD(\alpha,\beta,\gamma)$ is that

$$\frac{\Gamma(c-|a|-|b|)\Gamma(c)}{\Gamma(c-|a|)\Gamma(c-|b|)}\left[\frac{\alpha\gamma(|a|)_3(|b|)_3}{(c-3-|a|-|b|)_3}+\frac{(\alpha+\gamma+4\alpha\gamma)(|a|)_2(|b|)_2}{(c-2-|a|-|b|)_2}+\frac{(6\alpha\gamma+2\alpha+2\gamma+1)|a||b|}{c-1-|a|-|b|}\right]$$
$$\leq 2-\beta. \tag{9}$$

Proof: For each$f \in S, I_{a,b,c}(f)(z) = zF(a,b;c;z) * f(z) = z + \sum_{n=1}^{\infty}\frac{(a)_n(b)_n}{(c)_n(1)_n}a_{n+1}z^{n+1}$.

Let $A_n = \frac{(a)_{n-1}(b)_{n-1}}{(c)_{n-1}(1)_{n-1}}a_n$.

Therefore

$$\sum_{n=2}^{\infty}[\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)]|A_n|$$
$$=\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]|A_{n+1}|$$
$$=\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]\frac{(a)_n\ (b)_n}{(c)_n(1)_n}|a_{n+1}| \tag{9}$$

Since $f(z) = z + \sum_{n=2}^{\infty}a_nz^n \in S, |a_n| \leq n,\ \forall n \geq 2$. Using this in (9),

$$\sum_{n=2}^{\infty}[\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)]|A_n|$$
$$\leq \sum_{n=2}^{\infty}[\alpha\gamma n^2+(\alpha+(1-2\alpha)\gamma)n+(1-\alpha)(1-\gamma)]\frac{(|a|)_n\ (|b|)_n}{(|c|)_n(1)_n}(\mathrm{n+1})$$
$$=\sum_{n=1}^{\infty}[\alpha\gamma(n+1)^2+(\alpha+(1-2\alpha)\gamma)(n+1)+(1-\alpha)(1-\gamma)]\frac{(|a|)_n\ (|b|)_n}{(|c|)_n(1)_n}(n+1)$$
$$=\alpha\gamma\sum_{n=1}^{\infty}(n+1)^3\frac{(|a|)_n\ (|b|)_n}{(|c|)_n(1)_n}+(\alpha+(1-2\alpha)\gamma)\sum_{n=1}^{\infty}(n+1)^2\frac{(|a|)_n\ (|b|)_n}{(c)_n(1)_n}+$$
$$(1-\alpha)(1-\gamma)\sum_{n=1}^{\infty}(n+1)\frac{(|a|)_n\ (|b|)_n}{(c)_n(1)_n}$$
$$=\frac{\Gamma(c-|a|-|b|)\Gamma(c)}{\Gamma(c-|a|)\Gamma(c-|b|)}\left[\frac{\alpha\gamma(|a|)_3(|b|)_3}{(c-3-|a|-|b|)_3}+\frac{(\alpha+\gamma+4\alpha\gamma)(|a|)_2(|b|)_2}{(c-2-|a|-|b|)_2}+\frac{(6\alpha\gamma+2\alpha+2\gamma+1)|a||b|}{c-1-|a|-|b|}\right]-1$$

$< 1-\beta$, by (9).

CONCLUSION

In this paper, we have defined a new unified subclass of analytic univalent functions and examined its properties. As an application, we have obtained sufficient conditions for Hypergeometric functions and Hovlov operator to be in the subclass considered.

REFERENCES

1. *Y.C.Kim and S.Ponnusamy*, ***"****Sufficiency Gaussian hypergeometric functions to be uniformly convex", Int. J. Math. And Math.Sci., Vol 22, No.4. (1999), 765 – 773.*
2. *T.Rosy, " Studies on subclasses of starlike and convex functions", Ph.D. Thesis (Submitted to University of Madras), 2001.*
3. *T.Rosy, B.A.Stephen and K.G.Subramanian, "Classes of convex functions",Int. J. Math. And Math.Sci., Vol 23, No.12. (2000), 819 – 825.*
4. *H.Silverman, "Univalent functions with negative coefficients", Proc. Of. The. Amer.Math,soc., Vol 51, No.1 (1975), 109 – 116.*
5. *A.Swaminathan, "Sufficient conditions for hypergeometric functions to be in a certain class of analytic functions", Comp. and Math. With Applns., Vol. 59 (2010), 1578 – 1583.*

DOI No. : 10.5958/2394-9309.2022.00004.X
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A THEORY OF CONSTRAINTS APPROACH USING SIMULATION IN SAMPLING INSPECTION FOR V-TYPE PRODUCTION PLANTS

S. Ravi Sankar, Assistant Professor, and D. Maheshwari, Research Scholar
*Department of Statistics, Government Arts College, Coimbatore, Tamil Nadu, India.
E-mail id : maheswari.statistics@gmail.com

ABSTRACT

In this paper, an attempt is made to improve the sampling inspection forV-type production plants through Theory of Constraints approach using simulation.The major four elements ofthe Theory of Constraintsnamely,throughput,inventory, operating expenses andconstraintsare considered as the major components of the simulation model and are applied, using Goldsim (12.0), a simulation software, in sampling inspection for the V-type production plants.

Keywords: Acceptance Sampling, Theory of constraints, V-type Production Plants, Simulation Modeling.

1. INTRODUCTION

A sampling plan is a statistical approach for determining whether a large number of materials should be accepted or rejected. Acceptance sampling is carried out whileeither the raw material enters into the production unit, or at any stage of semi-finishedgoods or product leaves the production plants. It is usually used when testing is destructive; the cost of 100% inspection is very high; and 100% inspection takes too long.In a sampling inspection plan, the lot is accepted or rejected based on its samplingproperties. The different types of sampling plans are used and compared in respect of the cost involvedin V-typeproduction plants.

Goldratt (1980) developed the concept of Theory of Constraints(TOC).It is often used by various industrial companies for the continuous improvement of their processes to increase efficiency. The goal of any industrial organization is to make profit. TOC commonly uses financial measures, goal of making money, profit, net profit,Return on Investment (ROI) and the cost involved.

Throughput (T) is the rate at which a system produces the target units. Money is the objective unit, and throughput is thedifference between the Sales (S) and raw material cost (C).i e., T= S – C. Throughput exists only when a product or service is sold.

Investment(I) is the money tied up in the system with the inventory, machinery, buildings, other assets and liabilities. Operating Expense (OE) is the money in which system spends in generating “goal units”. For physical products, OE is the expenses,except the cost of the raw materials it includes maintenance, rent, taxes, etc.,

These three measures are very powerful and decisions taken by managers in order to improve the profit. Improving a measure is increasing, and decreasing for Throughput, Investment and Operating Expenses respectively.

The TOC tools help to manage through the process of implementing a project. This process on going improvement, have been applied to manufacturing, project management, supply chain that generated specific solutions. The TOC applications in the fields of marketing, sales and finance.

The production plants for VATI analysis using V-type production plants are discussed in raw materials and working stations as well as the Theory of Constraints technique for sampling plans in the context of simulation software systems. We can improve throughput and cost effectiveness by locating sampling inspections plans.

2. REVIEW OF LITERATURE

Single sampling plans are widely used for appraising the product quality. In such situations, it is logical to have a comparison between various sampling plans.

Peachand Littauer (1946) have given a method for finding acceptance number ‘c’ of the single sampling plan involving minimum consumer’s risk and producer’s risk.Gloub (1953) has given a method for finding the acceptance number ‘c’ for a single sampling plan involving minimum sum of producer’s risk and consumer’s risk. Guenther (1969) evolved procedure for Single sampling plan with Binomial, Hypergeometric and Poisson models. Stephans(1981) suggested aprocedure for finding the sample size and acceptance number for single sampling plan used for Binomial Model. Wasserman (1990) proposed an operating procedure for a chain sampling plan with a single long proportion sampling plan at the level of two target operations. Radhakrishnanand RaviSankar (2009)have proposed the selection of three classes attribute single sampling plan for the given Average Outgoing Quality Level.

Goldratt andJacobs (1980) havescheduled algorithm for industrial production known as Optimized Production Schedule. The throughput, significant reduction in inventory, defects associated with quality, cycle time and lead time respectively.

He has published "The Goal"(1984) and it illustrates American Production Plants introduces the principles of Theory of constraints. Another book was published by RobertFox's(1986)"The Race" book. The TOC tools, techniques and the implementation of the production process. This theory became a tool for continuous improvement focused on processes (1990), be complemented by the use of logic and Thinking Processes (TPs) as the bases for the resolution of the focused.He defines TOC indicators as the regarding performance measurement systems, and based on the assumption that the goal of any business is to generate money.

3. TYPES OF PRODUCTION PLANTS

In the TOC model, there are four types of production plants which specify the flow of materials through the operation. The VATI analysis uses the bottom-up shapes of the letters V, A, T, and I to characterise the plant types.

V-plant: The raw material is one-to-many, a plant can take one raw material and turn it into a variety of final products. One procedure, immediately after a diverging point material meant for the other operations, is the complication of V-plants. Once the material has been processed, it cannot be returned and reprocessed without extensive rework.

A-plant: The raw material is one-to-many, numerous sub-assemblies converge in a plant to form a final assembly. Converging lines in A-plants are an issue because they must all supply the final assembly point at the same time.

T-plant: The T-plant is made up of several assembly. The majority of parts produced are used in several assemblies. For example, Computers are good instances of customised devices. T-plants are affected by both A-plant synchronisation issues and V-plant robbing issues.

I-plant: In an assembly line, this material flows in a specific order. The I-plant is carried out in a sequential manner. The slowest operations are the constraints.

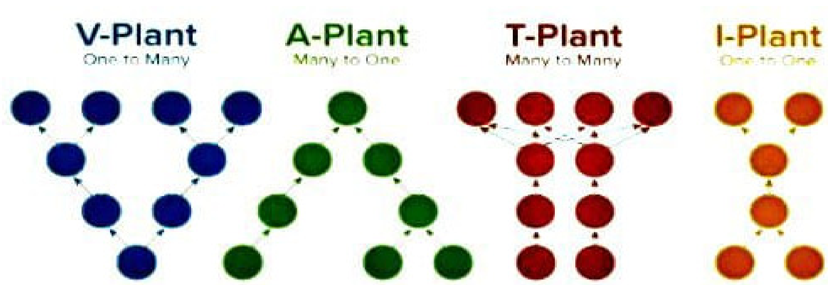


Figure 3.1: Types of Production Plants for VATI Analysis

4. RESEARCH METHODOLOGY

In this section,TOC model isevaluated using the basic simulation model to examine the effect of TOC on throughput level. This model is developed forthe V-type production plants to make management decisionon Return on Investment. Simulationsoftware can help to determine the drum buffer in the TOC to reduce the costs so as to satisfy throughput.

The main goal of TOC management production system is to increase the profit, Return on Investment and throughputby maximizing the productivity, and by reducingthe operating expenses.

Using sampling inspection technique,in production plants of V-type for single sampling plan. To find the proportion defective values and the Product A, B, C and Return on Investment is used a simulation software.

The production plants of system that uses a simulation model in V-type plants to evaluate the performance and the VATI analysis that can help the scheduling problems and plant types describe the raw material flow throughout the factory.

In this study, the simulation software, Goldsim is used to design the models for the V-type production plants.This implementation go under process and different work stations with different capacity drum buffers, drum beats and drum ropes. After this process,the sampling inspectionreduces the defective items to improve the quality of the products.

Goldsim(12.0) isa simulation software for carrying out dynamic, probabilistic simulation to support management and decision-making in engineering, science and business. It is highly-graphical one that runs on personal computers using windows operating system and it can be a powerful tool for understanding and managing complex systems. Simulation is defined as the process of a model of an existing of proposed system in order to identify and to understand those factors control the system to predict the future behaviour.

Goldsim is a powerful and flexible platform for visualizing and dynamically simulating in production plants for business. It is like a ‘visual spreadsheet’ that create graphs, manipulate data and equations. This system can be quantitively described using equations or rules can be simulated. In this study,Goldsim is used for the realistic simulation of discrete events such as financial transactions and amount of resources utilized.

5. Development of Simulation Model for V-type Production Plants

A simulation modelfor V-type production plants is developed and is presented in thefigure 5.1. Suppose, in a V-type production process, it is planned to produce three Products A, B and C which are to be processed through four different common work stations WC1, WC2, WC3, and WC4. Further, the

Product A is to be passed through WCA5, the Product B is to be passed through WCB2 and WCB3 and the Product C is to be passed through WCC5 also.

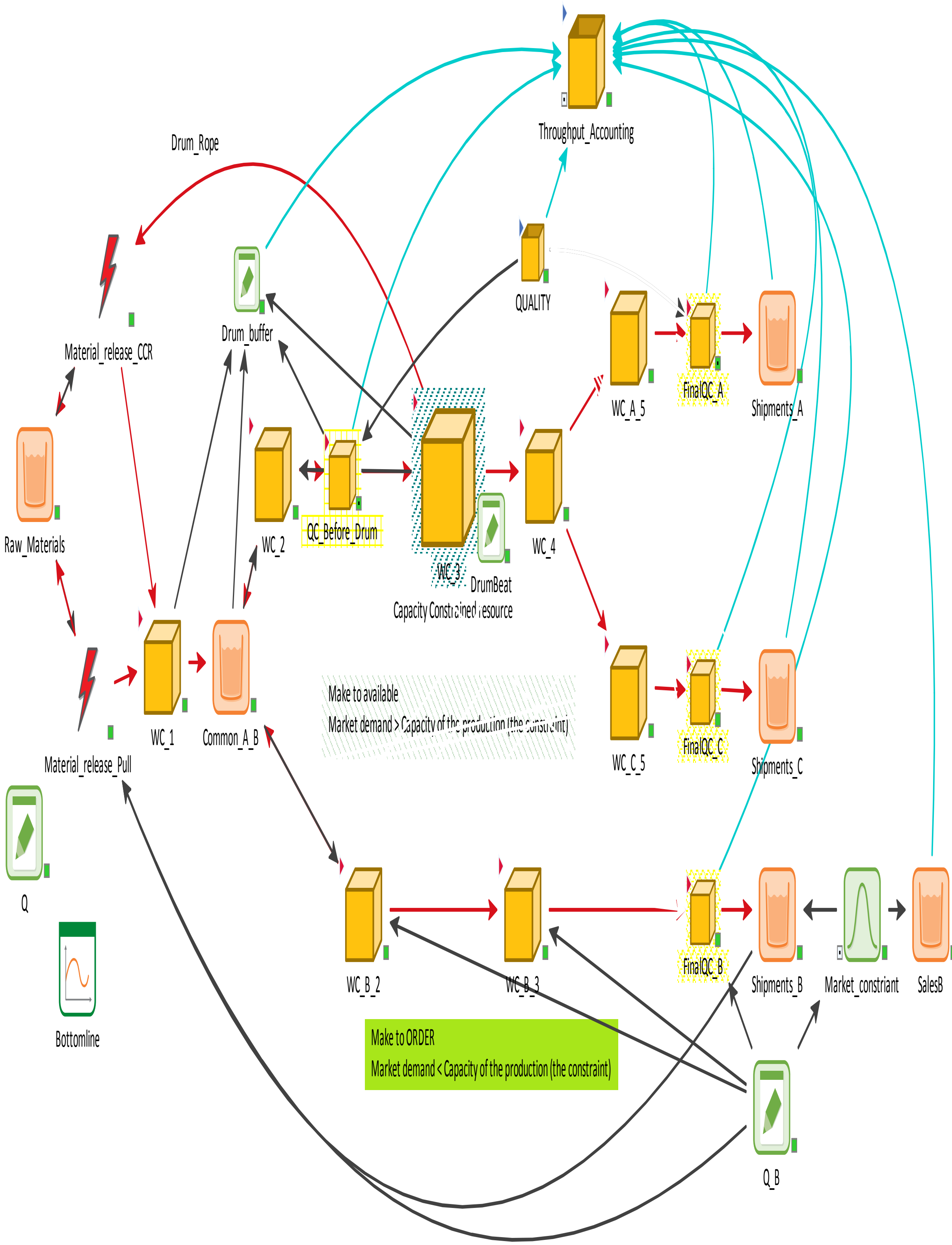


Figure 5.1 Simulation Model for the V- type Production Process

In the figure 5.1, a simulation model for V-type production plants is presented along withsampling inspection plan. The simulation model has different work stations with various capacities. This model has the work stationsWC1, WC2, WC3, WCA4, WCA5,WCB2 and WCB3. Also,this model involves Drum buffers, Drum Beats and Drum Rope to improve the quality of the Products A, B and C. After purchasing, the raw material is processed through WC1 and WC2. The quality is inspected before going to WC3. WC3 is a main constraint resource. It is a mandatory to check the product A and B before it moves to WC4.

The WC4 has two main branches viz., WCA5 and WCC5. After processing throughWCA5, finally the good quality Product A is checked and now ready for the Shipment A. Directly the ProductA moves to the throughput. After processing in WCC5,the quality of the ProductC is checked at the final stage. If there is no defect in the ProductC, the best quality product is moved to ShipmentC and directly ProductC will move to the throughput. Again, thepurchased raw material has to be released from the material release pull andis moved to the ShipmentB. When the raw material makes to market demand the market,there is another constrained resource. The Marketis a mandatory for the sales B and the product directly moves to the throughput. From thework station WC1, moved the product for a drumcommon for the Products A and B. The material which is common to the Products A and B moves to WCB2 and WCB3 after the quality of the products is checked, it is moved to ShipmentB.From there it moves to the market constraint and to sales B move to the throughput.

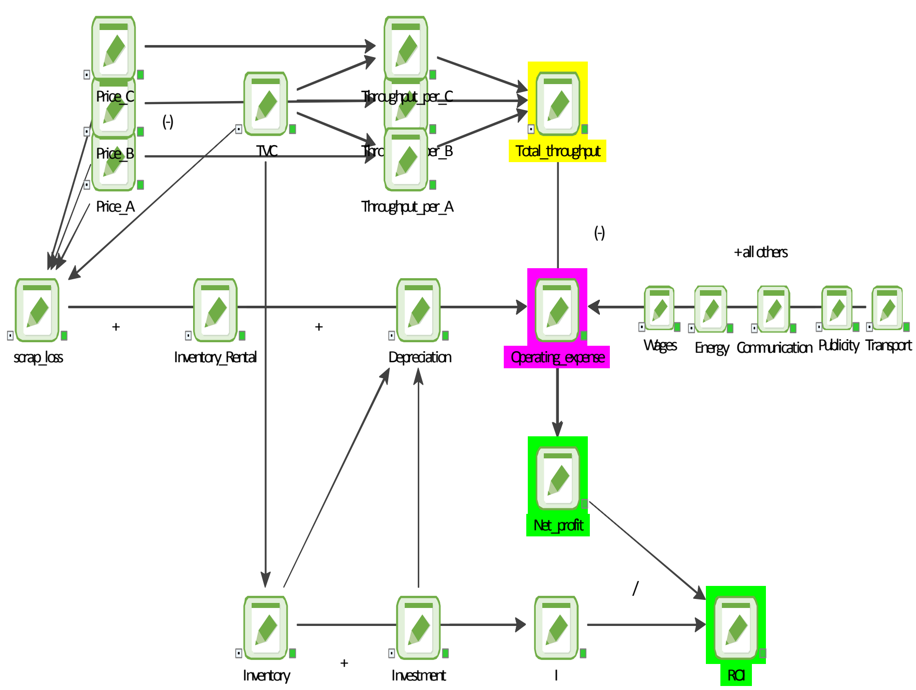


Figure5.2:Throughput Model for the V-type Plants

The figure 5.2showsthroughput model for the V-type plants having three Prices,viz.,- Price A,Price B,and Price C with their respective Throughputs viz., -Throughput A,Throughput B, and Throughput C. Total Variable Cost(TVC) can be arrivedusing prices and throughput values. TVC is the difference between prices and the throughput values. Twoseparate input elements are operating expenses and inventory. Operating expenses generate scrap loss, inventory rental, depreciation,wages, energy, etc.The inventory moves to Investment and Return on Investment.

The ROI values (in percentages), for various p values, the proportion defective values,are calculated and presented in Table 5.1 by applying single sampling plan with $(n_{1,}c_1)$, $(n_{2,}c_2)$, $(n_{3,}c_3)$ and(n_4, c_4) being the pairs of sample size and acceptance number for the Products A,B and C respectively of lot size 100.

Single sampling plan is implemented when the underlying quality features distributed with fixed parameters. For example, from the belowTable 5.1, in a production process for the proportion defective p = 0.05and the lot size=100, one can get the SSP (91,3), (82,1), (76,10), (81,7) with ROIas 159.46%.Similarly, for the proportion defective p = 0.20,&the lotsize100, one can get the SSP (67,0), (10,9), (40,3), (51,1) with ROIas 146.484%. The ROI decreases, from table 5.1, it is observed that when the proportion defective value increases.

Table 5.1:ROI% using single sampling plan with$(n_{1,}c_1)$, $(n_{2,}c_2)$, $(n_{3,}c_3)$,$(n_{4,}c_4)$ for the lot size of 100.

	Constraint		Product A		Product B		Product C		
p	n_1	c_1	n_2	c_2	n_3	c_3	n_4	c_4	ROI %
0.01	69	0	100	0	59	1	55	1	160.699
0.02	83	1	92	5	71	1	94	1	159.881
0.03	23	0	100	1	69	1	93	4	159.827
0.04	91	3	82	1	76	10	81	7	159.461
0.05	91	3	82	1	76	10	81	7	159.461
0.06	77	4	54	1	57	6	69	6	158.593
0.07	100	0	100	0	40	8	82	1	157.456
0.08	33	2	41	1	78	5	61	7	157.761
0.09	40	3	57	2	81	8	59	5	157.293
0.1	100	0	18	0	49	0	40	4	155.835
0.12	38	4	90	5	62	8	86	7	155.981
0.14	74	10	100	7	40	9	85	10	154.942
0.16	41	6	56	4	48	8	67	5	154.298
0.18	79	0	41	3	68	5	53	6	148.86
0.2	67	0	10	9	40	3	51	1	146.484
0.22	47	0	62	5	67	5	61	3	141.707
0.24	100	0	16	0	40	5	40	0	141.417
0.26	84	0	33	5	60	6	55	1	138.922
0.28	78	0	14	4	84	2	40	2	138.357
0.3	91	1	42	0	65	2	52	8	135.833
0.4	86	0	80	0	44	6	100	1	123.747
0.5	100	5	31		82	8	51	3	113.733

6. CONCLUSION

In this study, a table having single sampling plans withROI for various proportion defective (p) is constructed for easy selection of a suitable single sampling plan.The quality of the products, to attain the profit based on throughput,is being checked before shipment of all the three Products A,B and C. The idea regarding production scheduling and the utilization of the available capacity will help the production managers to think differently about how to utilize the available resources properly.

This study can further be extended by considering other type of production plantsviz., A-type plants, T-type plants, and I- type plants. The constraints suchas market, labour and finance are given more importance than machine constraints. It is further concluded that the Theory of Constraints would maximize the overall profit by minimizing cost through the constraints.

REFERENCES:

1. *Golub, A. (1953). Designing single-sampling inspection plans when the sample size is fixed. Journal of the American Statistical Association, 48(262), 278-288.*
2. *Goldratt, E. M., & Cox, J. (1984). The goal: excellence in manufacturing. North River Press.*
3. *Goldratt, E.M. and Fox, J. (1986). The Race, North River Press, New York.*
4. *Goldratt, E., (1980). The Goal. New York: North River Press.*
5. *Guenther, W. C. (1969). Use of the Binomial, Hypergeometric and Poisson Tables to obtain Sampling Plans, Journal of Quality Technology, 1(2).*
6. *Goldratt, E., Jacobs, F.R. (1980). The OPT Scheduling System: A review of a new production scheduling system. P&IM. J., 24(3), 47-51.*
7. *Peach, P. and Littauer, S. B. (1946). A note on sampling inspection. The Annals of Mathematical Statistics, 17(1), 81- 84.*
8. *Radhakrishnan, and RaviSankar, S. (2009). Single Sampling Plan for Three Attribute Classes Indexed Through Acceptable Quality Level. International Journal of Statistics and Management System, 4(1), 150-164.*
9. *Stephens, K.S. (1981). CSP -1 for Consumer Protection, Journal of Quality Technology, 13(4), 249-253.*
10. *Wasserman, G. S. (1990). Matching performance of continuous sampling plans with single sampling plans. Communications in Statistics-Simulation and Computation, 19(4), 1303-1317.*
11. *Software@goldsim.com*

DOI No. : 10.5958/2394-9309.2022.00005.1
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A CASE STUDYOF COVID-19 VACCINATION: FIRST AND SECOND DOSE OF INDIA

Renjini Harish* and Manimannan G**
*Teacher, Department of Mathematics , The Emirates National School, Sharjah, United Arab Emirates
**Assistant Professor, Department of Mathematics, Apollo College of arts and Science Chennai

ABSTRACT

This case study is attempted to identify COVID-19 vaccination of first and second dos of India. The data were collected from 16th January 2021 to 6th July 2021 and it is based on the population as of Census 2011. The database features are Total Vaccine administered population, gendered based features, age group based featured and dose based features are considered for this study. India began its vaccination programme on 16 January 2021, operating 3,006 vaccination centres on the onset.

This case study is to identify the types of vaccine administered in India where it signifies the three main vaccines incubated in India - COVISHIELD, COVAXIN and Sputnik V. The visualization of graph shows that India's population has been highly immunized by COVISHIELD, moderately by COVAXIN and least by Sputnik V. The first and second dose vaccine administered in India where it signifies that the people of India are being highly vaccinated for first dose and are least vaccinated for the second dose. The first dose helps your body create an immune response, while the second dose is a booster that strengthens your immunity to the virus.

In different age group of vaccine administered in India where it signifies that the Indian population ranging from age 18-45 years are being highly inoculated, population ranging from age 45-60 years are moderately inoculated and population ranging from 60 years and above are least inoculated by Covid-19 Vaccine. For many older adults and those with on-going conditions like heart disease and diabetes, the vaccine can prevent severe illness or death from the corona virus. It also helps the body producing antibodies to the corona virus. These antibodies help your immune system fight the virus if you happen to be exposed, so it reduces your chance of getting the disease. Vaccine administered with different gender in India where it signifies that male individuals of the country have been highly injected, female individuals are moderately injected and transgender individuals are least injected with Covid-19 vaccine.In addition, based on the database to visualize in simple bar and line diagrams using MS-Excel.

Keywords: Vaccination, COVAXIN, COVISHIELD, First Dose and Second Dose, Descriptive Analysis

1. INTRODUCTION

COVID-19 is the disease caused by new corona virus called Severe Acute Respiratory Syndrome-Corona virus-2 (SARS-CoV-2). The World Health Organisation first learned of this new virus on 31 December, 2019, following a report of a cluster of cases of ‘viral pneumonia’ in Wuhan, People’s Republic of China. This is a new corona virus that has not been previously identified in humans.

Corona viruses are a large family of viruses that are known to cause illness ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS)[1]. The main aim of this study is to classify the COVID-19 vaccinated people of India on the basis of dosage, vaccine, age and gender.

1.1 CAUSES

Infection with the new corona virus – Severe Acute Respiratory Syndrome-Corona virus-2 (SARS-CoV-2) causes corona virus disease 2019 (COVID-19).The virus that causes COVID-19 spreads easily among people, and more continues to be discovered over time about how it spreads. Data has shown that it spreads mainly from person to person among those in close contact (within about 6 feet, or 2 meters). The virus spreads by respiratory droplets released when someone with the virus coughs, sneezes, breathes, sings or talks. These droplets can be inhaled or land in the mouth, nose or eyes of a person nearby.

In some situations, the COVID-19 virus can spread by a person being exposed to small droplets or aerosols that stay in the air for several minutes or hours — called airborne transmission. It's not yet known how common it is for the virus to spread this way.It can also spread if a person touches a surface or object with the virus on it and then touches his or her mouth, nose or eyes, but the risk is low.Some reinfections of the virus that causes COVID-19 have happened, but these have been uncommon.

1.2 SYMPTOMS

The most common symptoms of COVID-19 are:

- ✓ Fever
- ✓ Dry cough
- ✓ Fatigue

Symptoms of severe COVID-19 disease include:

- ✓ Shortness of breath
- ✓ Loss of appetite
- ✓ Confusion
- ✓ Persistent pain or pressure in the chest
- ✓ High body temperature (above 38°C)

1.3RISK FACTORS

Risk factors for COVID-19 appear to include:

- ✓ Close contact (within 6 feet, or 2 meters) with someone who has COVID-19.
- ✓ Being coughed or sneezed by an infected person.

1.4 COMPLICATIONS

Although most people with COVID-19 have mild to moderate symptoms, the disease can cause severe medical complications and lead to death in some people. Older adults or people with existing medical conditions are at greater risk of becoming seriously ill with COVID-19.

Complications can include:

- ✓ Pneumonia and trouble breathing
- ✓ Organ failure in several organs
- ✓ Heart problems
- ✓ A severe lung condition that causes a low amount of oxygen to go through your bloodstream of your organs (acute respiratory distress syndrome)
- ✓ Blood clots
- ✓ Acute kidney injury
- ✓ Additional viral and bacterial infections [3]

2. THE IMMUNE SYSTEM

The immune system is the body's defence mechanism, protecting against invaders like bacteria and viruses to keep us healthy.

Cells are the main building blocks of our body. Our immune system relies on many different types of cells, each playing an important role. Many of these can be found in our bloodstream, especially white blood cells, which are the main component of the human immune system.

2.1IMMUNISATION

The purpose of immunisation is to prevent people from getting sick. It helps to protect people against the complications of becoming ill, including developing chronic diseases, cancer, and death.

2.2VACCINATION

Vaccines work by stimulating the body's defence mechanisms to provide protection against infection and illness. These defence mechanisms are collectively referred to as the immune system. Vaccines mimic and sometimes improve the protective response normally mounted by the immune system after infection. The great advantage of immunisation over natural infections is that immunisation has a much lower risk of harmful outcomes [8].

2.2.1 Vaccination is disease specific

A healthy immune system can generate hundreds of millions of T and B cells, each of which targets one particular antigen.

However, pathogens can sometimes overwhelm the immune response. Vaccines give the immune system a head start by allowing it to learn and remember what a pathogen looks like, providing valuable protection against aggressive pathogens.

These immune responses are very specific, so we need to have a separate vaccine for each disease. The immune system can respond independently to each pathogen it encounters. This is why the system cannot be 'overloaded' or damaged by giving the full range of currently available vaccines or by having multiple antigens in one vaccine.

2.2.2 Vaccines work with the immune system's ability to remember

Most of the cells involved in immune responses live for only a few days, but a small number of lymphocytes survive for months or years after the infection has been cleared away. These lymphocytes either continue to produce antibodies or retain a 'memory' of the invading pathogen. In case of measles, for example, that memory has been shown to last for more than 60 years [2].

2.3HERD IMMUNITY

'Herd immunity', also known as population immunity is the indirect protection from an infectious disease that happens when a population is immune either through vaccination or through immunity developed through previous infection.

World Health Organization supports achieving 'herd immunity' through vaccination, not by allowing a disease to spread through any segment of the population as this would result in unnecessary cases and deaths.

To safely achieve herd immunity against COVID-19, a substantial proportion of population would need to be vaccinated, lowering the overall amount of virus able to spread in the whole population.

The percentage of people who need to be immune in order to achieve herd immunity varies with each disease. For example: herd immunity against Measles requires about 95% of a population to be vaccinated. The remaining 5% will be protected by the fact that Measles will not spread among those who have vaccinated. For Polio, the threshold is about 80%. The proportion of the population that must be vaccinated against COVID-19 to begin inducing herd immunity is not known**.** This is an important area of research and will likely vary according to the community, the vaccine, the populations, prioritized for vaccination and other factors [4].

3. DATABASE

The database were collectedfrom 16th January 2021 to 6th July 2021 and it is based on the population as of Census 2011. The database features are Total Vaccine administered population, gendered based features, age group based featured and dose based features are considered for this study. India began its vaccination programme on 16 January 2021, operating 3,006 vaccination centres on the onset. In addition, based on the database to visualize in simple bar and line diagrams using MS-Excel.

3.1 TYPES OF VACCINE ADMINISTERED:

India has ramped up its corona virus vaccine production amidst the warnings of a third wave.

India has so far given more than 500 million doses of three approved vaccines – COVISHIELD, COVAXIN and Russia's Sputnik V. COVISHIELD is also known as s ChAdOx1 nCoV-19, or AZD1222, developed by Oxford University in partnership with AstraZeneca. Its manufacturing and trial partner is the Serum Institute of India, Pune, and Indian Council of Medical Research (ICMR). In the following table 1, shows that the detailed Description Of CIVISHELD Vaccine [5].

Table 1. Detailed Description of COVISHIELD Vaccine

Developed By	University of Oxford and AstraZeneca in collaboration with the SII
Vaccine Type	Modified chimpanzee adenovirus vector
Efficacy	DCGI: 70.42% overall
Storage Temperature	2-8°C
Dosage	Two doses (Gap 2.5-3 Months)
Routes of administration	Intramuscular injection

COVAXINis India's first indigenous COVID-19 vaccine. It is an inactivated vaccine that suppresses the virus' capability to duplicate yet keeps it unimpaired so that the immune system can still recognize it and create an immune reaction. COVAXIN helps in increasing the production of antibodies in the host body.Table 2, shows that the detailed Description of COVAXIN Vaccine [6].

Table 2. Detailed description of Covaxin Vaccine

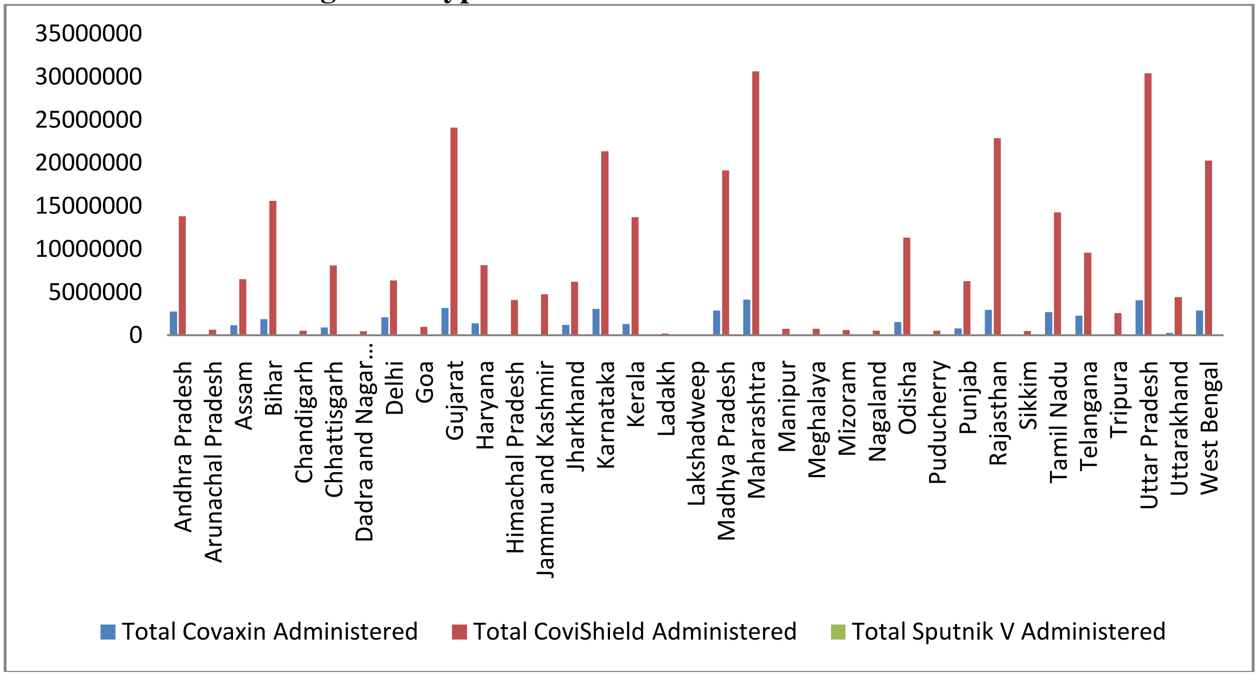
Developed By	Bharat Biotech in association with ICMR and NIV
Vaccine Type	Inactivated whole virus
Efficacy	N/A
Storage Temperature	2-8°C
Dosage	Two Doses (0, 14 Days)
Routes of administration	Intramuscular injection

In Table 3, Sputnik V is the globe's first registered COVID-19 vaccine based on a viral two-vector vaccine based on two human adenoviruses. The vaccine is manufactured in Russia but is approved for use in India and is being imported. Sputnik V is also known as Gam-COVID-Vac. It is progressing to the end of its clinical trials and soon will begin mass production. Sputnik V vaccine is titled after the 1st Soviet space satellite [7]. The result of Figure 1 is Types of Vaccine Administered in India.

Table 3. Detailed Description of Sputnik V -Vaccine

Developed By	Gamaleya Research Institute of Epidemiology and Microbiology, Russia. (Approved to be used in India)
Vaccine Type	Non-replicating viral vector (adenovirus)
Efficacy	91.4%
Storage Temperature	-18°C
Dosage	Two Doses
Routes of administration	Intramuscular injection

Figure 1. Types of Vaccine Administered in India



4. RESULT AND DISCUSSION

4.1 COVISHIELD

The gap between the two doses of Covid-19 vaccines has become a crucial issue as the Centre has revised the gap between two doses of **COVISHIELD** twice since the beginning of the vaccination drive in January. First, the two doses were meant to be administered four to six weeks apart. Later, the Centre revised it to four to eight weeks. Now the gap has been increased to 12 to 16 weeks.

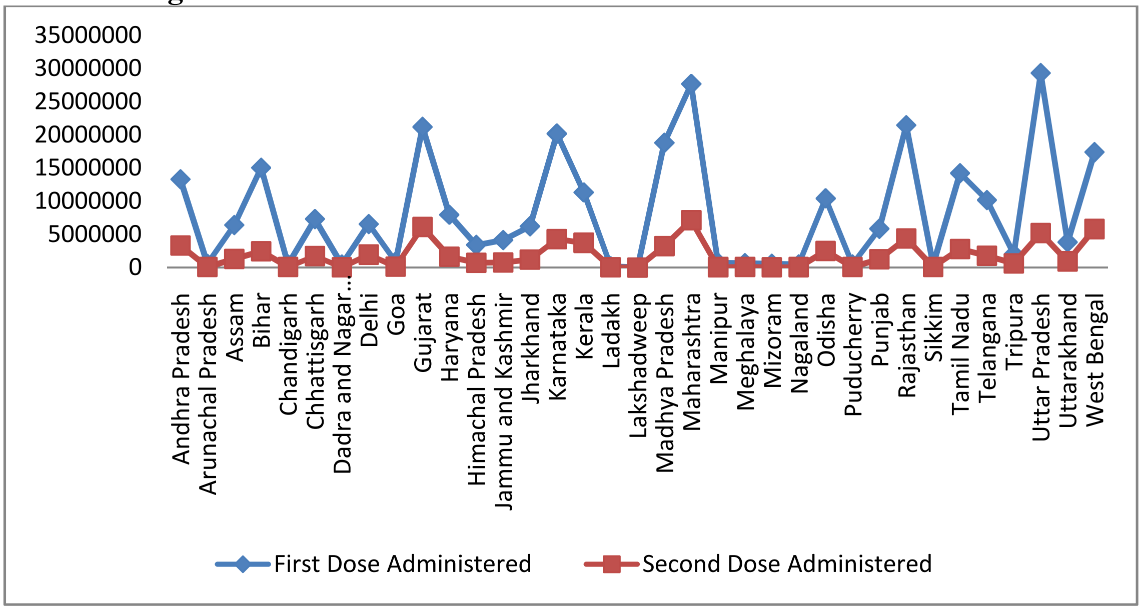
4.2 COVAXIN

COVAXIN, a COVID-19 vaccine manufactured by Bharat Biotech, requires a gap of only 28 days between two doses compared to Serum Institute of India's **COVISHIELD** that requires at least 84 days. The gap has created a significant increase in demand for **COVAXIN**, while the supply is unable to match it.

4.3 Sputnik V

The gap between two doses of Russian Covid-19 vaccine, Sputnik V, can be extended to 180 days, said Russian Direct Investment Fund (RDIF), which markets the vaccine globally. The gap between the two doses of Sputnik V can be widened and the jab will remain effective. In the following Figure 1.is shows that the results of Vaccine Administered in India of COVISHIELD followed by COVAXIN and Sputnik V respectively.

Figure 1. First and Second Dose Vaccine Administered in India

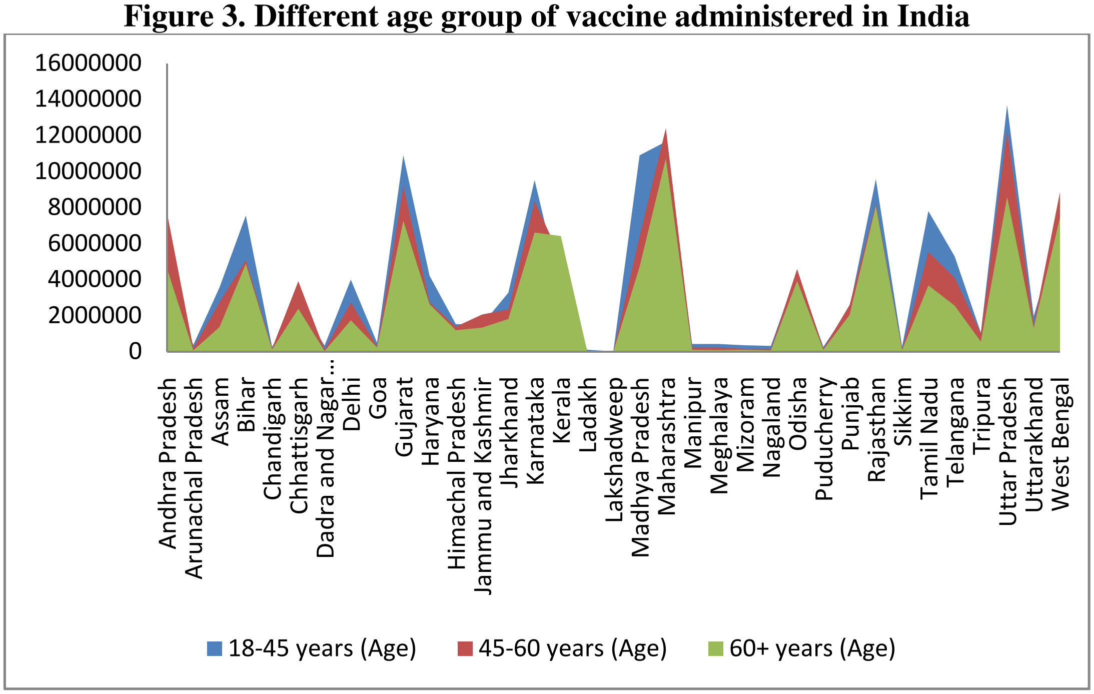


4.4 VACCINATED POPULATION OF DIFFERENT AGE GROUPS:

In figure 2, the results are exposes that about 49 percent of the population aged 60 years and above has been vaccinated with the first dose of COVID-19 vaccine. Out of estimated 59.7 crore people in the age group of 18-44 years, about 15 percent have been vaccinated with the first dose of the vaccine.

The COVID-19 vaccines are safe for most people 18 years and older, including those with pre-existing conditions of any kind, including auto-immune disorders. These conditions include: hypertension, diabetes, asthma, pulmonary, liver and kidney disease, as well as chronic infections that are stable and controlled.

Figure 3. Different age group of vaccine administered in India

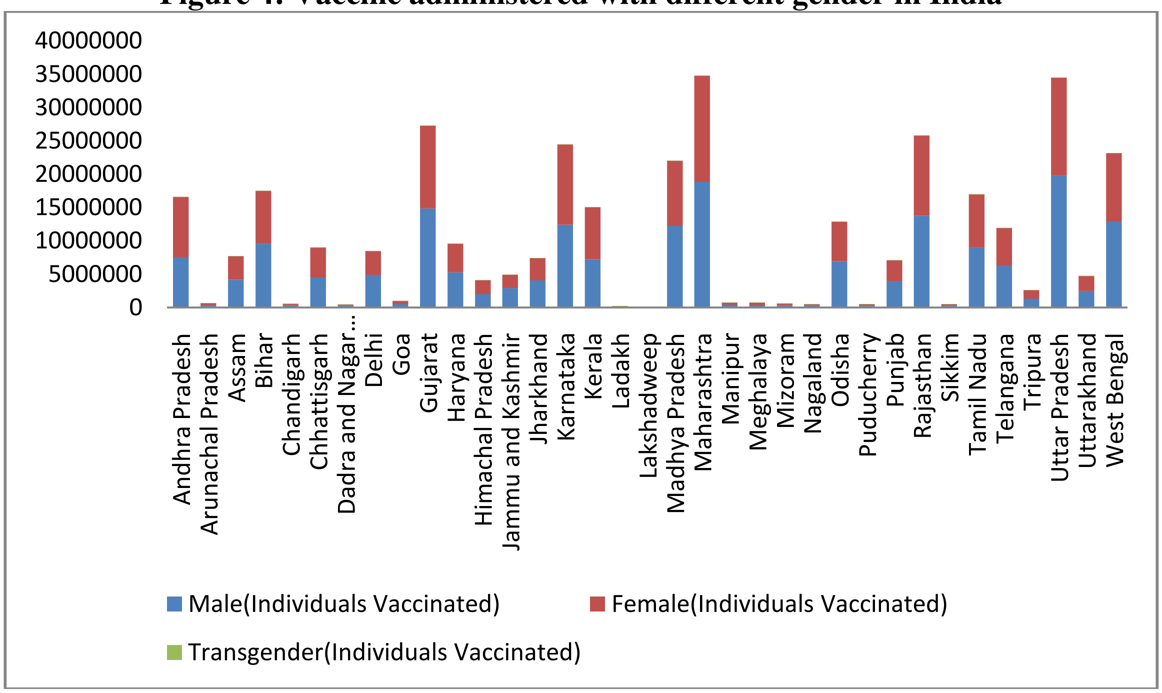


4.5 VACCINATED POPULATION OF GENDERS

At the global level, there is no significant difference in immunisation coverage for boys and girls. Yet in some countries and communities, gender discrimination means that boys have greater access to vaccines than do girls. In others, the opposite is true – girls have greater access.Gender-related barriers can also have an indirect impact on immunisation. Social and cultural norms, and the unequal status of women in many societies, can reduce the chances of children being vaccinated, by preventing their caregivers from accessing immunisation services.

The United Nations' Gender Development Index confirms that countries with a high level of gender equality have higher immunisation coverage.According to the vaccination data available in Union Health Ministry around 14.99 crore women and 17.48 men have received COVID-19 vaccines so far. While the government is also trying to bring guidelines for vaccination pregnant women as well, public health experts have raised concerns over the lower number of women vaccinations against COVID-19

Figure 4: Vaccine administered with different gender in India



5. CONCLUSION

In Figure 1. Illustrates the types of vaccine administered in India where it signifies the three main vaccines incubated in India – COVISHIELD, COVAXIN and Sputnik V. The visualization of graph shows that India's population has been highly immunized by COVISHIELD, moderately by COVAXIN and least by Sputnik V.

In Figure 2. illustrates the first and second dose vaccine administered in India where it signifies that the people of India are being highly vaccinated for first dose and are least vaccinated for the second dose. The first dose helps your body create an immune response, while the second dose is a booster that strengthens your immunity to the virus.

In Figure 3. illustrates the different age group of vaccine administered in India where it signifies that the Indian population ranging from age 18-45 years are being highly inoculated, population ranging from age 45-60 years are moderately inoculated and population ranging from 60 years and above are least inoculated by Covid-19 Vaccine.

For many older adults and those with on-going conditions like heart disease and diabetes, the vaccine can prevent severe illness or death from the corona virus. It also helps the body producing antibodies to the corona virus. These antibodies help your immune system fight the virus if you happen to be exposed, so it reduces your chance of getting the disease.

In Figure 4. illustrates the Vaccine administered with different gender in India where it signifies that male individuals of the country have been highly injected, female individuals are moderately injected and transgender individuals are least injected with Covid-19 vaccine.

REFERENCES

1. *Corona virus disease (COVID-19) available from: https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963#:~:text=Coronavirus%20is%20a%20family%20of,East%20respiratory%20syndrome%20(MERS)*
2. *Immunisation and Vaccination available from: https://www.science.org.au/learning/immunisation-and-climate-change/science-immunisation/what-is-immunisation*
3. *Symptoms and more information on COVID-19 available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19*
4. *Herd immunity and COVID-19 available from: https://www.who.int/news-room/q-a-detail/herd-immunity-lockdowns-and-covid-19?gclid=Cj0KCQjwvO2IBhCzARIsALw3A Sr3ZjdWHMKjoR3KX4w97CqtXW61b0NopbQ-MF9CZ3WxxiG7KAJE_rYaAn76EALw_wcB#*
5. *India's first Indigenous COVID-19 Vaccine available from: https://www.bharatbiotech.com/covaxin.html*

6. *Covishield vaccine efficacy rate, dose gap and side effects available from: https://www.mpnrc.org/covishield-vaccine-registration/*
7. *India's new COVID-19 vaccines available from: https://www.bbc.com/news/world-asia-india-55748124*
8. *COVID-19 vaccines in India available from: https://www.biotecnika.org/2021/01/coronavirus-vaccines-manufactured-in-india-and-their-types/*

DOI No. : 10.5958/2394-9309.2022.00006.3
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ROBUSTNESS OF GENERALIZED GROUP DIVISIBLE DESIGNS

Sreejith V.
Govt. College for Women, Thiruvananthapuram, Kerala

ABSTRACT

Generalized Group Divisible Designs have gained prominence in the theory of design and analysis of experiments. A new method of construction of generalized group divisible designs with s classes is presented in this article. The method is based on balanced incomplete block designs. Optimality, robustness and efficiency of generalized group divisible designs constructed using this method are also discussed.

1. INTRODUCTION

In this paper, we present a method of construction of Generalized Group Divisible Design with s groups (GGDD(s)). The robustness of such designs is also discussed. Here robustness of GGDD(s) is investigated when all the observations in a block are lost. One of the criteria of robustness is regarding connectedness. This was introduced by Ghosh(1982). According to this criterion, an incomplete block design is said to be robust against the loss of t(≥ 1) observations, if the residual design obtained by deleting these t observations remains connected. Another criterion of robustness is related to efficiency of the residual design. As per the efficiency criterion, a design is said to be robust if the efficiency of the residual design relative to the original one is not too small. Hedayat and John(1974), John(1976), Kageyama(1980, 1986, 1987), Saha and Kageyama(1984), Kageyama and Saha(1987), Dey and Dhall(1988), Whittinghill(1989), Mukerjee and Kageyama(1990), Srivastava, Gupta and Dey(1990), Ghosh, Kageyama and Mukherjee(1991) have made elaborate studies on robustness of designs.

In this study, we introduce a new method of construction of generalized group divisible design with s groups in section 2. In section 3, e-optimality of such designs is discussed. Robustness has been examined thoroughly in section 4. Robustness against loss of all the observations in a block has been discussed.

2. METHOD OF CONSTRUCTION

Consider s balanced incomplete block designs di(vi, bi, ri, k, λi), i = 1, 2, . . . , s. Merging of these BIB designs yields a generalized group divisible design with s groups, where the treatments of di are numbered 1, 2, . . . , vi (i = 1, 2, . . . , s) and v1 <v2 <v3 <. . .<vs.

Proof:

Let us first consider the treatments in the resulting design, say d□. There will be vs treatments. The first v1 of them are common to all the s BIBDs, each replicating ri times in di (i = 1, 2, . . . , s). Thus, in d□, they will be replicated r1 + r2 + . . . + rs = r□1 (say) times. Now consider the treatments v1 + 1, v1 + 2, . . . , v2. These treatments are common to all the BIBDs except the first one. Hence the number of replications for these treatments is r2 + r3 + . . . + rs = r□2(say). In a similar manner, we can see that the treatments v2 + 1, v2 + 2, . . . , v3 are replicated r3 + r4 + . . . + rs = r□3 times, the treatments v3 + 1, v3 + 2, . . . , v4 are replicated r4 + r5 + . . . + rs = r□4 times, . . . , and finally the treatments vs−1 + 1, vs−1 + 2, . . . , vs are replicated r□s times. On this basis, let us now classify the treatments in to s groups, viz., V1 = (1, 2, . . . , v1), V2 = (v1 + 1, v1 + 2, . . . , v2), . . . , Vs = (vs−1 + 1, vs−1 + 2, . . . , vs). Now let us look in to the concurrence λij of treatments.

For i, j ∈V1, λij = λ1 + λ2 + . . .+ λs = γ11 (say), i, j = 1, 2, . . . , v1

For i, j ∈V2, λij = λ2 + λ3 + . . .+ λs = γ22 (say), i, j = v1 + 1, v1 + 2, . . . , v2

...

For i, j ∈Vs, λij = λs = γss (say), i, j = vs−1 + 1, vs−1 + 2, . . . , vs.

For i ∈Vg and j ∈Vh, where 1 ≤ g <h ≤ s, λgh = λh + λh+1 + . . . + λs = γgh (say); i = vg−1 + 1, vg−1 + 2, . . . , vg, h = vh−1 + 1, vh−1 + 2, . . . , vh

Thus V1, V2, . . . , Vs form the s groups of a GGDD(s) and hence d∗ is a GGDD(s).

Example 2.1:

Consider the BIBDs $d_1(4,4,3,3,2)$, $d_2(7,7,3,3,1)$, $d_3(9,12,4,3,1)$.

The plan of these designs are given below

$$d_1: \begin{array}{cccc} 1 & 1 & 1 & 2 \\ 2 & 2 & 3 & 3 \\ 3 & 4 & 4 & 4 \end{array} \qquad d_2: \begin{array}{ccccccc} 1 & 1 & 1 & 2 & 2 & 3 & 3 \\ 2 & 4 & 6 & 4 & 5 & 4 & 5 \\ 3 & 5 & 7 & 6 & 7 & 7 & 6 \end{array} \quad \text{and}$$

$$d_3: \begin{array}{cccccccccccc} 1 & 4 & 7 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ 2 & 5 & 8 & 4 & 7 & 5 & 4 & 8 & 6 & 4 & 7 & 6 \\ 3 & 6 & 9 & 8 & 6 & 9 & 7 & 5 & 9 & 9 & 5 & 8 \end{array}$$

Merging these three designs, we get the following design, say d*:

$$\begin{array}{ccccccccccccccccccccccc} 1 & 1 & 1 & 2 & 1 & 1 & 1 & 2 & 2 & 3 & 3 & 1 & 4 & 7 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ 2 & 2 & 3 & 3 & 2 & 4 & 6 & 4 & 5 & 4 & 5 & 2 & 5 & 8 & 4 & 7 & 5 & 4 & 8 & 6 & 4 & 7 & 6 \\ 3 & 4 & 4 & 4 & 3 & 5 & 7 & 6 & 7 & 7 & 6 & 3 & 6 & 9 & 8 & 6 & 9 & 7 & 5 & 9 & 9 & 5 & 8 \end{array}$$

Obviously d* is a GGDD(3) with groups $V_1 = (1,2,3,4)$, $V_2 = (5,6,7)$ and $V_3 = (8,9)$.

The parameters of d* are v = 9, b = 23, k = 3, $r_1^* = 10$, $r_2^* = 7$, $r_3^* = 4$, $\gamma_{11} = 4$, $\gamma_{22} = 2$, $\gamma_{33} = 1$, $\gamma_{12} = 2$, $\gamma_{13} = 1$, $\gamma_{23} = 1$

3. E - OPTIMALITY

Theorem: The GGDD(s) constructed through this method is E-optimal in D(r*1 ,r*2,...,r*v; b; k) For proof, consider the following lemma Lemma (Jacroux, 1980): Let d ∈ D(r1,r2,...,rv; b; k) have information matrix Cα. Of the entries of NdN′ d= (λij) satisfy the condition that λij ≥rp(k-1)/(v-1) for all i≠j, where rp is the smallest number of replications, then d is the E-optimal in D(r1,r2,...,rv;b; k).

For the GGDD(s) constructed by the above method, the smallest number of replications is for the sth group and hence rp=r*s . Similarly the smallest off diagonal entry in NdN′ d is γss= λs. We know that λs(vs-1)=r*s (k-1) or λs=rp(k−1) (v−1) since r*s = rp and vs=vk.

Thus d* satisfies the condition of lemma with equality always. Hence d* is E-optimal.

Example 3.1:

Consider the GGDD(3) constructed in example 2.1. The smallest value of replication for this design is $r_p = 4$. Therefore, $\frac{r_p(k-1)}{(v-1)} = 1$

The concurrence matrix of d* is given by

$$N_{d*}N'_{d*} = \begin{bmatrix} 10 & 4 & 4 & 4 & 2 & 2 & 2 & 1 & 1 \\ 4 & 10 & 4 & 4 & 2 & 2 & 2 & 1 & 1 \\ 4 & 4 & 10 & 4 & 2 & 2 & 2 & 1 & 1 \\ 4 & 4 & 4 & 10 & 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 7 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 7 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 7 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 4 \end{bmatrix}$$

The smallest off - diagonal entry is 1, thereby satisfying the equality $\lambda_s = \frac{r_p(k-1)}{(v-1)}$. Hence d* is E - optimal.

4. ROBUSTNESS

Theorem: The GGDD(s) constructed this way is robust against the loss of all observations in a block if rs ≥ 2 and k ≥ 3. Proof:We'll prove that the GGDD(s) constructed this way is robust. For that consider the following theorem by Dey(1993). Theorem: A design d is robust against the loss of all observations in a block as per\the connectedness criterion, if the smallest eigen value of the information matrix is strictly larger than unity.

The C matrix of GGDD(s) constructed through the method described in section 2 is given by

$$C^* = \begin{bmatrix} r_1^* - \frac{1}{k}\left[(r_1^* - \gamma_{11})I_{v_1\times v_1} + \gamma_{11}E_{v_1\times v_1}\right] & -\frac{1}{k}\gamma_{12}E_{v_1\times(v_2-v_1)} & \cdots & -\frac{1}{k}\gamma_{1s}E_{v_1\times(v_s-v_{s-1})} \\ -\frac{1}{k}\gamma_{12}E_{(v_2-v_1)\times v_1} & r_2^* - \frac{1}{k}\left[(r_2^* - \gamma_{22})I_{(v_2-v_1)} + \gamma_{22}E_{(v_2-v_1)\times(v_2-v_1)}\right] & \cdots & -\frac{1}{k}\gamma_{2s}E_{(v_2-v_1)\times(v_s-v_{s-1})} \\ \vdots & \vdots & \vdots & \cdots \\ -\frac{1}{k}\gamma_{1s}E_{(v_s-v_{s-1})\times v_1} & -\frac{1}{k}\gamma_{2s}E_{(v_s-v_{s-1})\times(v_2-v_1)} & \cdots & r_s^* - \frac{1}{k}\left[(r_s^* - \gamma_{ss})I_{(v_s-vs-1)} + \gamma_{ss}E_{(v_s-v_{s-1})\times(v_s-v_{s-1})}\right] \end{bmatrix}$$

The v_s -1 non zero eigen values of C are r_1^* - $\frac{r_1^*-\gamma_{11}}{k}$ with multiplicity v_1 - 1, r_2^* - $\frac{r_2^*-\gamma_{22}}{k}$ with multiplicity v_2 - v_1, r_3^* - $\frac{r_3^*-\gamma_{33}}{k}$ with multiplicity v_3 - v_2, $\ldots$, r_s^* - $\frac{r_s^*-\gamma_{ss}}{k}$ with multiplicity v_s - v_{s-1}. Clearly the minimum among them is r_s^* - $\frac{r_s^*-\gamma_{ss}}{k}$. Note that r_s^* - $\frac{r_s^*-\gamma_{ss}}{k} > 1$ whenever $r_s^* \geq 2$ and $k \geq 3$. Hence our GGDD(s) is robust against the loss of all observations in a block.

Example 4.1:

The C matrix of the GGDD(3) shown in example 2.1 is

$$C^* = \begin{bmatrix} \frac{20}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{4}{3} & \frac{20}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{4}{3} & -\frac{4}{3} & \frac{20}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & \frac{20}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & \frac{14}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & \frac{14}{3} & -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & -\frac{2}{3} & \frac{14}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{8}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & \frac{8}{3} \end{bmatrix}$$

The non zero eigen values of C are 8($= r_1^*$ - $\frac{r_1^*-\gamma_{11}}{k}$) with multiplicity 3($= v_1$ - 1), $\frac{16}{3}$($= r_2^*$ - $\frac{r_2^*-\gamma_{22}}{k}$) with multiplicity 3($= v_2$ - v_1) and 3($= r_3^*$ - $\frac{r_3^*-\gamma_{33}}{k}$) with multiplicity 2($= v_3$ - v_2). The minimum non zero eigen value is $3 > 1$ and hence d^* is robust against the loss of all the observations in a block.

5 Efficiency

Suppose that all the observations in a block of design d are missing and assume that these observations corresponds to treatments belonging to the last group, V_s. Let d** be the resulting design and C** be its C - matrix. Then the eigen values of C_1 will be r_1^* - $\frac{r_1^*-\gamma_{11}}{k}$ with multiplicity v_1 - 1, r_2^* - $\frac{r_2^*-\gamma_{22}}{k}$ with multiplicity v_2 - v_1, r_3^* - $\frac{r_3^*-\gamma_{33}}{k}$ with multiplicity v_3 - v_2, ..., r_{s-1}^* - $\frac{r_{s-1}^*-\gamma_{s-1s-1}}{k}$ with multiplicity v_{s-1} - v_{s-2}, r_s^* - $\frac{r_s^*-\gamma_{ss}}{k}$ with multiplicity v_s - v_{s-1} - $(k-1)$ and $r_s^* - 1$ - $\frac{r_s^*-\gamma_{ss}}{k}$ with multiplicity $k-1$. Therefore, the efficiency of d** with respect to d* is

$$E = \frac{\displaystyle\sum_{i=1}^{s} \frac{v_i - v_{i-1}}{r_i^* - \frac{r_i-\gamma_{ii}}{k}}}{\displaystyle\sum_{i=1}^{s-1} \frac{v_i - v_{i-1}}{r_i^* - \frac{r_i^*-\gamma_{ii}}{k}} + \frac{v_s - v_{s-1} - (k-1)}{r_s^* - \frac{r_s^*-\gamma_{ss}}{k}} + \frac{k-1}{r_s^* - 1 - \frac{r_s^*-\gamma_{ss}}{k}}}$$

assuming $v_0 = 1$.

6 Illustration:

Consider the BIBDs d_1(5, 10, 6, 3, 3) and d_2(13, 26, 6, 3, 1) whose blocks are as follows:

$$d_1: \begin{array}{cccccccccc} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 3 \\ 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 \\ 3 & 4 & 5 & 4 & 5 & 4 & 5 & 5 & 5 & 5 \end{array}$$

and

$$d_2: \begin{array}{cccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 1 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 1 & 3 \\ 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 1 & 2 & 3 & 4 & 8 \end{array}$$

$$\begin{array}{cccccccccccc} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 1 & 2 \\ 9 & 10 & 11 & 12 & 13 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}$$

Let us merge d_1 and d_2 to get the design d* with blocks

$$d^*: \begin{array}{cccccccccccccccccc} 1 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 3 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 5 & 4 & 5 & 4 & 5 & 5 & 5 & 5 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \end{array}$$

9	10	11	12	13	1	2	3	4	5	6	7	8	9	10	11	12	13
10	11	12	13	1	3	4	5	6	7	8	9	10	11	12	13	1	2
13	1	2	3	4	8	9	10	11	12	13	1	2	3	4	5	6	7

Clearly d* is a GGDD(2) with v = 13, b = 36, k = 3, V_1 = (1, 2, 3, 4, 5), V_2 = (6, 7, 8, 9, 10, 11, 12, 13), $r_1^* = 12$, $r_2^* = 6$, $\gamma_{11} = 4$, $\gamma_{22} = 1$ and $\gamma_{12} = 1$. The C matrix of d* is

$$C^* = \begin{bmatrix}
8 & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & 8 & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & 8 & -\frac{4}{3} & -\frac{1}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & 8 & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & 8 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4 & -\frac{1}{3} \\
-\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 4
\end{bmatrix}$$

Its non zero eigen values are 9.33($= r_1^* - \frac{r_1^* - \lambda_1}{k}$) with multiplicity 4(= v_1 - 1) and 4.33($= r_2^* - \frac{r_2^* - \lambda_2}{k}$) with multiplicity 8(= v_2 - v_1) Now let us delete the block (9, 10, 13) from d* to get the residual design d**:

$$d^{**}: \begin{array}{cccccccccccccccccc}
1 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 3 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
3 & 4 & 5 & 4 & 5 & 4 & 5 & 5 & 5 & 5 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12
\end{array}$$

REFERENCE:

1. *Dey, A. and Dhall, S.P.(1988), Robustness of augmented BIB designs, Sankhya, B50, 376-381.*
2. *Ghosh, S.(1982), Robustness of BIB designs against non-availability of data, Journal of Statistical Planning and Inference, 6, 29-32.*
3. *Ghosh, S., Kageyama, S. and Mukerjee, R.(1992), Efficiency of connected binary block designs when a single observation is unavailable, Annals of Institute of Statistics and Mathematics, 14(3), 593 - 603.*
4. *Hedayat, A. and John, P.W.M.(1974), Resistant and susceptible BIB designs, Annals of Statistics, 2, 148 - 158.*
5. *John, P.W.M.(1976), Robustness of incomplete block designs, Annals of Statistics,4, 960 - 962.*
6. *Kageyama, S.(1980), Robustness of connected balanced block designs, Annals of Institute of Statistics and Mathematics, A, 32, 255 - 261.*
7. *Kageyama, S.(1986), Robustness of 3-associate PBIB designs of two types,Journal of Japan Statistical Society, 16, 21 - 24.*
8. *Kageyama, S.(1987), Some characterizations of locally resistant BIB designs of degree one, Annals of Institute of Statistics and Mathematics, A, 39, 661 - 669.*
9. *Kageyama, S. and Saha, G.M.(1987), On resistant t - designs, Ars Combinatoria,23, 81 - 92.*
10. *Mukerjee, R. and Kageyama, S.(1990), Robustness of group divisible designs, Communications in Statistics - Theory and Methods, 19, 3189 - 3203.*
11. *Saha, G.M. and Kageyama, S.(1984), On Connectedness - conditions of PSIB designs, Journal of Japan Statistical Society, 14. 169 - 178.*
12. *Srivastava, R., Gupta, V.K. and Dey, A.(1990), Robustness of some designs against missing observations, Communications in Statistics - Theory and Methods, 19(1), 121 - 126.*
13. *Whittinghill, D.C., III(1989), Balanced block designs robust against the loss of a single observation, Journal of Statistical Planning and Inference, 22, 71 - 80.*

Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
DOI No. : 10.5958/2394-9309.2022.00007.5
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

EQUITABLE COLORING OF CARTESIAN PRODUCT OF SEMI-TOTAL POINT GRAPH, SEMI-TOTAL LINE GRAPH AND TOTAL GRAPH OF CYCLE WITH PATH GRAPH

G. Murali*, M.S. Paulraj**
* & **Assistant Professor, Department of Mathematics, A.M. Jain College, Chennai, India
E-mail : murali87maths@gmail.com

ABSTRACT

Let G be graph and let $\{V_i : 1 \leq i \leq k\}$ be a properk-coloring of G. If $\left||V_i| - |V_j|\right| \leq 1, i, j = 1, 2, \ldots, k$, then this coloring is called an Equitable k-coloring. The smallest integer k for which G is equitably k-colorable is known as Equitable Chromatic number of Gand is denoted by $\chi_{=}(G)$. In this paper, an attempt has been made to establish the acceptance of equitable coloring to the cartesian product of semi-total point graph, semi-total line graph and total graph of cycle with path. The equitable chromatic number for the said graphs has also been calculated.

Keywords: Cartesian product, Semi-total point graph, Semi-total line graph, Total graphs of a graph
Mathematics Subject Classification 05C15

1. INTRODUCTION

The concept of coloring in graphs has fascinated the graph theoretic researchers around the globe, motivating them to come up with different types of coloring to different graphs. The notion of equitable coloring was first defined and a conjecture was proposed by Meyer in 1973[10]. The proof for his conjecture was attempted by a wide-spectrum of researchers and later concluded that it is improbable to calculate the equitable chromatic number and compare the same with chromatic number. The very existence of equitable coloring in product graphs was first given by Hanna Furmanczyk [6], in 2006. This paved way for Wu-Hsiung Lin and Gerard J. Chang [14] in 2012, to prove that the cartesian product of graphs admit equitable coloring. These results provided the required impetus to attempt the acceptance of equitable coloring for some graphs.

All Graphs Considered In This Research Paper Are Finite, Simple And Undirected.

2. BASIC DEFINITIONS

Definition 2.1 The semi-totalpoint graph $T_2(G)$ [11] of G is the graph whose vertex set is $V(G) \cup E(G)$. For $a, b \in V(T_2(G))$, a and b are adjacent if and only if the following conditions hold

(i) $a, b \in V(G)$, a, b are adjacent vertices of G.

(ii) $a \in V(G)$ and $b \in E(G)$, b is incident with $a$in G.

Definition 2.2 The semi-total line graph $T_1(G)$ [11] of G is the graph whose vertex set is $V(G) \cup E(G)$. For $a, b \in V(T_1(G))$,a and b are adjacent if and only if the following conditions hold

(i) $a, b \in E(G)$, a, b are adjacent edges of G.

(ii) $a \in V(G)$ and $b \in E(G)$, b is incident with $a$in G.

Definition 2.3 The total graph $T(G)$ [11] of G is a graph whose vertex set is $V(G) \cup E(G)$. For$a, b \in V(T(G))$, a and b are adjacentif and only if the following conditions hold

(i) $a, b \in V(G)$, a, b are adjacent vertices of G.

(ii) $a, b \in E(G)$, a, b are adjacent edges of G.

(iii) $a \in V(G)$ and $b \in E(G)$, b is incident with $a$in G.

Definition 2.4 For graphs G_1 and G_2, the Cartesian products of graphs G_1 and G_2 will be denoted by $G_1 \; G_2$ with vertex set $V(G_1 \; G_2) = \{(x, y): x \in V(G_1), y \in V(G_2)\}$ and edge set

$E(G_1 \; G_2) = \{(x, y)(u, v): x = u$ and $yv \in E(G_2)$ or $y = v$ and $xu \in E(G_1)\}$.

In this paper, we used $u_i v_j$ instead of (x, y).

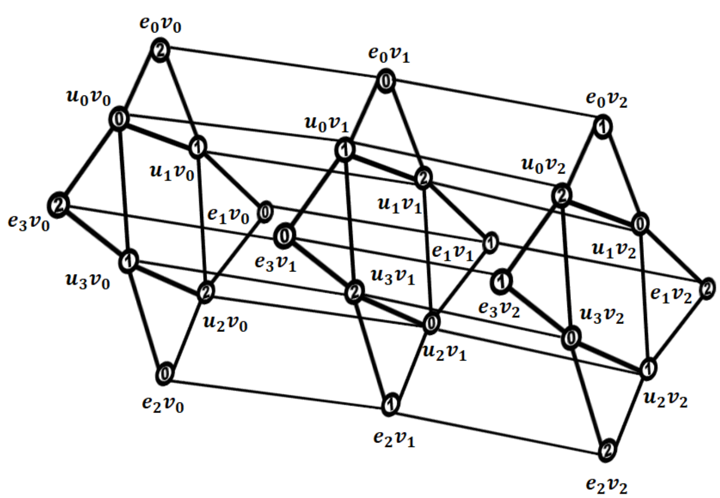


Fig.1 Example of Cartesian product$T_2(C_4) \; P_3$

3. PRE-REQUISITES

The following results are useful in proving our main result.

Result3.1[10]IfG is a connected graph, different from$C_{2n+1}$and$K_n \forall\, n \geq 1$, then$\chi_=(G) \leq \Delta(G)$.

Result3.2 [3]The Equitable Δ- Coloring Conjecture (EΔCC): A connected graph G is equitable$\Delta(G)$ – colorable if G is different from C_{2n+1}, K_n and $K_{2n+1,2n+1} \forall\, n \geq 1$.

Result3.3 [14] $\chi_=(G_1 \; G_2) \leq \chi(G_1)\chi(G_2)$ for connected graph G_1 and G_2.

4. MAIN RESULTS

In this paper, the equitable chromatic number of the following graphs have been calculated:

(i) $T_2(C_m)\ \ P_n$

(ii) $T_1(C_m)\ \ P_n$

(iii)$T(C_m)\ \ P_n$

For a non-negativeinteger $m$and$n$, the vertex set$\{u_0, u_1, u_2 \dots, u_{m-1}\}$and edge set$\{e_0, e_1, e_2 \dots, e_{m-1}\}$ of C_m, the vertex set of P_n is $\{v_0, v_1, v_2 \dots, v_{n-1}\}$.

Equitable coloring of Cartesian product of graphs

The equitable chromatic number for Cartesian product of graphs for the followingis being attempted;

4.1. Product of $T_2(C_m)$ with P_n

Cartesian product of semi-totalpoint graph of cycle with path

The semi-totalpoint graph of cycle $T_2(C_m)$ on $2m$ vertices and $3m$ edges having the vertex set and edge set is given by

$V(T_2(C_m)) = \bigcup_{i=0}^{m-1}\{u_i; e_i\}$ and $E(T_2(C_m)) = \left(\bigcup_{i=0}^{m-1}\{u_i u_{i+1(mod\ m)}; u_i e_i; e_i u_{i+1(mod\ m)}\}\right)$

Theorem 4.1.1 If m and n are non-negative integers, $m \geq 3, n \geq 2$ then the equitable chromatic number of $T_2(C_m)\ \ P_n$ is 3

Proof. Clearly,$|V(T_2(C_m)\ \ P_n)| = 2mn$ and $|E(T_2(C_m)\ \ P_n)| = m(5n-2)$.

Let the vertex set and edge set of the Cartesian product $T_2(C_m)\ \ P_n$ be

$$V(T_2(C_m)\ \ P_n) = \bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{u_i v_j; e_i v_j\}$$

$$E(T_2(C_m)\ \ P_n) = \left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{(u_i v_j)(u_{i+1(mod\ m)} v_j); (u_i v_j)(e_i v_j); (e_i v_j)(u_{i+1(mod\ m)} v_j)\}\right)$$
$$\cup \left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-2}\{(u_i v_j)(u_i v_{j+1}); (e_i v_j)(e_i v_{j+1})\}\right)$$

Fig.2 Cartesian product of semi-totalpoint graph of cycle with path, $T_2(C_m)\ \ P_n$.

Define the coloring using the map$f : V(T_2(C_m)\ \ P_n) \rightarrow \{0,1,2\}$ of $T_2(C_m)\ \ P_n$ as follows:

Case 1: When $m \equiv 0(mod\ 3)$

$$f(u_i v_j) = (i+j)(mod\ 3)$$

$$f(e_i v_j) = (i+j+2)(mod\ 3)$$

$\forall\ i = 0,1,2, \dots, (m-1)$ and $j = 0,1,2, \dots, (n$ -1)

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0 = \begin{cases} u_i v_j ; & (i+j) \equiv 0(mod\ 3) \\ e_i v_j ; (i+j+2) \equiv 0(mod\ 3) \end{cases}$$

$$V_1 = \begin{cases} u_i v_j ; & (i+j) \equiv 1(mod\ 3) \\ e_i v_j ; (i+j+2) \equiv 1(mod\ 3) \end{cases}$$

$$V_2 = \begin{cases} u_i v_j ; & (i+j) \equiv 2(mod\ 3) \\ e_i v_j ; (i+j+2) \equiv 2(mod\ 3) \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.2$u_i v_j, u_{i+1(mod\ m)} v_j \in V(T_2(C_m)\ P_n)$ are adjacent in $T_2(C_m)\ P_n$.Consider two vertices of type $u_i v_j$and $u_k v_j$ where $k = (i+1)(mod\ m)$ .Clearly, if $(i+j) \equiv 0(mod\ 3)$ then $(k+j) \not\equiv 0(mod\ 3)$, also if $(i+j) \equiv 1(mod\ 3)$ then $(k+j) \not\equiv 1(mod\ 3)$ and if $(i+j) \equiv 2(mod\ 3)$ then $(k+j) \not\equiv 2(mod\ 3)$.Hence any two vertices of type $u_i v_j$and $u_{i+1(mod\ m)} v_j$does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color class has been shown below,

(i) $u_i v_j$and $e_i v_j$

(ii) $e_i v_j$and $u_{i+1(mod\ m)} v_j$

(iii) $u_i v_j$and $u_i v_{j+1}$

(iv) $e_i v_j$and $e_i v_{j+1}$

Case 2: When $m \equiv 1(mod\ 3)$

$$f(u_i v_j) = \begin{cases} (i+j)(mod\ 3) ; 0 \le i \le m-2 \\ (j+1)(mod\ 3) ; \quad i = m-1 \end{cases}$$

$$f(e_i v_j) = \begin{cases} (i+j+2)(mod\ 3) ;\ 0 \le i \le m-3 , i = m-1 \\ j(mod\ 3) ; \quad i = m-2 \end{cases}$$

$\forall\ i = 0,1,2, \dots, (m-1)$ and $j = 0,1,2, \dots, (n$ -1)

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0 = \begin{cases} u_i v_j\ ; & (i+j) \equiv 0(mod\ 3) \text{if } 0 \le i \le m-2 \text{ and} (j+1) \equiv 0(mod\ 3) \text{if } i = m-1 \\ e_i v_j ; & (i+j+2) \equiv 0(mod\ 3) \text{if } 0 \le i \le m-3, i = m-1 \text{ and } j \equiv 0(mod\ 3) \text{if } i = m-2 \end{cases}$$

$$V_1 = \begin{cases} u_i v_j\ ; & (i+j) \equiv 1(mod\ 3) \text{if } 0 \le i \le m-2 \text{ and} (j+1) \equiv 1(mod\ 3) \text{if } i = m-1 \\ e_i v_j ; & (i+j+2) \equiv 1(mod\ 3) \text{if } 0 \le i \le m-3, i = m-1 \text{ and } j \equiv 1(mod\ 3) \text{if } i = m-2 \end{cases}$$

$$V_2 = \begin{cases} u_iv_j\ ; & (i+j) \equiv 2(mod\ 3)\text{if } 0 \le i \le m-2 \text{ and}(j+1) \equiv 2(mod\ 3)\text{if } i = m-1 \\ e_iv_j; & (i+j+2) \equiv 2(mod\ 3)\text{if } 0 \le i \le m-3, i = m-1 \text{ and } j \equiv 2(mod\ 3)\text{if } i = m-2 \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.2 $e_iv_j, u_{i+1(mod\ m)}v_j \in V(T_2(C_m)\ P_n)$ are adjacent in $T_2(C_m)\ P_n$.Consider two vertices of type $e_iv_j$and u_kv_j where $k = (i+1)(mod\ m)$.

Clearly, if $(i+j+2) \equiv 0(mod\ 3)\text{if } 0 \le i \le m-3, i = m-1$ and $j \equiv 0(mod\ 3)\text{if } i = m-2$ then $(k+j) \not\equiv 0(mod\ 3)\text{if } 0 \le k \le m-2$ and$(j+1) \not\equiv 0(mod\ 3)\text{if } k = m-1$,

Also if $(i+j+2) \equiv 1(mod\ 3)\text{if } 0 \le i \le m-3, i = m-1$ and $j \equiv 1(mod\ 3)\text{if } i = m-2$ then $(k+j) \not\equiv 1(mod\ 3)\text{if } 0 \le k \le m-2$ and$(j+1) \not\equiv 1(mod\ 3)\text{if } k = m-1$ and

If $(i+j+2) \equiv 2(mod\ 3)\text{if } 0 \le i \le m-3, i = m-1$ and $j \equiv 2(mod\ 3)\text{if } i = m-2$ then $(k+j) \not\equiv 2(mod\ 3)\text{if } 0 \le k \le m-2$ and$(j+1) \not\equiv 2(mod\ 3)\text{if } k = m-1$. Hence any two vertices of type $e_iv_j$and $u_{i+1(mod\ m)}v_j$does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color classis established below,

(i) $u_iv_j$and e_iv_j

(ii) $u_iv_j$and $u_{i+1(mod\ m)}v_j$

(iii) $u_iv_j$and u_iv_{j+1}

(iv) $e_iv_j$and e_iv_{j+1}

Case 3: When $m \equiv 2(mod\ 3)$

$$f(u_iv_j) = (i+j)(mod\ 3)$$

$$f(e_iv_j) = \begin{cases} (i+j+2)(mod\ 3)\ ; & 0 \le i \le m-2 \\ (j+2)(mod\ 3)\ ; & i = m-1 \end{cases}$$

$\forall\ i = 0,1,2,\dots,(m-1)$ and $j = 0,1,2,\dots,(n\ \text{-}1)$

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0 = \begin{cases} u_iv_j\ ; & i+j \equiv 0(mod\ 3) \\ e_iv_j\ ; & (i+j+2) \equiv 0(mod\ 3)\text{if } 0 \le i \le m-2 \text{ and}(j+2) \equiv 0(mod\ 3)\text{if } i = m-1 \end{cases}$$

$$V_1 = \begin{cases} u_iv_j\ ; & i+j \equiv 1(mod\ 3) \\ e_iv_j\ ; & (i+j+2) \equiv 1(mod\ 3)\text{if } 0 \le i \le m-2 \text{ and}(j+2) \equiv 1(mod\ 3)\text{if } i = m-1 \end{cases}$$

$$V_2 = \begin{cases} u_iv_j\ ; & i+j \equiv 2(mod\ 3) \\ e_iv_j\ ; & (i+j+2) \equiv 2(mod\ 3)\text{if } 0 \le i \le m-2 \text{ and}(j+2) \equiv 2(mod\ 3)\text{if } i = m-1 \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.2 $e_iv_j, e_iv_{j+1} \in V(T_2(C_m)\ P_n)$ are adjacent in $T_2(C_m)\ P_n$ .Consider two vertices of type $e_iv_j$and e_iv_{j+1} .Clearly, if $(i + j + 2) \equiv 0(mod\ 3)$if $0 \leq i \leq m - 2$ and$(j + 2) \equiv 0(mod\ 3)$if $i = m - 1$ then $(i + j + 3) \not\equiv 0(mod\ 3)$if $0 \leq i \leq m - 2$ and $(j + 3) \not\equiv 0(mod\ 3)$if $i = m - 1$,

Also if $(i + j + 2) \equiv 1(mod\ 3)$if $0 \leq i \leq m - 2$ and$(j + 2) \equiv 1(mod\ 3)$if $i = m - 1$ then $(i + j + 3) \not\equiv 1(mod\ 3)$if $0 \leq i \leq m - 2$ and$(j + 3) \not\equiv 1(mod\ 3)$if $i = m - 1$and

If $(i + j + 2) \equiv 2(mod\ 3)$if $0 \leq i \leq m - 2$ and $(j + 2) \equiv 2(mod\ 3)$if $i = m - 1$ then $(i + j + 3) \not\equiv 2(mod\ 3)$if $0 \leq i \leq m - 2$ and$(j + 3) \not\equiv 2(mod\ 3)$if $i = m - 1$.

Hence any two vertices of type $e_iv_j$and e_iv_{j+1}does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color classis proved below,

(i) $u_iv_j$and e_iv_j

(ii) $e_iv_j$and $u_{i+1(mod\ m)}v_j$

(iii) $u_iv_j$and $u_{i+1(mod\ m)}v_j$

(iv) $u_iv_j$and u_iv_{j+1}

(a) If $n \equiv 0(mod\ 3)$ then $|V_0| = \frac{2mn}{3}$, $|V_1| = \frac{2mn}{3}$and $|V_2| = \frac{2mn}{3}$.

(b) If $n \equiv 1(mod\ 3)$ then $|V_0| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m-1}{3}\right\rceil\right)$, $|V_1| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m}{3}\right\rceil\right)$ and $|V_2| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m-2}{3}\right\rceil\right)$.

(c) If $n \equiv 2(mod\ 3)$ then $|V_0| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2(m-1)}{3}\right\rceil\right)$, $|V_1| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2m}{3}\right\rceil\right)$ and $|V_2| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2m-1}{3}\right\rceil\right)$.

Obviously, from (a), (b) and (c) f satisfies the condition$\left||V_i| - |V_j|\right| \leq 1, i \neq j$. Clearly, f is a proper coloring of $T_2(C_m)\ P_n$and so $\chi_=(T_2(C_m)\ P_n) = 3$.

Hence the theorem holds.

4.2. Product of $T_1(C_m)$ with P_n

Cartesian product of semi-total linegraph of cycle with path.

The semi-total line graph of cycle$T_1(C_m)$ on $2m$ vertices and $3m$ edges with vertex set and edge set is given by

$V(T_1(C_m)) = \bigcup_{i=0}^{m-1}\{u_i; e_i\}$ and $E(T_1(C_m)) = \left(\bigcup_{i=0}^{m-1}\{u_ie_i; e_iu_{i+1(mod\ m)}; e_ie_{i+1(mod\ m)}\}\right)$

Theorem 4.2.1 If m and n are non-negative integers, $m \geq 3$, $n \geq 2$then the equitable chromatic number of $T_1(C_m)\ \ P_n$ is 3

Proof. Clearly,$|V(T_1(C_m)\ \ P_n)| = 2mn$ and $|E(T_1(C_m)\ \ P_n)| = m(5n-2)$.

Let the vertex set and edge set of the Cartesian product $T_1(C_m)\ \ P_n$ be

$$V(T_1(C_m)\ \ P_n) = \bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{u_iv_j; e_iv_j\}$$

$$E(T_1(C_m)\ \ P_n) = \left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{(u_iv_j)(e_iv_j); (e_iv_j)(u_{i+1(mod\ m)}v_j); (e_iv_j)(e_{i+1(mod\ m)}v_j)\}\right)$$
$$\cup\left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-2}\{(u_iv_j)(u_iv_{j+1}); (e_iv_j)(e_iv_{j+1})\}\right)$$

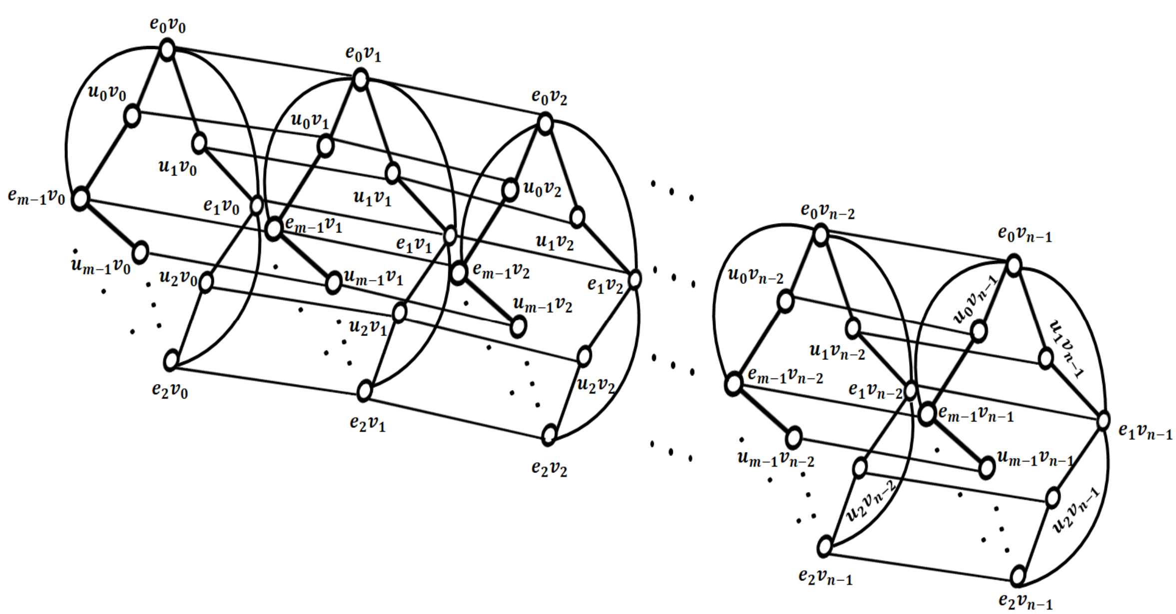


Fig.3 Cartesian product of semi-total line graph of cycle with path,$T_1(C_m)\ \ P_n$.

Define the coloring using the map$f : V(T_1(C_m)\ \ P_n) \rightarrow \{0,1,2\}$ of $T_1(C_m)\ \ P_n$ as follows:

Case 1: When $m \equiv 0(mod\ 3)$

$$f(u_iv_j) = (i + j + 2)(mod\ 3)$$

$$f(e_iv_j) = (i + j + 1)(mod\ 3)$$

$$\forall\, i = 0,1,2, \dots, (m-1) \text{ and } j = 0,1,2, \dots, (n-1)$$

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0 = \begin{cases} u_i v_j \,;\ (i+j+2) \equiv 0 (mod\ 3) \\ e_i v_j \,;\ (i+j+1) \equiv 0 (mod\ 3) \end{cases}$$

$$V_1 = \begin{cases} u_i v_j \,;\ (i+j+2) \equiv 1 (mod\ 3) \\ e_i v_j \,;\ (i+j+1) \equiv 1 (mod\ 3) \end{cases}$$

$$V_2 = \begin{cases} u_i v_j \,;\ (i+j+2) \equiv 2 (mod\ 3) \\ e_i v_j \,;\ (i+j+1) \equiv 2 (mod\ 3) \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.3 $e_i v_j, e_{i+1(mod\ m)} v_j \in V(T_1(C_m)\ P_n)$ are adjacent in $T_1(C_m)\ P_n$ .Consider two vertices of type $e_i v_j$and $e_k v_j$ where $k = (i+1)(mod\ m)$ .Clearly, if $(i+j+1) \equiv 0(mod\ 3)$ then $(k+j+1) \not\equiv 0(mod\ 3)$, also if $(i+j+1) \equiv 1(mod\ 3)$ then $(k+j+1) \not\equiv 1(mod\ 3)$and if $(i+j+1) \equiv 2(mod\ 3)$ then$(k+j+1) \not\equiv 2(mod\ 3)$. Hence any two vertices of type $e_i v_j$and $e_{i+1(mod\ m)} v_j$does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color class has is established below,

(i) $u_i v_j$and $e_i v_j$

(ii) $e_i v_j$and $u_{i+1(mod\ m)} v_j$

(iii) $u_i v_j$and $u_i v_{j+1}$

(iv) $e_i v_j$and $e_i v_{j+1}$

Case 2: When $m \equiv 1(mod\ 3)$

$$f(u_i v_j) = \begin{cases} (j+2)(mod\ 3) \,; & i = 0 \\ (i+j+1)(mod\ 3) \,; & 1 \le i \le m-2 \\ j(mod\ 3) \,; & i = m-1 \end{cases}$$

$$f(e_i v_j) = \begin{cases} (i+j)(mod\ 3) \,; & 0 \le i \le m-2 \\ (j+1)(mod\ 3) \,; & i = m-1 \end{cases}$$

$$\forall\, i = 0,1,2, \dots, (m-1) \text{ and } j = 0,1,2, \dots, (n-1)$$

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0=\begin{cases}u_iv_j\ ;(j+2)\equiv 0(mod\ 3)\text{if } i=0,(i+j+1)\equiv 0(mod\ 3)\text{if } 1\le i\le m-2 \text{ and } j\equiv 0(mod\ 3)\text{if } i=m-1\\ e_iv_j\ ; \qquad (i+j)\equiv 0(mod\ 3)\text{if } 0\le i\le m-2 \text{ and}(j+1)\equiv 0(mod\ 3)\text{if } i=m-1\end{cases}$$

$$V_1=\begin{cases}u_iv_j\ ;(j+2)\equiv 1(mod\ 3)\text{if } i=0,(i+j+1)\equiv 1(mod\ 3)\text{if } 1\le i\le m-2 \text{ and } j\equiv 1(mod\ 3)\text{if } i=m-1\\ e_iv_j\ ; \qquad (i+j)\equiv 1(mod\ 3)\text{if } 0\le i\le m-2 \text{ and}(j+1)\equiv 1(mod\ 3)\text{if } i=m-1\end{cases}$$

$$V_2=\begin{cases}u_iv_j\ ;(j+2)\equiv 2(mod\ 3)\text{if } i=0,(i+j+1)\equiv 2(mod\ 3)\text{if } 1\le i\le m-2 \text{ and } j\equiv 2(mod\ 3)\text{if } i=m-1\\ e_iv_j\ ; \qquad (i+j)\equiv 2(mod\ 3)\text{if } 0\le i\le m-2 \text{ and}(j+1)\equiv 2(mod\ 3)\text{if } i=m-1\end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.3 $e_iv_j, e_iv_{j+1}\in V(T_1(C_m)\ P_n)$ are adjacent in $T_1(C_m)\ P_n$ .Consider two vertices of type e_iv_j and e_iv_{j+1} .Clearly, if $(i+j)\equiv 0(mod\ 3)$if $0\le i\le m-2$ and$(j+1)\equiv 0(mod\ 3)$if $i=m-1$ then $(i+j+1)\not\equiv 0(mod\ 3)$if $0\le i\le m-2$ and$(j+2)\not\equiv 0(mod\ 3)$if $i=m-1$,

Also if $(i+j)\equiv 1(mod\ 3)$if $0\le i\le m-2$ and $(j+1)\equiv 1(mod\ 3)$if $i=m-1$ then $(i+j+1)\not\equiv 1(mod\ 3)$if $0\le i\le m-2$ and$(j+2)\not\equiv 1(mod\ 3)$if $i=m-1$and

If $(i+j)\equiv 2(mod\ 3)$if $0\le i\le m-2$ and $(j+1)\equiv 2(mod\ 3)$if $i=m-1$ then $(i+j+1)\not\equiv 2(mod\ 3)$if $0\le i\le m-2$ and$(j+2)\not\equiv 2(mod\ 3)$if $i=m-1$. Hence any two vertices of type $e_iv_j$and e_iv_{j+1}does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color classis proved below,

(i) $u_iv_j$and e_iv_j

(ii) $e_iv_j$and $u_{i+1(mod\ m)}v_j$

(iii) $u_iv_j$and u_iv_{j+1}

(iv) $e_iv_j$and $e_{i+1(mod\ m)}v_j$

Case 3: When $m\equiv 2(mod\ 3)$

$$f(u_iv_j)=\begin{cases}(j+2)(mod\ 3)\ ; & i=0\\ (i+j+1)(mod\ 3)\ ; & 1\le i\le m-1\end{cases}$$

$$f(e_iv_j)=(i+j)(mod\ 3)$$

$$\forall\ i=0,1,2,\dots,(m-1) \text{ and } j=0,1,2,\dots,(n-1)$$

Then the vertex set is partitioned into V_0, V_1andV_2 as below

$$V_0=\begin{cases}u_iv_j\ ;\ (j+2)\equiv 0(mod\ 3)\text{if } i=0 \text{ and } (i+j+1)\equiv 0(mod\ 3)\text{if } 1\le i\le m-1\\ e_iv_j\ ; \qquad (i+j)\equiv 0(mod\ 3)\end{cases}$$

$$V_1 = \begin{cases} u_i v_j \ ; \ (j+2) \equiv 1(mod\ 3) \text{if } i = 0 \text{ and } (i+j+1) \equiv 1(mod\ 3) \text{if } 1 \le i \le m-1 \\ e_i v_j \ ; \qquad (i+j) \equiv 1(mod\ 3) \end{cases}$$

$$V_2 = \begin{cases} u_i v_j \ ; \ (j+2) \equiv 2(mod\ 3) \text{if } i = 0 \text{ and } (i+j+1) \equiv 2(mod\ 3) \text{if } 1 \le i \le m-1 \\ e_i v_j \ ; \qquad (i+j) \equiv 2(mod\ 3) \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.3 $u_i v_j, e_i v_j \in V(T_1(C_m)\ P_n)$ are adjacent in $T_1(C_m)\ P_n$ .Consider two vertices of type$u_i v_j$and $e_i v_j$.Clearly, if $(j+2) \equiv 0(mod\ 3)$if $i = 0$ and $(i+j+1) \equiv 0(mod\ 3)$if $1 \le i \le m-1$ then $(i+j) \not\equiv 0(mod\ 3)$, also if $(j+2) \equiv 1(mod\ 3)$if $i = 0$ and $(i+j+1) \equiv 1(mod\ 3)$if $1 \le i \le m-1$ then $(i+j) \not\equiv 1(mod\ 3)$ andif $(j+2) \equiv 2(mod\ 3)$if $i = 0$ and $(i+j+1) \equiv 2(mod\ 3)$if $1 \le i \le m-1$ then $(i+j) \not\equiv 2(mod\ 3)$. Hence any two vertices of type $u_i v_j$and $e_i v_j$does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color class has been shown below,

(i) $e_i v_j$and $e_i v_{j+1}$

(ii) $e_i v_j$and $u_{i+1(mod\ m)} v_j$

(iii) $u_i v_j$and $u_i v_{j+1}$

(iv) $e_i v_j$and $e_{i+1(mod\ m)} v_j$

(a) If $n \equiv 0(mod\ 3)$ then $|V_0| = \frac{2mn}{3}$, $|V_1| = \frac{2mn}{3}$and $|V_2| = \frac{2mn}{3}$.

(b) If $n \equiv 1(mod\ 3)$ then $|V_0| = \frac{2m(n-1)}{3} + \left(m - \left\lceil \frac{m-1}{3} \right\rceil\right)$, $|V_1| = \frac{2m(n-1)}{3} + \left(m - \left\lceil \frac{m}{3} \right\rceil\right)$ and $|V_2| = \frac{2m(n-1)}{3} + \left(m - \left\lceil \frac{m-2}{3} \right\rceil\right)$.

(c) If $n \equiv 2(mod\ 3)$ then $|V_0| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil \frac{2(m-1)}{3} \right\rceil\right)$, $|V_1| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil \frac{2m}{3} \right\rceil\right)$ and $|V_2| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil \frac{2m-1}{3} \right\rceil\right)$.

Obviously, from (a), (b) and (c) f satisfies the condition$\left||V_i| - |V_j|\right| \le 1, i \ne j$. Clearly, f is a proper coloring of $T_1(C_m)\ P_n$and so $\chi_=(T_1(C_m)\ P_n)$ = 3. Hence the theorem holds.

4.3. Product of $T(C_m)$ with P_n

Cartesian product of total graph of cycle with path

The total graph of cycle $T(C_m)$ on $2m$ vertices and $4m$ edges with vertex set and edge set is given by

$V(T(C_m)) = \cup_{i=0}^{m-1}\{u_i; e_i\}$ and $E(T(C_m)) = \left(\cup_{i=0}^{m-1}\{u_iu_{i+1(mod\ m)}; u_ie_i; e_iu_{i+1(mod\ m)}; e_ie_{i+1(mod\ m)}\}\right)$

Theorem 4.3.1 If m and n are non-negative integers, $m \geq 3, n \geq 2$ thenthe equitable chromatic number of $T(C_m)\ P_n$ is 3. The theorem holds only for those m that are multiples of 3.

Proof. Clearly,$|V(T(C_m)\ P_n)| = 2mn$ and $|E(T(C_m)\ P_n)| = 2m(3n-1)$.

Let the vertex set and edge set of the Cartesian product $T(C_m)\ P_n$ be

$$V(T(C_m)\ P_n) = \bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{u_iv_j; e_iv_j\}$$

$$E(T(C_m)\ P_n) \\ = \left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-1}\{(u_iv_j)(u_{i+1(mod\ m)}v_j); (u_iv_j)(e_iv_j); (e_iv_j)(u_{i+1(mod\ m)}v_j); (e_iv_j)(e_{i+1(mod\ m)}v_j)\}\right) \\ \cup\left(\bigcup_{i=0}^{m-1}\bigcup_{j=0}^{n-2}\{(u_iv_j)(u_iv_{j+1}); (e_iv_j)(e_iv_{j+1})\}\right)$$

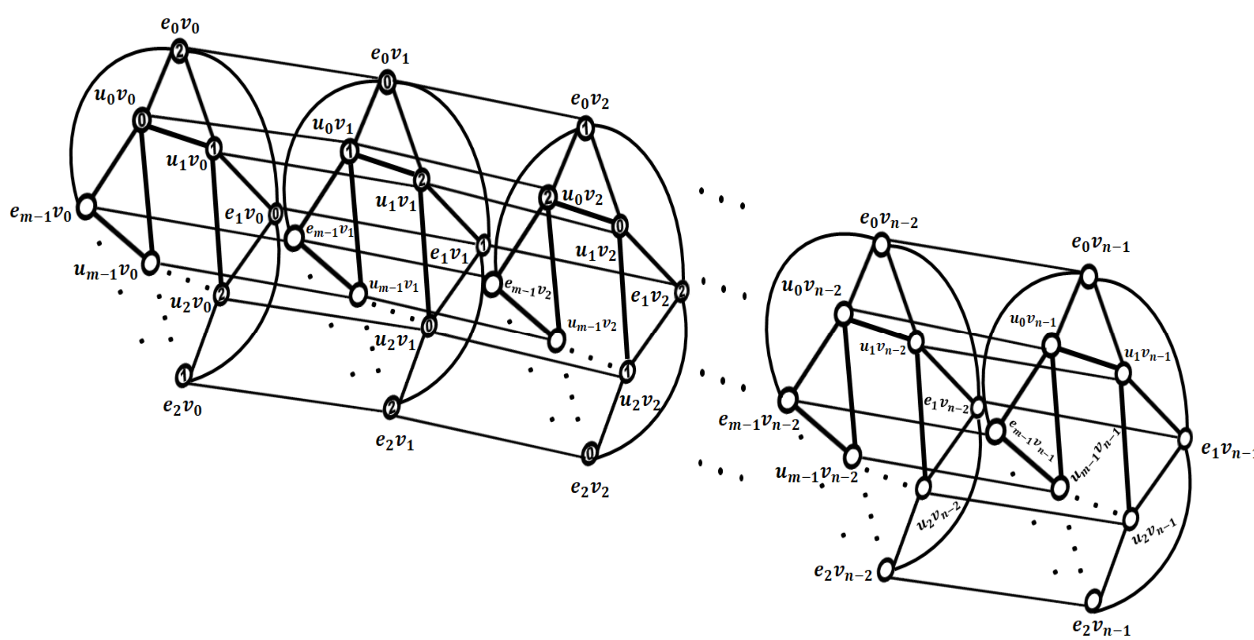


Fig.4 Cartesian product of total graph of cycle with path$T(C_m)\ P_n$,m is a multiple of 3.

Define the coloring using the map$f : V(T(C_m)\ P_n) \rightarrow \{0,1,2\}$ of $T(C_m)\ P_n$ as follows:

$$f(u_iv_j) = (i+j)(mod\ 3)$$

$$f(e_i v_j) = (i + j + 2)(mod\ 3)$$

$\forall\ i = 0,1,2, \dots, (m-1)$ and $j = 0,1,2, \dots, (n$ -1)

where m is a multiple of 3.

Then the vertex set is partitioned into V_0, V_1 and V_2 as below

$$V_0 = \begin{cases} u_i v_j\ ; & (i + j) \equiv 0(mod\ 3) \\ e_i v_j\ ; & (i + j + 2) \equiv 0(mod\ 3) \end{cases}$$

$$V_1 = \begin{cases} u_i v_j\ ; & (i + j) \equiv 1(mod\ 3) \\ e_i v_j\ ; & (i + j + 2) \equiv 1(mod\ 3) \end{cases}$$

$$V_2 = \begin{cases} u_i v_j\ ; & (i + j) \equiv 2(mod\ 3) \\ e_i v_j\ ; & (i + j + 2) \equiv 2(mod\ 3) \end{cases}$$

where the vertices in V_0 receives color 0,V_1 receives color 1 and V_2 receives color 2.

In Fig.3 $u_i v_j, u_i v_{j+1} \in V(T(C_m)\ P_n)$ are adjacent in $T(C_m)\ P_n$ .Consider two vertices of type $u_i v_j$ and $u_i v_{j+1}$. Clearly, if $(i + j) \equiv 0(mod\ 3)$ then $(i + j + 1) \not\equiv 0(mod\ 3)$, also if $(i + j) \equiv 1(mod\ 3)$ then $(i + j + 1) \not\equiv 1(mod\ 3)$ and if $(i + j) \equiv 2(mod\ 3)$ then $(i + j + 1) \not\equiv 2(mod\ 3)$. Hence any two vertices of type $u_i v_j$ and $u_i v_{j+1}$ does not belong to the same color class.

Similarly, the following adjacent edges does not belong to the same color class has been shown below,

(i) $u_i v_j$ and $e_i v_j$

(ii) $e_i v_j$ and $u_{i+1(mod\ m)} v_j$

(iii) $u_i v_j$ and $u_{i+1(mod\ m)} v_j$

(iv) $e_i v_j$ and $e_i v_{j+1}$

(v) $e_i v_j$ and $e_{i+1(mod\ m)} v_j$

(a) If $n \equiv 0(mod\ 3)$ then $|V_0| = \frac{2mn}{3}$, $|V_1| = \frac{2mn}{3}$ and $|V_2| = \frac{2mn}{3}$.

(b) If $n \equiv 1(mod\ 3)$ then $|V_0| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m-1}{3}\right\rceil\right)$, $|V_1| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m}{3}\right\rceil\right)$ and $|V_2| = \frac{2m(n-1)}{3} + \left(m - \left\lceil\frac{m-2}{3}\right\rceil\right)$.

(c) If $n \equiv 2(mod\ 3)$ then $|V_0| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2(m-1)}{3}\right\rceil\right)$, $|V_1| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2m}{3}\right\rceil\right)$ and $|V_2| = \frac{2m(n-2)}{3} + \left(2m - \left\lceil\frac{2m-1}{3}\right\rceil\right)$.

Obviously, from (a), (b) and (c) f satisfies the condition $\left||V_i| - |V_j|\right| \leq 1, i \neq j$. Clearly, f is a proper coloring of $T(C_m)\ P_n$ and so $\chi_{=}(T(C_m)\ P_n) = 3$.

Hence the theorem holds.

Inference:

Note that in the theorem 4.3.1,m holds only for the multiples of 3, otherwise $\chi_{=}(T(C_m)\ \ P_n) \neq 3$ as the adjacent vertices receive the same color in this pattern, defying the definition of coloring.

CONCLUSION:

In this paper, Cartesian product of semi-total point graph, semi-total line graph, total graph of cycle with path has been defined and the acceptance of equitable coloring to those graphs has been proved.The equitable chromatic number for the above-defined graphs has also been calculated.The concept can be extended to different graphs.

REFERENCES

1. *Bor-Liang Chan, Chih-Hung Yen, Equitable Δ-coloring of graphs, Discrete Mathematics 312 (2012) 1512-1517.*
2. *Brooks, R.L., On coloring the nodes of a network. Proc. Camb. Philos. Soc. 37,194-197 (1941).*
3. *Chen B.L., Lih K.W., Wu P.L., Equitable coloring and the maximum degree, Eur. J. Comb. 15, 443-447 (1994).*
4. *Bor-Liang Chen, Ko-Wei Lih, Jing-Ho Yan, Equitable Coloring of Interval Graphs and Products of Graphs, arXiv:0903.1396v1 [math.co] 8 Mar 2009.*
5. *Furmanczyk. H, Equitable Coloring of graphs, in: M. Kubale(ed.), Optymalizacja dyskretna. Modele I metody kolorowania grafow, WNT Warszawa 2002, 72-92 (in polish).*
6. *Furmanczyk. H, Equitable Coloring of graph products, Opuscula Mathematica 26(1) (2006), 31-44.*
7. *Hajnal. A, Szemerdi. E, Proof of a conjecture of Erds, in: Renyi. A, Soc. V.T.(Eds.), Combinatorial Theory and Its Application, Vol. II, in: Colloq. Math. Soc. Jnos Bolyai, Vol. 4, North-Holland, Amsterdam. pp. 601-623 (1970).*
8. *Jensen. T.R, Toft. B, Graph Coloring Problems, Wiley-Interscience, New York (1995).*
9. *Loura Jency and Benedict Michael Raj, Equitable Coloring On Rooted Product Of Graphs Vol.12 No.10(2021), 5956-5959.*
10. *Meyer. W, Equitable coloring, Am. Math. Mon. 80, 920-922(1973).*
11. *E. Sampathkumar and S. B. Chikkodimath, Semi-Total Graphs of a graph-I, Reprinted from The Karnatak University Journal: Science – Vol. XVIII, 1973.*

12. *Vernold Vivin. J, Kaliraj. K, Akbar Ali. M. M, Equitable Coloring on Total Graphs of Bigraphs and Central Graph of Cycles and Paths, International Journal of Mathematics and Mathematical Sciences Volume 2011, Article ID 279246, 5 pages doi:10.1155/2011/279246.*
13. *Wang. W.F, Lih. K.W, Equitable list coloring of graphs, Taiwanese J.Math. 8(4) (2004) 747-759.*
14. *Wu-Hsiung Lin, Gerard J. Chang, Equitable colorings of Cartesian products of graphs, Discrete Applied Mathematics 160 (2012) 239-247.*

DOI No. : 10.5958/2394-9309.2022.00008.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

BAYESIAN ESTIMATION ON ROBUST REGRESSION USING LEAST ABSOLUTE DEVIATIONS METHOD

Eakambaram. S* and Ashok. B**
* Department of Statistics, Periyar Arts College, Cuddalore 607 001
**Department of Community Medicine, Melmaruvathur Adiparasakthi Institute of Medical Sciences and Research Melamaruvathur - 603 319
E-mail : bashokstats@gmail.com

ABSTRACT
Recent days we can obtain easily large sample high dimensional data sets from cheap sensors. However, the growth of variables can prevent us from constructing a parsimonious model which provides good interpretation about the system. One important stream of statistical research requires effective variable selection procedures to improve both accuracy and interpretability of the learning technique. Variable selection is an important research topic in linear regression especially for model selection in high-dimensional data situation. In this paper,it is proposed to develop a procedure for Bayesian estimation based on robust regression by leastabsolute deviations method.
Keywords: Variable selection, Linear Regression, Robust Regression, Least Absolute Deviations.

1. INTRODUCTION

Recent days we can obtain easily large sample high dimensional data sets from cheap sensors. However, the growth of variables can prevent us from constructing a parsimonious model which provides good interpretation about the system. One important stream of statistical research requires effective variable selection procedures to improve both accuracy and interpretability of the learning technique. Variable selection is an important research topic in linear regression especially for model selection in high-dimensional data situation.

Tibshirani (1996) proposed the least absolute shrinkage and selection operator (LASSO), which can simultaneously select valuable covariates and estimate regression parameters. Traditional model selection criteria such as Akaike Information Criterion (AIC) [Akaike, 1973] and Bayesian Information Criterion (BIC) [Schwarz, 1978] have major drawbacks that parameter estimation and model selection are two separate processes. The LASSO is a regularisation with $l1$-type penalization and it becomes extremely popular, because it shrinks the regression coefficients toward zero with the possibility of setting some coefficients equal to zero, resulting in a simultaneous estimation and variable selection.

Many literatures show successful applications using the LASSO. However, the LASSO can be biased for the coefficients whose absolute values are large. Fan & Li (2001) proposed a penalized regression with the Smoothly Clipped Absolute Deviation (SCAD) penalty function and showed that it has better theoretical properties than the LASSO with $L1$-type penalty. The penalized regression with SCAD not only selects important covariates consistently but also produces parameter estimators as efficient as if the true model were known, i.e., the oracle property. The LASSO does not satisfy the oracle property. The SCAD function is a non-convex penalty function to make up for the deficiencies of the LASSO.

The penalized regression consists of a loss function and a penalty function. In traditional regression setting it is well known that the least squares method is sensitive to even single outlier. Alternative to the least squares method is the Least Absolute Deviation (LAD) method. Wang et al., (2007) developed a robust algorithm with the LAD loss function and $L1$-type penalty (LAD-L1). They showed that the LAD-L1 is resistant to non-normal error distributions and outliers. Jung proposed a robust method with the LAD loss function and the SCAD penalty (LAD-SCAD) estimate. It shows that the SCAD penalty function is more efficient than the LAD-L1.

Recently statisticians often treat data sets with a non-normal response variable or covariates that may contain multiple outliers or leverage points. Even though the LAD is more robust than the least squares method, the unbounded loss function of the LAD affects strongly the LAD estimator. In this paper we consider a weight method for the bounded loss function. The weight in our algorithm attenuates the influence of outliers on the estimator and works the same influence of non-outliers as the un-weighted LAD-SCAD method. The weighted LAD loss function with the SCAD penalty function (WLAD-SCAD) improves the error performance of the LAD-SCAD. The proposed method combines the robustness of the weighted LAD and the oracle property of the SCAD penalty. The tuning parameter controls the model complexity and plays an important role in the variable selection procedure. We propose a BIC-type tuning parameter selector using a data-driven method, the WLAD-SCAD with the BIC tuning parameter can identify the most parsimonious correct model.

In this paper it is proposed to develop a procedure for Bayesian estimation based on robust regression by least absolute deviations method for a detailed study on Bayesian robust regression method. The existing results can be validated by comparing Akaike Information Criterion (AIC); Bayesian Information Criterion (BIC); Least Absolute Deviation (LAD); LAD and the $L1$-type penalty (LAD-L1); LAD and the SCAD penalty (LAD-SCAD); Least Absolute Shrinkage and Selection Operator (LASSO); Local Linear Approximation (LLA); Local Quadratic Approximation (LQA); Least Squares Estimator (LSE); Smoothly Clipped Absolute Deviation (SCAD); the Weighted LAD and the SCAD penalty (WLAD-SCAD).

2. THE MODEL

Consider the linear regression model

$y_i=\alpha+x_i^T\beta+\varepsilon_i$,i=1,$\cdots$, n,

where $x_i=(x_{i1},\cdots,x_i)^T$ is the d-dimensional covariate, $\beta=(\beta_1,\cdots,\beta_d)^T$, d is the number of covariates and n is the number of observations. Let $y=(y_1,\cdots,y_n)^T$and let X be a $n\times(d+1)$ matrix whose i-th row $(1,x_i^T)$. The

Least Squares Estimator (LSE) $(XTX)^{-1}X^{T}y$ which minimizes the sum of squares of residuals can be distorted by the heady-tailed probability distribution of errors or even single outlier. The criterion can be written by

$$\sum_{i=1}^{n}(y_i - (\alpha + x_i^T\beta))^2.$$

One alternative to the Least Squares method is the LAD method which minimises the criterion function, the sum of absolute deviations of the errors

$$\sum_{i=1}^{n}|y_i - (\alpha + x_i^T\beta)|.$$

The major advantage of the LAD method lies in its robustness relative to the LSE. The LAD estimates are less affected by the presence of a few outliers or influential observations. However, both the LSE and the LAD cannot be useful for model selection when especially the number of covariates is very large.

Tibshirani (1996) proposed a penalty based on the $L1$ norm $|\beta j|dj$=1 for automatically deleting unnecessary covariates. The LASSO criterion is a simply penalized least squares with the $L1$ penalty

$$\sum_{i=1}^{n}(y_i - (\alpha + x_i^T\beta))^2 + \lambda\sum_{j=1}^{d}|\beta_j|, \qquad (1)$$

where λ>0 is the tuning parameter which controls the trade-off between model fitting and model sparsity. When the tuning parameter is large, the criterion focuses on model sparsity. Traditionally in model selection, cross-validation and information criteria-including the AIC [Akaike, 1973] and BIC [Schwarz, 1978]-are widely applied. Shao (1997) showed that the BIC can identify the true model consistently in linear regression with fixed dimensional covariates, but the AIC may fail due to over-fitting. Yang (2005) showed that cross-validation is asymptotically equivalent to the AIC and so they behave similarly.

LASSO is not asymptotically consistent and so the LASSO can be biased for large coefficients. Fan & Li (2001) addressed this problem and proposed the SCAD penalty function. They described the conditions of a good penalty function (a) unbiasedness: the resulting estimator is nearly unbiased when the true unknown parameter is large; (b) sparsity: the resulting estimator is a thresholding rule, which automatically sets small estimated coefficients to be zero; (c) continuity: the resulting estimator is continuous in the data. The LSE with the SCAD penalty function minimizes the criterion function

$$\sum_{i=1}^{n}(y_i - (\alpha + x_i^T\beta))^2 + \sum_{j=1}^{d}p_\lambda(|\beta_j|), \qquad (2)$$

where

$$p_\lambda(|\beta|) = \begin{cases} \lambda|\beta|, & if\ 0 \le |\beta| < \lambda \\ \dfrac{(a^2-1)\lambda^2 - (|\beta| - a\lambda)^2}{2(a-1)}, & if\ \lambda \le |\beta| < a\lambda \\ \dfrac{1}{2}(a+1)^2\lambda^2, & if\ |\beta| > a\lambda \end{cases}$$

and so its derivative becomes

$$p'_\lambda(|\beta|) = \begin{cases} \lambda, & if\ 0 \le |\beta| < \lambda \\ \dfrac{a\lambda - |\beta|}{a-1}, & if\ \lambda \le |\beta| < a\lambda \\ 0, & if\ |\beta| > a\lambda \end{cases}$$

where a can be chosen using cross-validation or generalized cross-validation. However, the simulation of Fan & Li (2001) gives us a=3.7 which is approximately optimal. In this article we set a=3.7.

Similar to that the LAD is more robust than the LSE in non-penalized regression model, Wang et al., (2007) proposed the LAD with the SCAD penalty which minimises the criterion function

$$\sum_{i=1}^{n} | y_i - (\alpha + x_i^T \beta)| + \sum_{j=1}^{d} p_\lambda(|\beta_j|). \qquad (3)$$

Even though the LAD is robust, its breakdown point is also $1/n$ which is equivalent to the LSE. Weighted LAD with the SCAD penalty, because the simulation results of Giloni et al., (2006) show that in non-penalized linear regression the performance of the weighted LAD estimator is competitive with that of high breakdown regression estimators, particularly in the presence of outliers located at leverage points. A weighted LAD penalized estimator which combines the weighted LAD estimator with the $L1$ penalty and it used the criterion function

$$\sum_{i=1}^{n} w_i | y_i - (\alpha + x_i^T \beta)| + \sum_{j=1}^{d} p_\lambda(|\beta_j|), \qquad (4)$$

where the weight wi depends on the space of covariates. We call the solution of (4) the weighted LAD with the SCAD (WLAD-SCAD). The proposed method uses the weight for robustness which is resistant to leverage points or influential observations, because the weight reduces the effects of the observations having large deviations or the leverage points. The objective function will give the robustness of weight methods and the advantages of the SCAD penalized function.

3. METHODS

The solution of (4) can be obtained by a standard optimization program if the criterion function is convex and differentiable. Unfortunately, the absolute function in (4) is not differentiable at zero and the SCAD penalty function is not convex in β. Approximation of the absolute function and the SCAD function transforms the objective function into linear equations and so we can obtain efficiently an iterative solution of the WLAD-SCAD estimator.

The absolute function

$$|u| \approx \frac{u^2}{2|u^0|} + \frac{1}{2|u^0|}$$

for nonzero u^0 near u gives

$$|y_i - (\alpha + x_i^T \beta)| \approx \frac{(y_i - (\alpha + x_i^T \beta))^2}{2|y_i - (\alpha^0 + x_i^T \beta^0)|} + \frac{1}{2}|y_i - (\alpha_0 + x_i^T \beta_0)|, \quad (5)$$

for initial values α_0, β_0 near the minimization of (3). Also, the Taylor expansion at the non-zero βj_0 yields

$$p_\lambda(|\beta_j|) \approx p_\lambda(|\beta_j^0|) + \frac{p'_\lambda(|\beta_j^0|)}{2|\beta_j^0|}(\beta_j^2 - \beta_j^{02}), \quad (6)$$

and we set β_j=0 if β_j^0 is near zero. Assume that the log-likelihood function is smooth with respect to β and its first two partial derivatives are continuous. The linear quadratic approximation (LQA) is a modification of the Newton-Raphson algorithm [Fan & Li, 2001]. The LQA is broadly useful for solving optimization problems with the non-differentiable criterion function. Then the criterion function (4) becomes

$$\frac{1}{2n}\sum_{i=1}^{n} w_i \frac{(y_i - (\alpha + x_i^T \beta))^2}{|y_i - (\alpha^0 + x_i^T \beta^0)|} + \frac{1}{2n}\sum_{i=1}^{n} w_i |y_i - (\alpha^0 + x_i^T \beta^0)| + \sum_{j=1}^{d}[p_\lambda(|\beta_j^0|) + \frac{p'_\lambda(|\beta_j^0|)}{2|\beta_i^0|}(\beta_j^2 - \beta_j^{02})]. \quad (7)$$

and we obtain the criterion function with up to constants

$$\frac{1}{n}(y - \tilde{X}\tilde{\beta})^T W (y - \tilde{X}\tilde{\beta}) + \frac{1}{2}\beta^T Q \beta, \quad (8)$$

where $\tilde{\beta} = (\alpha, \beta^T)^T, W = \mathrm{diag}\left(\frac{w_i}{|y_i - (\alpha^0 + x_i^T \beta^0)|}\right)$

4. RESULTS

Based on the methods sited above in section (3) the results are given in the following table with exponentiated exponential distribution.

Table 1: Simulation results with exponentiated exponential without outliers

Exponentiated Exponential Distribution	Methods	LASSO	SCAD	LAD-SCAD	WLAD-SCAD	BAYESIAN
n = 350 k=1, α=2, λ=2.5	AMAD	0.510 (0.084)	0.412 (0.076)	0.389 (0.076)	0.280 (0.077)	0.167 (0.067)
n = 550 k=1, α=2.5, λ=3	AMAD	0.410 (0.384)	0.312 (0.376)	0.289 (0.376)	0.180 (0.377)	0.077 (0.367)
n = 750 k=1, α=3, λ=3.5	AMAD	0.271 (0.284)	0.263 (0.276)	0.298 (0.266)	0.260 (0.277)	0.217 (0.267)

5. CONCLUSION

A robust algorithm was proposed for the penalized regression model with the LAD loss function and the SCAD penalty function. We used a weight function for the robust loss function and improved the effectiveness of the proposed algorithm through numerical simulations and derived two approximations for objective functions to treat the non-convex optimization problem. One is LLA and the other is LQA. Since the former is linear, it is so easy to implement it. Both two methods are robust to outliers. The numerical simulations shows that the proposed method is more robust than other methods from the point view of finding exact non-zero coefficients. Thus, the proposed method gives a method of variable selection from thousands of input variables appeared in the applications of biometrical experiments and for the further study considers the Huber function for the loss function with the SCAD penalty function.

REFERENCE

1. ***Akaike, H.*** *(1973). Information theory and the maximum likelihood principle in* 2nd International Symposium on Information Theory.*University of California, Los Angeles*
2. ***Fan, J. and Li, R.*** *(2001). Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. JASA, 96, pp. 1348-1360.*
3. ***Gilono, A., Simonoff, J. F and Sengupta, B.*** *(2006).Robust Weighted LAD Regression. Journal of Computational Statistics & Data Analysis, 50(11), pp. 3124-3140.*
4. ***Tibshirani, R.*** *(1996).Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical SocietySeries B (Methodological), pp. 267-288.*
5. ***Wang, Z., Wu, Y.H. and Zhao, L.C.*** *(2007). Change Point Estimation for Censored Regression Model. Science in China Series A: Mathematics,**50**(1). pp. 63-72.*
6. ***Schwarz, G.*** *(1978). Estimating a Dimension of a Model. The Annals of Statistics, Vol 6, No 2, pp. 461-464.*
7. ***Shao, J.*** *(1997). An Asymptotic Theory for Linear Model Selection. StatisticaSinica, 7, pp. 221-264.*
8. ***Yang L.****(2005). Log-Linear Models for Word Alignment. In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), pages 459–466, Ann Arbor, Michigan. Association for Computational Linguistics.*

DOI No. : 10.5958/2394-9309.2022.00009.9
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

COMPOSITION OF NEUTROSOPHIC SOFT MATRICES

M. Kavitha*, P. Murugadas**
*Department of Mathematics, Bharath Institute of Higher Education and Research, Chennai, India
**Department of Mathematics, Government Arts and Science College, veerapandi, India
E-mail : kavithakathir3@gmail.com*, bodi_muruga@yahoo.com**

ABSTRACT

In this paper we extend the results on fuzzy matrices under max-productcomposition to Neutrosophic Soft Matrix (NSM). We stress the significance of NSM-theoretical approach and show how several previous results for matricesof special types can be obtained in a united way. We show how to constructanidempotent NSM from a given on ethrough the max-product composition of NSMs.

2010AMSSubjectClassification:Primary03E72;Secondary15B15.

Keywords : Neutrosophic Soft Set (NSS), Neutrosophic Soft Matrix(NSM), Neutrosophic Soft Relation (NSR) Idempotent Neutrosophic Soft Matrix (INSM)

1. INTRODUCTION

The concept of fuzzy sets was introduced by Zadeh(1965). Since then the fuzzy sets and fuzzy logic have been applied in many real life problems in uncertain, am-biguous environment.The traditional fuzzy sets is characterized by the member ship value or the grade of membership value.Some times it may be very difficult to assign the membership value for a fuzzy sets. Conse quently the concept of intervalvalued fuzzy sets was proposed (Turksen1986) to capture the uncertainty of grade of membership value. In some real life problems in expert system,belief system information fusion and so on , we must consider the truth-membership as well asthe falsity-membership for proper description of an object in uncertain, ambiguous environment. Neither the fuzzy sets nor the inter valvalued fuzzy setsis appropriate for such asituation.Intuitionistic fuzzysets (IFS)introduced by Atanassov (1986) is appropriate for such a situation. But the intuitionistic fuzzy sets can only han-dle the incomplete information considering both the truth-membership (or simplymembership) and falsity-membership (or non-membership) values.It does nothandle the indeterminate and inconsistent information which exists in belief sys-tem. Smarandache (2005) introduced the concept of neutrosophicset (NS) which isa mathematical tool for handling problems involving imprecise, indeterminacy and in consistent data.

In our regular everyday life we face situations which require procedure sallow-ing certain flexibility in information processing capacity. Soft set theory (Molodtsov1999; Maji et al.2002, 2003) addressed those problems successfully. In their early work soft set wa sdescribed purely asa mathematical method to model uncertain-ties.There searcher scan pick any kind of parameter sof any nature they wish in order to facilitate the decision making procedure as there is a varied way of picturingthe objects.Maji et al.(2002, 2003) have done further research on soft set theory. Presence of vagueness demanded Fuzzy Soft Set(FSS) (Majietal.2001;Basuetal.2012) to come into picture.

But satis factory evaluation nof membershipvaluesis not always possible because of the insufficiency in the available information (be-sides the presence of vagueness) situation. Evaluation of non-membership values isalso not always possible for the same reason and as a result there exists an in deterministic partupon which he sitation survives.Certainly fuzzy soft set theory is not suitable to solve such problems. In those situations Intuitionistic Fuzzy Soft Set theory (IFSS) (Majietal.2001) may be more applicable. No win the perlance of soft set theory there is hardly any limitation to select the nature of the criteria and asmost of the parameter sorcriteria (which are words or sentences) are neutron sophic in nature,Maji (2012,2013) has been motivated to combine the concept of soft setand neutrosophic set to make the new mathematical model neutron sophic soft set and has give nanalg orithmto solve a decision making problem.

Deli (2014) defined neutron sophic parameter ized neutron sophic soft sets (npn-soft sets) which is the combination of neutrosophic set and a soft set.Deli and Broumi(2014) redefined the notion of neutrosophic set in a new way and put forward the concept of neutron sophic softmatrix and different ty pesofmatrice sinneutro sophic soft theory.They have in traduced some one woperations and propertie sonthese matrices.The minimal solution wasd one by Kavithaet.al,(2017)based on the notion of FNSM given by Sumathi and Arokiarani (2013,2014). As the time goes a some works on FNSM were done by Kavithaet.al,(2017,2018,2020).The Mono-to neinter valfuzzy neutron sophic softeigen problem and Monotone fuzzy neutron sophics of teigens pacest ructures in max-minalge brawere investigated by Murugadaset.al, (2019).Also, two kinds of fuzzy neutron sophics of tmatric espresented byUmaet.al,.

In this paper we study and prove some properties of the product-max composition of neutron sophics of tmatrices. However, this composition is studied inadual way to the max-product composition and it is useful for studying the complement of neutrosophic softrelations. Also we define near lyirreflexive and nearly constant neutron sophics of tmatrices and prove some properties of it.One of the seresults enablesusto construct anidempotent neutron sophics of tmatrixf romagive none and this is The main result in the paper

2 PRELIMINARIES

This area essentially depicts Neutrosophic Set (NS), Fuzzy Neutro sophic Soft Set (FNSS),Fuzzy Neutro sophic Soft Matrix (NSFM) and FNSMs of type-I, Type-II.For the fundamental definitions and outline see([11,12,13,14,15,16,17,21]

3. Main Result

Definition 1:

Let A =$\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle$ and D =$\langle d_{ij}^{T}, d_{ij}^{I}, d_{ij}^{F}\rangle$ are NSM of order m× n. Now here, some special types of NSMs define as,

$$A \ominus G=\begin{cases}\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle \; if \; \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle > \langle d_{ij}^{T}, d_{ij}^{I}, d_{ij}^{F}\rangle \\ \langle 0,0,1\rangle if \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle \leq \langle d_{ij}^{T}, d_{ij}^{I}, d_{ij}^{F}\rangle\end{cases}$$

$$A^{\alpha} =\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle^{\langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle}=\begin{cases}\langle 1,1,0\rangle \quad if \; \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle \geq \langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle \\ \langle 0,0,1\rangle if \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle < \langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle\end{cases}$$

$$A_{\alpha}=\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle_{\langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle}=\begin{cases}\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle \quad if \; \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle \geq \langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle \\ \langle 0,0,1\rangle if \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle < \langle \alpha_{ij}^{T}, \alpha_{ij}^{I}, \alpha_{ij}^{F}\rangle\end{cases}$$

Definition 2:

ForanySquareNeutrosophicSoftMatrix(SNSM)$R=\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ *we have*

(a) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$*is*istransitiveIfandonlyif $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{2} \leq \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$

(b) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$isaidempodentifandonlyif $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{2} = \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$

(c) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$isreflexiveifandonlyif $\langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle = \langle 1,1,0\rangle$ for all iϵ {1,2, …,n}

(d) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$isirreflexiveifandonlyif $\langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle = \langle 1,1,0\rangle$ for all iϵ {1,2, …,n}

(e) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$is a nearly irreflexiveif and only if $\langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle \leq \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ for all iϵ {1,2, …,n}

(f) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$is asymmetricifandonlyif $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle' = \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$.

Definition 3: For any NSM $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle_{m\times n}$ is constant if and only if $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle$=$\langle a_{kj}^{T}, a_{kj}^{I}, a_{kj}^{F}\rangle$ for every I,k ϵ {1,2, …,m} $j\epsilon$ {1,2, …,n} , $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle$ is nearly constant if and only if $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle$=$\langle a_{kj}^{T}, a_{kj}^{I}, a_{kj}^{F}\rangle$ where I $\neq j$ k $\neq$j.

Proposition 1: Let $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle_{m\times n}$be symmetricandnearlyirreflexiveNSM.

Thenwehavethefollowing:

(a) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{*} \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle \leq \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ that is $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ is max-min transitive,

(b) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{*} \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$symmetricandnearlyirreflexive,

(c) $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{2}$ is weakly reflexive.

Proof:

(a) Suppose $\langle s_{ij}^{T}, s_{ij}^{I}, s_{ij}^{F}\rangle$ =$\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle^{*} \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$. Then

$\langle s_{ij}^{T}, s_{ij}^{I}, s_{ij}^{F}\rangle \prod_{k=1}^{n}\langle r_{ik}^{T}, r_{ik}^{I}, r_{ik}^{F}\rangle + \langle r_{kj}^{T}, r_{kj}^{I}, r_{kj}^{F}\rangle \leq \langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle + \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle = \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$

so that $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle * \langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle \leq \langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle$.

(b) We have as in (a)

$\langle s_{ji}^T, s_{ji}^I, s_{ji}^F\rangle = \prod_{k=1}^n \langle r_{jk}^T, r_{jk}^I, r_{jk}^F\rangle + \langle r_{ki}^T, r_{ki}^I, r_{ki}^F\rangle = \prod_{k=1}^n \langle r_{kj}^T, r_{kj}^I, r_{kj}^F\rangle + \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle$

$=\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle$, so that $\langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$ is symmetric.

Also $\langle s_{ii}^T, s_{ii}^I, s_{ii}^F\rangle = \prod_{k=1}^n \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle + \langle r_{ki}^T, r_{ki}^I, r_{ki}^F\rangle = \prod_{k=1}^n \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle$

$=\langle r_{ii}^T, r_{ii}^I, r_{ii}^F\rangle \leq \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle + \langle r_{kj}^T, r_{kj}^I, r_{kj}^F\rangle = \langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$ and $\langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$ is thus nearly irreflexive.

(c) Let $\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle = \langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle^2$. Then

$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle = \sum_{k=1}^n \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle \langle r_{kj}^T, r_{kj}^I, r_{kj}^F\rangle = \langle r_{ih}^T, r_{ij}^I, r_{ih}^F\rangle \langle r_{hj}^T, r_{hj}^I, r_{hj}^F\rangle$

for some h ϵ {1,2,…,n}

$\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle = \sum_{k=1}^n \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle = \sum_{k=1}^n \langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle = \langle t_{il}^T, t_{il}^I, t_{il}^F\rangle$

for some l ϵ {1,2,…,n}

$\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle = \langle r_{il}^T, r_{il}^I, r_{il}^F\rangle \geq \langle r_{ih}^T, r_{ih}^I, r_{ih}^F\rangle \geq \langle r_{ih}^T, r_{ih}^I, r_{ih}^F\rangle \langle r_{hj}^T, r_{ih}^I j, r_{hj}^F\rangle = \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$

that is the NSM $\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle = \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle^2$ is weakly reflexive.

Proposition 2: For NSMs$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle_{m\times n}$ and $\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{m\times n} \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle_{n\times l}$ and $\langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle_{p\times m}$ we have the following:

(a) $(\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle * \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle)' = \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle' \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle'$;

$(b) If\ \langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle \leq \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$then$\langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle * \langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle$

$\leq \langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$and$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle$

$\leq \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle * \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle$

Proof: Let $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle' \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle'$ and

$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle = \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle * \langle c_{ij}^T, c_{ij}^I, c_{ij}^F\rangle$. Then

$\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \prod_{k=1}^n (\langle c_{ki}^T, c_{ki}^I, c_{ki}^F\rangle + \langle b_{jk}^T, b_{jk}^I, b_{ij}^F k\rangle)$,

$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle = \prod_{k=1}^n (\langle b_{jk}^T, b_{jk}^I, b_{ij}^F k\rangle + \langle c_{ki}^T, c_{ki}^I, c_{ki}^F\rangle)$,

That is $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle t_{jk}^T, t_{jk}^I, t_{jk}^F\rangle$and $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle'$.

(b) Let $\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle = \langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle * \langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle$ and

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = \langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$

$$\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle = \prod_{k=1}^n \langle d_{ik}^T, d_{ik}^I, d_{ik}^F\rangle + \langle a_{kj}^T, a_{kj}^I, a_{kj}^F\rangle$$

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = \prod_{k=1}^n \langle d_{ik}^T, d_{ik}^I, d_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle$

Since we have $\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle \leq \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$ we get

$$\langle d_{ik}^T, d_{ik}^I, d_{ik}^F\rangle + \langle a_{kj}^T, a_{kj}^I, a_{kj}^F\rangle \leq \langle d_{ik}^T, d_{ik}^I, d_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle$$

for every k ϵ {1,2,…,m}

Therefore $\prod_{k=1}^{n}\langle d_{ik}^{T}, d_{ik}^{I}, d_{ik}^{F}\rangle + \langle a_{kj}^{T}, a_{kj}^{I}, a_{kj}^{F}\rangle \leq \prod_{k=1}^{n}\langle d_{ik}^{T}, d_{ik}^{I}, d_{ik}^{F}\rangle + \langle b_{kj}^{T}, b_{kj}^{I}, b_{kj}^{F}\rangle$.

Similarly, we can shown that $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle * \langle c_{ij}^{T}, c_{ij}^{I}, c_{ij}^{F}\rangle \leq \langle b_{ij}^{T}, b_{ij}^{I}, b_{ij}^{F}\rangle * \langle c_{ij}^{T}, c_{ij}^{I}, c_{ij}^{F}\rangle$.

Theorem 1: For any NSM $\langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle, \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle * \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle'$ is nearly reflexive and symmetric.

Proof: Let $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle = \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle * \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle'$ that is

$\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle = \prod_{k=1}^{n}\langle a_{ik}^{T}, a_{ik}^{I}, a_{ik}^{F}\rangle + \langle a_{jk}^{T}, a_{jk}^{I}, a_{jk}^{F}\rangle' = \langle a_{il}^{T}, a_{il}^{I}, a_{il}^{F}\rangle + \langle a_{jl}^{T}, a_{jl}^{I}, a_{jl}^{F}\rangle'$

for some $1 \in \{1,2,\ldots,m\}$ and

$\langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle = \prod_{k=1}^{n}\langle a_{ik}^{T}, a_{ik}^{I}, a_{ik}^{F}\rangle + \langle a_{ik}^{T}, a_{ik}^{I}, a_{ik}^{F}\rangle' = \prod_{k=1}^{n}\langle a_{ik}^{T}, a_{ik}^{I}, a_{ik}^{F}\rangle = \langle a_{ih}^{T}, a_{ih}^{I}, a_{ih}^{F}\rangle$

for some $h \in \{1,2,\ldots,m\}$.

Thus we have $\langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle = \langle a_{ih}^{T}, a_{ih}^{I}, a_{ih}^{F}\rangle \leq \langle a_{il}^{T}, a_{il}^{I}, a_{il}^{F}\rangle \leq \langle a_{il}^{T}, a_{il}^{I}, a_{il}^{F}\rangle + \langle a_{jl}^{T}, a_{jl}^{I}, a_{jl}^{F}\rangle' = \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$.

Therefore, $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle = \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle * \langle a_{ij}^{T}, a_{ij}^{I}, a_{ij}^{F}\rangle'$ is nearly irreflexive.

The symmetric of $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$is obvious.

Theorem 3: Let $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle_{n\times n}$ be a symmetric and nearly irreflexive NSM. Then the NSM $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle = \langle b_{ij}^{T}, b_{ij}^{I}, b_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ is idempotent and nearly constant where the NSM $\langle b_{ij}^{T}, b_{ij}^{I}, b_{ij}^{F}\rangle$is obtained from the identity NSM I_n by multiplying each i^{th} row by an element $\langle a_{i}^{T}, a_{i}^{I}, a_{i}^{F}\rangle \epsilon [0,1]$.

Proof: Based on definition of the NSMs $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle, \langle b_{ij}^{T}, b_{ij}^{I}, b_{ij}^{F}\rangle$and $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$, we can write the elements of $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$ in terms of that $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ and the constants $\langle a_{i}^{T}, a_{i}^{I}, a_{i}^{F}\rangle$as follows:

$$\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle = \begin{cases} \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle \; if \;\; i \neq j, \\ \langle a_{i}^{T}, a_{i}^{I}, a_{i}^{F}\rangle \; if \; i = j \; and \; \langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle \\ \prod_{k\neq i}\langle r_{ik}^{T}, r_{ik}^{I}, r_{ik}^{F}\rangle \; if \; i = j \; and \; \leq \; \prod_{k\neq i}\langle r_{ik}^{T}, r_{ik}^{I}, r_{ik}^{F}\rangle \; \leq \; \langle a_{i}^{T}, a_{i}^{I}, a_{i}^{F}\rangle \; , \\ \langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle \; if \; i = j \; and \; \langle a_{i}^{T}, a_{i}^{I}, a_{i}^{F}\rangle \; \leq \; \langle r_{ii}^{T}, r_{ii}^{I}, r_{ii}^{F}\rangle. \end{cases}$$

The elements of $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle^2$ is calculated as

$\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle^2 = \sum_{k=1}^{n}\langle t_{ik}^{T}, t_{ik}^{I}, t_{ik}^{F}\rangle\langle t_{kj}^{T}, t_{kj}^{I}, t_{kj}^{F}\rangle = \langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle$

for some $h \in \{1,2,\ldots,m\}$.

Now we have several cases based on the definition of $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$.

Case 1: $i \neq j$, $h \neq j$. In this case we have

$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle = \langle t_{hh}^{T}, t_{hh}^{I}, t_{hh}^{F}\rangle\langle t_{jj}^{T}, t_{jj}^{I}, t_{jj}^{F}\rangle \geq \langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle\langle t_{jj}^{T}, t_{jj}^{I}, t_{jj}^{F}\rangle$.

But

$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle\langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$

$$= \begin{cases} \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\langle a_j^T, a_j^I, a_j^F\rangle = \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\ if\ \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle \le \langle a_j^T, a_j^I, a_j^F\rangle \le \prod_{k\neq j}\langle r_{jk}^T, r_{jk}^I, r_{jk}^F\rangle, \\ \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\left(\prod_{k\neq j}\langle r_{jk}^T, r_{jk}^I, r_{jk}^F\rangle,\right) = \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\ if\ \prod_{k\neq j}\langle r_{jk}^T, r_{jk}^I, r_{jk}^F\rangle, \le \langle a_j^T, a_j^I, a_j^F\rangle \\ \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle = \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle\ if\ \langle a_j^T, a_j^I, a_j^F\rangle \le \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle. \end{cases}$$

Thus, we have $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{jh}^T, t_{jh}^I, t_{jh}^F\rangle$=$\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$.

However, this implies $\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$.

Therefore, $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle$= $\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$=$\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle = \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$.

Case 2:i= $h \neq j$ and $\langle r_{ii}^T, r_{ii}^I, r_{ii}^F\rangle \le \langle a_i^T, a_i^I, a_i^F\rangle \le \prod_{k\neq i}\langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle$.

In this case, $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle = \langle a_i^T, a_i^I, a_i^F\rangle\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$.

Since we have i=h, $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle \ge \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle\langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$.

But this implies $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle \ge \langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$ and hence $\langle a_i^T, a_i^I, a_i^F\rangle \ge \langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$.

By the definition of $\langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$ *it is clear* that $\langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$.

Thus, $\langle a_i^T, a_i^I, a_i^F\rangle \ge \langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$.

Therefore, $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle$ =$\langle a_i^T, a_i^I, a_i^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$

$= \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$= $\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$

Case 3: i= h$\neq j$ and $\prod_{k\neq i}\langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle \le \langle a_i^T, a_i^I, a_i^F\rangle$.As in case 2, $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle \ge \langle t_{jj}^T, t_{jj}^I, t_{jj}^F\rangle$. But we have $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle = \prod_{k\neq i}\langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle$, so that

$$\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle = \left(\prod_{k\neq i}\langle r_{ik}^T, r_{ik}^I, r_{ik}^F\rangle\right)\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$$

$$= \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle = \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle.$$

Case 4: i= h$\neq j$ and $\langle a_i^T, a_i^I, a_i^F\rangle \le \langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle$.

Since $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle \ge \langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle \ge \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$ and $\langle t_{ii}^T, t_{ii}^I, t_{ii}^F\rangle = \langle r_{ii}^T, r_{ii}^I, r_{ii}^F\rangle$ we have $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle = \langle r_{ii}^T, r_{ii}^I, r_{ii}^F\rangle\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$

$= \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$=$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$

Case 5: i= h$\neq j$ and $\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle \le \langle a_h^T, a_h^I, a_h^F\rangle \le \prod_{k\neq i}\langle r_{hk}^T, r_{hk}^I, r_{hk}^F\rangle$. In this case we get $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle$=$\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle = \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$=$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$.

Case 6: i$\neq$ h= j and $\prod_{k\neq h}\langle r_{hk}^T, r_{hk}^I, r_{hk}^F\rangle \le \langle a_h^T, a_h^I, a_h^F\rangle$. In this case we get $\langle t_{ih}^T, t_{ih}^I, t_{ih}^F\rangle\langle t_{hj}^T, t_{hj}^I, t_{hj}^F\rangle$=$\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle(\prod_{k\neq h}\langle r_{hk}^T, r_{hk}^I, r_{hk}^F\rangle)$=$\langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle$

$= \langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle$=$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$

Case 7: i$\neq$ h= j and $\langle r_{jj}^T, r_{jj}^I, r_{jj}^F\rangle \le \langle r_{hh}^T, r_{hh}^I, r_{hh}^F\rangle$.We have in this case,

$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle$=$\langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle\langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle$=$\langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle$

$= \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle$ =$\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$

Case 8: i= h= $j\,.We\ have\ in\ this\ case$

$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle$=$\langle t_{ii}^{T}, t_{ii}^{I}, t_{ii}^{F}\rangle\langle t_{ii}^{T}, t_{ii}^{I}, t_{ii}^{F}\rangle$=$\langle t_{ii}^{T}, t_{ii}^{I}, t_{ii}^{F}\rangle = \langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$

Case 9: $i = j \neq h$. In this case we have

$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle = \langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle\langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \geq \langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle\langle t_{jj}^{T}, t_{jj}^{I}, t_{jj}^{F}\rangle$.

But we have

$$\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle\langle t_{jj}^{T}, t_{jj}^{I}, t_{jj}^{F}\rangle = \begin{cases} \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle\langle a_{j}^{T}, a_{j}^{I}, a_{j}^{F}\rangle = \langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle \ if\ \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \leq \langle a_{j}^{T}, a_{j}^{I}, a_{j}^{F}\rangle \leq \prod_{k\neq j}\langle r_{jk}^{T}, r_{jk}^{I}, r_{jk}^{F}\rangle \\ \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle\left(\prod_{k\neq j}\langle r_{jk}^{T}, r_{jk}^{I}, r_{jk}^{F}\rangle\right) = \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \ if\ \prod_{k\neq j}\langle r_{jk}^{T}, r_{jk}^{I}, r_{jk}^{F}\rangle \leq \langle a_{j}^{T}, a_{j}^{I}, a_{j}^{F}\rangle \\ \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle\langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle = \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \ if\ \langle a_{j}^{T}, a_{j}^{I}, a_{j}^{F}\rangle \leq \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \end{cases}$$ Th

erefore,

$$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle = \langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle\langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle \geq \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle.$$

However, this implies $\langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle \geq \langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle$ and so

$\langle t_{ih}^{T}, t_{ih}^{I}, t_{ih}^{F}\rangle\langle t_{hj}^{T}, t_{hj}^{I}, t_{hj}^{F}\rangle$=$\langle r_{hh}^{T}, r_{hh}^{I}, r_{hh}^{F}\rangle\langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle$= $\langle r_{jj}^{T}, r_{jj}^{I}, r_{jj}^{F}\rangle = \langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$.

From the above computations of $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle^2$=$\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$ in all cases and this means that $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle is\ idempotent.$

By the definition of $\langle t_{ij}^{T}, t_{ij}^{I}, t_{ij}^{F}\rangle$ is nearly constant. This completes the proof of the theorem.

The following corollaries are important is studying fuzzy relations (NSMs). However, they enable us to construct an idempotent Neutrosophic Soft Relation (NSR) from given one.

Corollary 1. Let $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle_{n\times n}$ NSM. Then we have

(a) I_m*($\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle'$) is a idempotent and nearly cinstant,

(b) $(\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle')^2$ is weakly reflexive.

Proof: By Theorem 1 and Theorem 2 and Proposition 1.

Corollary 2: Let $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle_{n\times n}$ NSM. Then we have $(\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle')$* I_m.

Proof: We have $((\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle') * I_m)' = I_m$*$\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle'$.

Then the above corollary the NSM $((\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle') * I_m)'$ is also idempotent.

But $((\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle'))'$= $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle'$. Thus, the NSM $((\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle') * I_m.)'$is idempotent and so $\langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle * \langle r_{ij}^{T}, r_{ij}^{I}, r_{ij}^{F}\rangle' * I_m.$ is also idempotent.

Example 1: $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle =$

$$\begin{bmatrix} 0.5,0.4,0.5 & 0.7,0.6,0.3 & 0.8,0.7,0.2 & 0.9,0.8,0.1 & 0.6,0.5,0.4 \\ 0.4,0.3,0.6 & 0.3,0.2,0.7 & 0.7,0.6,0.3 & 0.2,0.1,0.8 & 0.8,0.7,0.2 \\ 0.6,0.5,0.4 & 0.5,0.4,0.5 & 0.4,0.3,0.6 & 1,1,0 & 0.3,0.2,0.7 \\ 0.9,0.8,0.1 & 0.2,0.3,0.8 & 0.6,0.5,0.4 & 0.4,0.5,0.6 & 0.7,0.6,0.3 \end{bmatrix}$$

Then $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle'=$

$$\begin{bmatrix} 0.5,0.4,0.5 & 0.4,0.3,0.6 & 0.6,0.5,0.4 & 0.9,0.8,0.1 \\ 0.7,0.6,0.3 & 0.3,0.2,0.7 & 0.5,0.4,0.5 & 0.2,0.1,0.8 \\ 0.8,0.7,0.2 & 0.7,0.6,0.3 & 0.4,0.3,0.6 & 0.6,0.5,0.4 \\ 0.9,0.8,0.1 & 0.2,0.3,0.8 & 1,1,0 & 0.4,0.5,0.6 \\ 0.6,0.5,0.4 & 0.8,0.7,0.2 & 0.3,0.2,0.7 & 0.7,0.6,0.3 \end{bmatrix}$$

$\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle * \langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle'=$

$$\begin{bmatrix} 0.5,0.4,0.5 & 0.5,0.4,0.5 & 0.6,0.5,0.4 & 0.7,0.6,0.3 \\ 0.5,0.4,0.5 & 0.2,0.1,0.8 & 0.5,0.4,0.5 & 0.3,0.2,0.7 \\ 0.6,0.5,0.4 & 0.5,0.4,0.5 & 0.3,0.2,0.7 & 0.5,0.4,0.5 \\ 0.7,0.6,0.3 & 0.3,0.2,0.7 & 0.5,0.4,0.5 & 0.2,0.1,0.8 \end{bmatrix}$$

It is clear that $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle * \langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle'$ is nearly irreflexive and symmetric

Let $\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle=$

$$\begin{bmatrix} 0.7,0.6,0.3 & 0,0,1 & 0,0,1 & 0,0,1 \\ 0,0,1 & 0.4,0.3,0.6 & 0,0,1 & 0,0,1 \\ 0,0,1 & 0,0,1 & 0.5,0.4,0.5 & 0,0,1 \\ 0,0,1 & 0,0,1 & 0,0,1 & 0.2,0.1,0.8 \end{bmatrix}$$

$\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle=\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle * (\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle * \langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle^1)=$

$$\begin{bmatrix} 0.5,0.4,0.5 & 0.2,0.1,0.8 & 0.3,0.2,0.7 & 0.2,0.3,0.8 \\ 0.5,0.4,0.5 & 0.3,0.2,0.7 & 0.3,0.2,0.7 & 0.2,0.1,0.1 \\ 0.5,0.4,0.5 & 0.2,0.1,0.8 & 0.5,0.4,0.5 & 0.2,0.1,0.1 \\ 0.5,0.4,0.5 & 0.2,0.1,0.8 & 0.3,0.2,0.7 & 0.2,0.1,0.8 \end{bmatrix}.$$

and it is easy to see that the NSM $\langle t_{ij}^T, t_{ij}^I, t_{ij}^F\rangle$ is idempotent and nearly constant.

Proposition 3: For any NSM $\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle_{n\times m}$ and $\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{n\times l}$ we have

(a) $(\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle)^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}= \langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle} *$

$\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}$

(b) $(\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle)_{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}= \langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle_{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle} * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}$

Proof: (a) Suppose $\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle=\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$,

$\langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle = \langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}$,

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle=\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle} * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle}$

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle=\prod_{k=1}^{n}(\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle} * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle^{\langle\alpha_{ij}^T,\alpha_{ij}^I,\alpha_{ij}^F\rangle})$.

If we take $\langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle = 1$ so that $\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

that is $\prod_{k=1}^{n}\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle$=($\langle a_{ih}^T, a_{ih}^I, a_{ih}^F\rangle * \langle b_{hj}^T, b_{hj}^I, b_{hj}^F\rangle$)

$\geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$for some h$\epsilon$ {1,2,3,…,n}.

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle$=$\prod_{k=1}^{n}\left(\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}\right)$

= $\langle a_{is}^T, a_{is}^I, a_{is}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{sj}^T, b_{sj}^I, b_{sj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$ for some s ϵ {1,2,3,…,n}.

$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = 0$, then $\langle a_{is}^T, a_{is}^I, a_{is}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{sj}^T, b_{sj}^I, b_{sj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$=0

That is, $\langle a_{is}^T, a_{is}^I, a_{is}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} = 0 \; and \; \langle b_{sj}^T, b_{sj}^I, b_{sj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} = 0$

so that $\langle a_{is}^T, a_{is}^I, a_{is}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$ and $\langle b_{sj}^T, b_{sj}^I, b_{sj}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$.

Thus we have $\langle a_{is}^T, a_{is}^I, a_{is}^F\rangle$+$\langle b_{sj}^T, b_{sj}^I, b_{sj}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

which contradictions $\langle a_{ih}^T, a_{ih}^I, a_{ih}^F\rangle * \langle b_{hj}^T, b_{hj}^I, b_{hj}^F\rangle) \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$,

therefore$\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = 1$

Moreover if we put $\langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle$=0, that is

$\langle a_{ih}^T, a_{ih}^I, a_{ih}^F\rangle + \langle b_{hj}^T, b_{hj}^I, b_{hj}^F\rangle) \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$,

then$\langle a_{ih}^T, a_{ih}^I, a_{ih}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$ and $\langle b_{hj}^T, b_{hj}^I, b_{hj}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

and so that $\langle a_{ih}^T, a_{ih}^I, a_{ih}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$=$\langle b_{hj}^T, b_{hj}^I, b_{hj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$=0.

Since $\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = \prod_{k=1}^{n}\left(\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}\right)$

we get $\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = \langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle^{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} = 0$.

Therefore $\langle g_{ij}^T, g_{ij}^I, g_{ij}^F\rangle = \langle d_{ij}^T, d_{ij}^I, d_{ij}^F\rangle$.

(b) If we assume that ($\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle$)=$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle$,

$\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$and

$\langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$=$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$.

That is $\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$,

$\langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$=$\prod_{k=1}^{n}\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle} + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle_{\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle}$

If $\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$, $e, i, .\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

for some k ϵ $\{1,2, \dots, n\}$,

then$\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$ or $\langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle \geq \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

or both for every k so that

$\langle a_{ik\alpha}^T, a_{ik\alpha}^I, a_{ik\alpha}^F\rangle + \langle b_{kj\alpha}^T, b_{kj\alpha}^I, b_{kj\alpha}^F\rangle$=$\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle$,

therefore$\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle$=$\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle = \langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle$.

If $\langle f_{ij}^T, f_{ij}^I, f_{ij}^F\rangle < \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$, that is $\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle + \langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle \le \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$

for some k ϵ $\{1,2,\dots,n\}$ then $\langle a_{ik}^T, a_{ik}^I, a_{ik}^F\rangle$<$\langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$ or

$\langle b_{kj}^T, b_{kj}^I, b_{kj}^F\rangle < \langle \alpha_{ij}^T, \alpha_{ij}^I, \alpha_{ij}^F\rangle$so that $\langle a_{ik\alpha}^T, a_{ik\alpha}^I, a_{ik\alpha}^F\rangle + \langle b_{kj\alpha}^T, b_{kj\alpha}^I, b_{kj\alpha}^F\rangle$=0,

hence$\langle r_{ij}^T, r_{ij}^I, r_{ij}^F\rangle = \langle s_{ij}^T, s_{ij}^I, s_{ij}^F\rangle = 0$.

Example 2: Let $\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle$=$\begin{bmatrix} 0.7,0.6,0.3 & 0.5,0.4,0.5 & 0.3,0.2,0.7 & 0.4,0.3,0.6 \\ 0.2,0.1,0.8 & 0.9,0.8,0.1 & 0.7,0.6,0.3 & 0.8,0.7,0.2 \\ 1,1,0 & 0.8,0.7,0.2 & 0.3,0.2,0.7 & 0.6,0.5,0.4 \end{bmatrix}$

$$\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle = \begin{bmatrix} 0.5,0.4,0.5 & 0.7,0.6,0.3 \\ 0.6,0.5,0.4 & 0.4,0.3,0.6 \\ 0.2,0.1,0.8 & 0.8,0.7,0.2 \\ 0.7,0.6,0.3 & 0.5,0.4,0.5 \end{bmatrix}$$

Then $\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle$*$\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{(0.6,0.5,0.4)}$=

$$\begin{bmatrix} 0.3,0.2,0.7 & 0.5,0.4,0.5 \\ 0.5,0.4,0.5 & 0.7,0.6,0.3 \\ 0.3,0.2,0.7 & 0.6,0.5,0.4 \end{bmatrix}_{(0.6,0.5,0.4)} = \begin{bmatrix} 0,0,1 & 0,0,1 \\ 0,0,1 & 0.7,0.6,0.3 \\ 0,0,1 & 0.6,0.5,0.4 \end{bmatrix}$$

$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle_{(0.6,0.5,0.4)}$*$\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle_{(0.6,0.5,0.4)}$=

$$\begin{bmatrix} 0.7,0.6,0.3 & 0,0,1 & 0,0,1 & 0,0,1 \\ 0,0,1 & 0.9,0.8,0.1 & 0.7,0.6,0.3 & 0.8,0.7,0.2 \\ 1,1,0 & 0.8,0.7,0.2 & 0,0,1 & 0.6,0.5,0.4 \end{bmatrix} * \begin{bmatrix} 0,0,1 & 0.7,0.6,0.3 \\ 0.6,0.5,0.4 & 0,0,1 \\ 0,0,1 & 0.8,0.7,0.2 \\ 0.7,0.6,0.3 & 0,0,1 \end{bmatrix} = \begin{bmatrix} 0,0,1 & 0,0,1 \\ 0,0,1 & 0.7,0.6,0.3 \\ 0,0,1 & 0.6,0.5,0.4 \end{bmatrix},$$

and$(\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle * \langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle)^{(0.6,0.5,0.4)}$=$\begin{bmatrix} 0,0,1 & 0,0,1 \\ 0,0,1 & 1,1,0 \\ 0,0,1 & 1,1,0 \end{bmatrix}$,

$\langle a_{ij}^T, a_{ij}^I, a_{ij}^F\rangle^{(0.6,0.5,0.4)}$*$\langle b_{ij}^T, b_{ij}^I, b_{ij}^F\rangle^{(0.6,0.5,0.4)}$=$\begin{bmatrix} 0,0,1 & 0,0,1 \\ 0,0,1 & 1,1,0 \\ 0,0,1 & 1,1,0 \end{bmatrix}$.

REFERENCES:

1. *I. Arockiarani and I. R. Sumathi,A Fuzzy Neutrosophic Soft Matrix Approach in Decision Making, JGRMA, 2(2), (2014), 14-23.*

2. *I. Arockiarani, I. R. Sumathi and Martina Jency, Fuzzy Neutrosophic Soft Topological Spaces, International Journal of Mathematical Archive, 4 (10) (2013), 225-238.*
3. *K.Atanassov., Intuitionistic Fuzzy Sets, Fuzzy Sets and System, 20 (1983), 87-96.*
4. *D. Dubois and H. Prade., Fuzzy Sets and Systems, Theory and Application (Academic Press, New York, 1980).*
5. *Z.T. Fan and De-Fu Liu., On the Power Sequence of a Fuzzy Matrix, I Convergent Power Sequence, Sumitted.*
6. *Z.T. Fan and De-Fu Liu., On the power sequence of a Fuzzy Matrix, II Oscillanting Power Sequence, Sumitted.*
7. *Z.T. Fan and De-Fu Liu., On the power sequence of a Fuzzy Matrix, III The Maximum Principle of the Second order, in preparation.*
8. *H. Hashimoto., Convergence of Powes of a Fuzzy Transitive Matrix, Fuzzy Transitive Matrix, Fuzzy Sets and Systems 9(1983), 153-160.*
9. *I. Deli, S. Broumi., Neutrosophic Soft Sets and Neutrosophic Soft Matrices Based on Decision Making, Journal of Intelligent and Fuzzy Systems, 28(5), (2014),1-28.*
10. *J.X. Li,. Some Results on the Nilpotent Fuzzy Matrices, Fuzzy Systems Math, 1 (1989), 52-55.*
11. *A. Kandel., Fuzzy Mathematical Techniques with Apllications (Addison-Wesley, Tokyo, 1986).*
12. *M. Kavitha, P. Murugadas and S. Sriram., Minimal Solution of Fuzzy Neutrosophic Soft Matrix, Journal of Linear and Topological Algebra, 6, (2017),171-189.*
13. *M. Kavitha, P. Murugadas and S. Sriram., On the _-Robustness of Fuzzy Neutrosophic Soft Matrix, International Journal of Fuzzy Mathematical Archive, 2, (2017), 267-287.*
14. *M. Kavitha, P. Murugadas and S. Sriram., Computing the Greatest Xeigenvector of Fuzzy Neutrosophic Soft Matrix, Internation Journal of Mathamatics and its Applications, 4, (2017), 893-907.*
15. *M. Kavitha, P. Murugadas and S. Sriram., On the Power of Fuzzy Neutrosophic Soft Matrix, Journal of Linear and Topological Algebra, 7, (2018), 133-147.*
16. *M. Kavitha., A Study on Fuzzy Neutrosophic Soft Matrices with Related Aspects on Eigenvectors of the Matrices, (2018).*
17. *M. Kavitha, P. Murugadas and S. Sriram., Priodicity of Interval Fuzzy Neutrosophic Soft Matrices, Advances in Mathematics: Scientific Journal, 9, (2020), 1661-1670.*
18. *A. Kaufmann, Introduction to the Theory of Fuzzy Subsets, Academic Press, New York, (1975).*
19. *Li Jain-xin, Controllable Fuzzy Matrices, Fuzzy Sets and Systems, 45, (1992), 313-319.*

20. *P. Murugadas, M. Kavitha, S. Sriram., Monotone Interval Fuzzy Neutrosophic Soft Eigenproblem, Malaya Journal of Matematik, S(1), (2019), 342-350.*
21. *P. Murugadas, M. Kavitha, S. Sriram., Monotone Fuzzy Neutrosophic Soft Eigenspace Structure in Max-Min Algebra, Recent Trends in Pure and Applied Mathematics, 2177, 020048(1-8), (2019).*
22. *P. Murugadas, M. Kavitha., Solvability of System of Neutrosophic Soft Linear Equations, Neutrosophic Sets and System, 40,(2021).*
23. *P. Murugadas, M. Kavitha., Convergence of Fuzzy Neutrosophic Soft Circulant Matrices, Journal of Physics: Conference Series, 40,(2021).*
24. *P. K. Maji, A. R. Roy and R. Biswas., On Intuitionistic Fuzzy Soft Set, InternationalJournal of Fuzzy Mathematics, 12(3), (2004), 669-684.*
25. *P. K. Maji., ANeutrosophic Soft Set Approach to a Decision Making Problem, Annauls Fuzzy Mathematics Information, 3(2), (2012), 313-319.*
26. *P. K. Maji, R. Biswas and A. R. Roy., Fuzzy Soft Sets, The journal of FuzzyMathematics, 9(3), (2001), 589-602.*
27. *D. Molodtsov., Soft Set Theory Tirst Results, Computer & Mathematics withApplications, 37 (4-5), 1999, 19-31.*
28. *M. A. Mishref and E.G. Emam., Transitivity and Subinverses in Fuzzy Matrices, Fuzzy Sets and System, 52, 337-343, (1992).*
29. *M. Z. Ragab and E. G Emam., The Determinant and Adjoint of a Square FuzzyMatrix, Fuzzy Sets and Systems 61, 297-307, (1994).*
30. *F. Smarandach., Neutrosophic Set a Generalization of the Intuitionistic Fuzzy Set, International Journal of Pure and Applied Mathematics, 24 (2005), 287-297.*
31. *Thomasson., Convergence of powers of a Fuzzy Matrix, Journal of Mathematic Annauls Application, 57, (1977), 476-480.*
32. *R. Uma, S. Sriram and P. Murugadas., Fuzzy Neutrosophic Soft Matrices of Type-I and Type-II, Fuzzy Information and Engineering, Volume 13 (2), (2021).*
33. *J. Verba, General Decomposition Problem of Fuzzy Relations, Fuzzy Sets and Systems 54, 69-79, (1993).*
34. *Zhu Reiying., Convergence of Powers of a Fuzzy Symmetric Matrix, Fuzzy mathematics,3 (1984), 45-49.*
35. *L. A. Zadeh, Fuzzy Sets, Information and Control, 8, (1965), 338-353.*

DOI No. : 10.5958/2394-9309.2022.00010.5
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

CONTROL CHARTS BASED ON NON-NORMAL WITH SPECIAL REFERENCE TO AILAMUJIA DISTRIBUTION

Joshua Daniel Egan L*, D. Venkatesan**
Department of Statistics, Annamalai University, India
E-mail: joshdanie85@gmail.com*, sit_san@hotmail.com**

ABSTRACT

Statistical Process control (SPC) is very useful in production processes in controlling the quality of products. Products are usually produced with a defined variance and nominal value, but still unexpected impacts do occur, and these might lead to changes in the production process. Control charts are hence used to detect the change in the process and to minimize any false alarms especially when a process is in control mode. Shewhart ControlCharts are often used to monitor processes when the quality of interest follows is assumed to followa normal distribution but in practice, it is not always true. In this paper, the use of a non-normal distribution method, namely size biased Ailamujia distribution, which is a mixture of exponential distribution and gamma distribution, has been considered to monitor the process.

Keywords : Control Charts, Size Biased Distribution, Mixture distribution, Ailamujia distribution

1. INTRODUCTION

With variation in production where two products manufactured can't be the exact replica of the other, the SPC is introduced. One can say that the process is "under control" when there are small deviations that are not easily identified due to random causes, and the process is "out of control" when these deviations are noticeable and the causes active.

During process monitoring, the products are sent to the system and wait for the inspection. Good process control also protects the product as it is an excellent opportunity to control batches at permissible quality levels of the product which depends on the variation of the production process but this variation in the process may be due to some controllable and uncontrollable factors. The control charts were introduced by Prof. Shewhart Walter in 1920s and have now become an important tool in quality improvement. Control charts are applied for monitoring a process during the manufacturing of a product. Timely action about the process should be taken based on control charts, that is, no action is taken when the analysis of a control chart shows that the process is in control. However, if the control chart declares that the process has shifted, then appropriate action should be taken to bring the process back under control. Therefore, a control chart indicates the right time or to which observation corrective action should be taken. The two control limits, namely the upper control limits (UCL) and the lower control limit (LCL), are used to monitor the process mean or variance. These limits are very helpful for reducing defective products and alternately for increasing the profits of industries. The maintenance of quality through control charts brings a good reputation for an industry in the market.

For decades, several researchers have developed different types of control charts. Nelson (1984) proposed the Shewhart control chart-tests for special causes. Santiago and Smith (2013) have proposed the control charts, called the t-chart, when the time between events follows the exponential distribution. They used the variable transformation proposed by Nelson (1994) to transform exponentially distributed data to an approximate normal data. Aslam*et al*. (2014) proposed a new control charts for the exponential distribution using the transformed variable and the repetitive sampling. Amin and Venkatesan (2017) have proposed the SPC using the length biased weighted Garima distribution. Amin and Venkatesan (2019) have discussed on the recent developments in control charts techniques. Amin and Venkatesan (2019) have proposed the SPC using transmuted generalized uniform distribution. Control charts are designed on the assumption that the quantitative trait of interest follows the normal distribution, which is not always the case in practice. The variable of interest may follow some non-normal distribution such as an exponential distribution or a gamma distribution or any other distribution. The use of control charts designed for a normal distribution may not be workable in this situation and may result in an increasing of the number of non-conforming products. In addition, the normal distribution is applied in situations where data are collected in subgroups, so that the central limit theorem can be applied when designing the control charts. Again, in practice, it is not always possible to collect data in groups. In many practical situations, classical distributions therefore offer insufficient adjustment for real data. For example, if the data are asymmetric, the normal distribution is not a good option. That's why different generators have been proposed based on one or more parameters to generate new distributions.

The concept of weighted distribution was presented for the first time by Fisher (1934) where the impact of inspection methods on distribution of recorded observations was examined where it is characterized into two types, that is, length biased and size biased distribution. Warren (1975) was the first to apply the size biased distributions in connection with sampling wood cells. Jeffrey (2003) reviewed some of the more recent results on size-biased distributions pertaining to parameter estimation in forestry. Rather*et al.* (2018) have proposed size biased distribution with applications in engineering and medical science. Ailamujia distribution is a lifetime distribution which was introduced by Lv*et al.* (2002) in the engineering application. Several distributions can be used to monitor the production process, and in this paper, the control limits are derived and control chart is monitored using the size biased Ailamujia distribution. The organization of this paper is as follows.

In section 2 the description of the distribution is given. The Performance measures of the size biased Ailamujia distribution is provided in section 3. Section 4provides Control limits using the size biased Ailamujia distribution. A numerical example is given in section 5, and section 6 provides the Conclusion.

2. DESCRIPTION OF THE DISTRIBUTION

Definition: A non-negative random variable X is said to have a size biased distribution if its probability density function (pdf) is given by

$$f_{sb}(x) = \frac{x\, f(x)}{E(x)} \;; \qquad x > 0 \qquad (1)$$

Definition 2: IfX isa non-negative random variable with probability density function $f(x)$, then the probability density function of the Size Biased Ailamujiadistribution,which is a mixture of exponential and gamma distributions, can be expressed bythe equation 1 as:

$$f_{sb}(x;\lambda) = \frac{x\, f(x;\lambda)}{E(x)}$$

where; $E(x) = 4\lambda^2 \int_0^\infty x^2 e^{-2\lambda x} dx = \frac{1}{\lambda}$

Therefore,

$$f_{sb}(x;\lambda) = 4\lambda^3 x^2 e^{-2\lambda x} \qquad x > 0,\ \lambda > 0 \qquad (2)$$

The corresponding cdf of size biased Ailamujia distribution is obtained as:

$$F_{sb}(x;\lambda) = \int_0^x 4\lambda^3 x^2 e^{-2\lambda x} dx$$

Therefore, after simplifications one can get, $F_{sb}(x;\lambda) = \frac{1}{2}\gamma(3, 2\lambda x)$

Where, $\lambda$is positive parameter and $\gamma(s,x) = \int_0^x z^{s-1} e^{-z} dz$ is a lower incomplete gamma function.

3. PERFORMANCE MEASURES OF THE SIZE BIASED AILAMUJIA DISTRIBUTION

From the above size biased Ailamujia density function (pdf) with parameterλ, then one can obtain the r^{th} order moment $E(X^r)$ of a size biased Ailamujiadistribution, thus,

$$\mu_r' = \int_0^\infty x^r f_{sb}(x;\lambda) dx$$

$$= \int_0^\infty x^r\, 4\lambda^3 x^2 e^{-2\lambda x} dx$$

Therefore,

$$\mu_r' = \frac{\Gamma(r+3)}{2^{r+1}\lambda^r} \qquad (3)$$

Then, putting r=1, 2 in equation 3, one can get the first and second moment which are,

$$E(X) = \mu_1' = \frac{3}{2\lambda} \qquad (4)$$

Then the second moment is given by

$$E\left(X^2\right)=\mu_2{'}=\frac{3}{\lambda} \tag{5}$$

Therefore, from equations (4) and (5), the variance $V(X)$ is given by

$$V(X)=E\left(X^2\right)-[E(X)]^2$$

Therefore,

$$V(X)=\frac{3}{4\lambda^2} \tag{6}$$

4. CONTROL LIMITS USING THE SIZE BIASED AILAMUJIA DISTRIBUTION

This control chart has several values and control limit sets,then, when the process is under control, almost all points are within the upper control (UCL) and lower (LCL) limits [Duncan (1986) and Montgomery (2012)]. Therefore, from equations (4) and (6), the control limits are given by

$$UCL=\frac{3}{2\lambda}+\frac{3}{2\lambda}\sqrt{3}$$

$$Centre\ Line\ (CL)=\frac{3}{2\lambda}$$

$$LCL=\frac{3}{2\lambda}-\frac{3}{2\lambda}\sqrt{3}$$

Where; $\lambda>0$

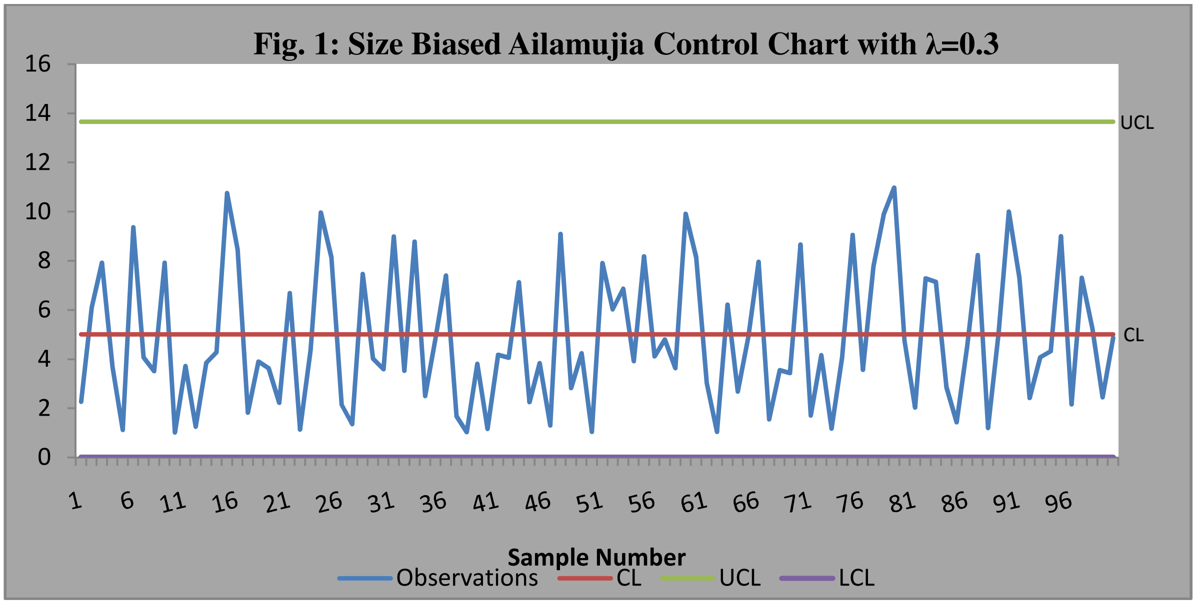
5. NUMERICAL ILLUSTRATION

The construction of control limits is considered, in this example, for illustrating the applications of the proposed method. The control limits of the size biased Ailamujia distribution are obtained for simulated data set. For parameter λbeing random variable the Table 1 is constructed. All the generated samples are reported in Table 1.

Table 1: Control limits using size biased Ailamujia Distribution

λ	CL	UCL	LCL
0.1	15	40.98	0
0.3	5	13.66	0
0.5	3	8.20	0
1	2.14	5.85	0
5	1.5	4.10	0
7	0.5	1.37	0
10	0.3	0.82	0
15	0.15	0.41	0

It is observed that from Table 1, the deviation of the control limits decreases whenever the parameter λ increases. The size biased Ailamujia Control Chart is shown in Fig. 1 for $\lambda = 0.3$.



It is also pointed out that, the observations of the process control must be with parameter $\lambda > 0$, otherwise the process will be out of control. That is, depends on the manufacturing products, the manufacturing engineers should fix the parameters value λ based on what type of data they are working with. As one can see on Table 1, the more the parameters values increase, the smaller control limits.

6. CONCLUSION

In this paper, the process control has been developed using mixture of distributions namely, asize biased Ailamujiadistribution. A novel algorithm has been developed for sentencing the process while manufacturing. The Control limits are given using the size biased Ailamujia distribution with different values of parameterλ. Table is constructed to help to the selection of the parameter based on the type of data the manufacturing engineer is faced with. The control chart is drawn by considering the parameterλ = 0.3 where all observations are showing to be in control.

REFERENCES

1. *A. J. Duncan (1986), "Quality Control and Industrial Statistics", 5th ed., Irwin, Homewood, IL.*
2. *Aafaq A. Rather and C. Subramanian (2018), "Length Biased Sushila Distribution", Universal Review Journal, Vol. 7, No. 12, pp. 1010-1023.*
3. *D. C. Montgomery (2012), "Introduction to Statistical Quality Control", 7th ed., Wiley, New York.*
4. *Eduardo Santiago & Joel Smith (2013), "Control Charts Based on the Exponential Distribution: Adapting Runs Rules for the t Chart", Quality Engineering, Vol. 25, No. 2, pp. 85-96.*
5. *H. Q.Lv,L. H. Gao, C. L. Chen (2002),"Эрланга distribution and its application in supportability data analysis",Journal of Academy of Armored Force Engineering, Vol. 16,No. 3,pp. 48-52.*
6. *Jeffrey H. Gove (2003), "Estimation and Applications of size biased distributions in forestry", In: Amaro, A; Reed, D.; Soares, P., eds. Modelling forest systems. Cambridge, MA: CABI,pp. 201-212.*

7. *L. S. Nelson (1984). "The Shewhart control chart-tests for special causes", Journal of Quality Technology, Vol. 16 pp. 238-239.*
8. *L. S. Nelson (1994), "A control chart for parts-per-million nonconforming items", Journal of Quality Technology, Vol. 26 pp. 239-240.*
9. *M. Aslam, M. Azam and C. H. Jun (2014), "A new exponentially weighted moving average sign chart using repetitive sampling", Journal of Process Control, Vol. 24, No.7, pp. 1149-1153.*
10. *R. Khatree (1989), "Characterization of Inverse-Gaussian and Gamma distributions through their length biased distribution", IEEE Transactions on Reliability, Vol. 38, pp. 610-611.*
11. *Shaka A. Amin and D. Venkatesan (2019), "Recent Developments in Control Charts Techniques", Universal Review Journal, Vol. 8, No. 4, pp. 746-756.*
12. *Shaka A. Amin and D. Venkatesan (2019), "SPC using the Length Biased Weighted Garima Distribution", Science, Technology and Development, Vol. 8, No. 9, pp. 01-08.*
13. *Shaka A. Amin and D. Venkatesan (2019), "SPC Using Transmuted Generalized Uniform Distribution", Journal of Information and Computational Science, Vol. 9, No. 8, pp. 332-339.*
14. *W. G. Warren (1975), "Statistical Distributions in Forestry and Forest Products Research", A Modern Course on Statistical Distributions in Scientific Work, Springer, Dordrecht, pp. 369-384.*

DOI No. : 10.5958/2394-9309.2022.00011.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Online) : 2394-9309

HARMONIC MEAN CORDIAL LABELING OF SOME LADDER RELATED GRAPHS

N. G. Gidheeba* and A.Vijayan**

*Research Scholar (Part time), Register Number :18233112092024, Nesamony Memorial Christian College, Marthandam, Kanyakumari District, Tamil Nadu- 629165.

**Research Department of Mathematics, Nesamony Memorial Christian College, Marthandam, Kanyakumari District, Tamil Nadu -629165

[Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012, Tamil Nadu, India]

ABSTRACT

Let G be a graph with vertex set V(G) and edge set E(G) and f be a function defined on vertex set from V(G) →{ 0,1,2 }. For each edge uv of G the possibility is to assign the label $\left\lceil \frac{2f(u)\,f(v)}{f(u)+f(v)} \right\rceil$ ***then f is called a harmonic mean cordial labeling if*** $|v_f(i) - v_f(j)| \leq 1$ ***and*** $|e_f(i) - e_f(j)| \leq 1$, ***where*** $v_f(x)$ ***and*** $e_f(x)$ ***denote the cardinality of vertices and edges labeled with x,*** $x \in \{0,1,2\}$ ***respectively. A graph which admits harmonic mean cordial labeling is called harmonic mean cordial graph. In this paper, we investigate about the existence of root cube mean cordiality of some family of graphs such as, Ladder Graphs, Open Ladder Graphs, Triangular Ladder Graphs, Open triangular Graphs, Slanting Ladder graph***

Keywords: Ladder Graphs, Open Ladder Graphs, Triangular Ladder Graphs, Open Triangular Graphs, Slanting Ladder Graph,, Harmonic Mean Cordial Labeling, Harmonic Mean Cordial Graphs.

1. INTRODUCTION

Here, the graphs take into account are finite, undirected and simple. The vertex set and edge set of a graph G are denoted by $V(G)$ and $E(G)$ respectively. For many applications such as coding theory, X- ray crystallography, astronomy, circuit design etc. the edges or vertices given labels are greatly beneficial in that domain. The concept of cordial labeling was first introduced by Cahit in the year 1987. In this paper, by the motivation of cordial and mean cordial labeling we introduce a new concept, harmonic mean cordial labeling. Also we investigate the cordiality behavior of some standard graphs. Motivated by the works of many researchers in the area of cordial labeling, we introduced a new type of labeling called root cube mean cordial labeling. In this paper we have discussed about the cordiality some ladder related graphs.

2. PRELIMINARIES

Definition 2.1

The Ladder graphL_n is defined by $L_n = P_n \times P_2$ on 2n vertices and $3n-2$ edges with vertex set $V(G) = \{u_i, v_i : 1 \leq i \leq n\}$ and edge set $E(G) = \{u_i u_{i+1}, v_i v_{i+1} 1 \leq i \leq n\} \cup \{u_i v_i, 1 \leq i \leq n\}$ where P_n is a path with n vertices and X denotes the Cartesian product

Definition 2.2

The Open ladder graph $O(L_n), n \geq 2$ is a graph obtained from two paths with $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and $E(G) = \{u_i u_{i+1}, v_i v_{i+1}: 1 \leq i \leq n-1\} \cup \{u_i v_i: 2 \leq i \leq n-1\}$

Definition 2.3

The slanting ladder $S(L_n)$ is the graph obtained from two paths $u_i u_2 \dots u_n$ and $v_1 v_2 \dots v_n$ with vertex set $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and edge set $E(G) = \{u_i u_{i+1}, v_i v_{i+1}: 1 \leq i \leq n-1\} \cup \{v_i u_{i+1}: 1 \leq i \leq n-1\}$.

Definition 2.4

The triangular ladder graph, $TL_n, n \geq 2$ is a graph obtained from the ladder graph by adding the edges $u_i v_{i+1}, 1 \leq i \leq n-1$ to L_n where L_n is the graph with vertex set $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and edge set $E(G) = \{u_i u_{i+1}, v_i v_{i+1}, 1 \leq i \leq n-1\} \cup \{u_i v_i: 1 \leq i \leq n\}$

Definition 2.5

An open Triangular ladder graph, $O(TL_n)$ is a graph obtained from open ladder graph by adding the edges$u_i v_{i+1}$ for $1 \leq i \leq n-1$

3. MAIN RESULTS

Theorem 3.1

The Ladder graph L_n is a harmonic mean cordial graph

Proof

Let $G = L_n$ be a graph with vertex set $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and $E(G) = \{u_i u_{i+1}, v_i v_{i+1}: 1 \leq i \leq n-1\} \cup \{u_i v_i: 1 \leq i \leq n\}$

Define $f : V(L_n) \rightarrow \{0,1,2\}$ as follows.

Case (i) $n \equiv 2(mod\ 3)$

Let $n = 3t + 2$

Define$f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t+1$

$f(u_{2t+1+i}) = 2, 1 \leq i \leq t$

$f(u_n) = 0$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t+1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t+1$

Here $v_f(0) = 2t+1$,$v_f(1) = 2t+2$,$v_f(2) = 2t+1$

$e_f(0) = 3t+2$,$e_f(1) = 3t+1$,$e_f(2) = 3t+1$

Also, $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1\ \forall\ i, j \in \{0,1,2\}$

Case (ii) $n \equiv 0(mod\ 3)$

Let $n = 3t$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(v_{2t+i}) = 2, 1 \leq i \leq t$

Here the necessary condition for cordiality is not satisfied.

Case (iii) : $n = 3t + 1$

Let $n \equiv 1 \ (mod\ 3)$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t + 1$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

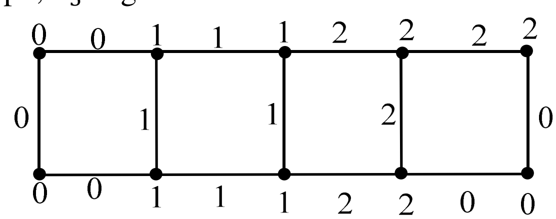
$f(v_{2t+1+i}) = 2, 1 \leq i \leq t$

Here, $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i, j \in \{0,1,2\}$ is not satisfied

Ladder graph L_n is a harmonic mean cordial labeled graph for $n \equiv 2 \ (mod\ 3)$

Example 3.2.

The harmonic mean cordial graph, L_5 is given below.



Here $v_f(0) = 3, v_f(1) = 4, v_f(2) = 3$

$e_f(0) = 5, e_f(1) = 4, e_f(2) = 4$

Also, $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i, j \in \{0,1,2\}$

Therefore L_5 is a harmonic mean cordial labeled graph.

Theorem 3.3

The Open Ladder graph $O(L_n)$ is a harmonic mean cordial labeled graph

Proof

Let $G = O(L_n)$ be a graph with vertex set $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and $E(G) = \{u_i u_{i+1}, v_i v_{i+1}: 1 \leq i \leq n - 1\} \cup \{u_i v_i: 1 \leq i \leq n\}$

Define $f : V(O(L_n)) \rightarrow \{0,1,2\}$ as follows.

Case(i):$n \equiv 0(mod\ 3)$

Let$n = 3t$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(v_{2t+i}) = 2, 1 \leq i \leq t$

Then $v_f(0) = 2t$,$v_f(1) = 2t$,$v_f(2) = 2t$

$e_f(0) = 3t - 1$,$e_f(1) = 3t - 2$,$e_f(2) = 3t - 1$

Therefore, $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1\ \forall\ i, j \in \{0,1,2\}$

Case (ii) :$n \equiv 1(mod\ 3)$

Let $n = 3t + 1$

Define$f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(u_n) = 0$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t$

Then $v_f(0) = 2t + 1$, $v_f(1) = 2t + 1$, $v_f(2) = 2t$

$e_f(0) = 3t$, $e_f(1) = 3t - 1$, $e_f(2) = 3t$

Therefore, $|v_f(i) - v_f(j)| \leq 1$, $|e_f(i) - e_f(j)| \leq 1\ \forall\ i, j \in \{0,1,2\}$

Case (iii) :$n \equiv 2\ (mod\ 3)$

Let $n = 3t + 2$

Define$f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(u_{2t+1+i}) = 2, 1 \leq i \leq t$

$f(u_n) = 0$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(v_{2t+i}) = 2, 1 \leq i \leq t+1$

Then $v_f(0) = 2t + 1$, $v_f(1) = 2t + 2$, $v_f(2) = 2t + 1$

$e_f(0) = 3t$, $e_f(1) = 3t + 1$, $e_f(2) = 3t + 1$

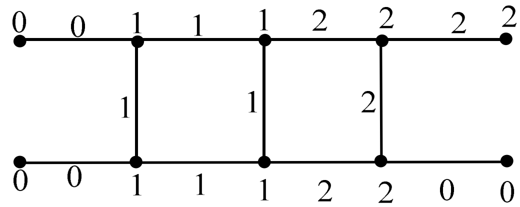
Therefore, $|v_f(i) - v_f(j)| \leq 1$

$|e_f(i) - e_f(j)| \leq 1 \ \forall \ i,j \in \{0,1,2\}$is satisfied

From all the above cases, the open ladder graph $O(L_n)$ is harmonic mean cordial graph.

Example 3.4

The harmonic mean cordial graph, $O(L_5)$ is given below.



Here $v_f(0) = 3, v_f(1) = 4, v_f(2) = 3$

$e_f(0) = 3, e_f(1) = 4, e_f(2) = 4$

Also, $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i,j \in \{0,1,2\}$

Therefore $O(L_5)$is a harmonic mean cordial labeled graph.

Theorem 3.5

The slanting ladder graph,SL_n is a harmonic mean cordial graph for $n \equiv 2 \ (mod\ 3)$

Proof

Let $G = SL_n$ be a slanting ladder graph with vertex set $V(G) = \{v_i, u_i: 1 \leq i \leq n\}$ and edge set $E(G) = \{(v_i u_{i+1}), (v_i v_{i+1}), (u_i u_{i+1}): 1 \leq i \leq n\}$

Define $f : V(SL_n) \to \{0,1,2\}$ as follows.

Case (i) :$n \equiv 0 \ (mod\ 3)$

Let $n = 3t$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(v_{2t+i}) = 2, 1 \leq i \leq t$

Here the necessary conditions for cordiality $|v_f(i) - v_f(j)| \leq 1$ and

$|e_f(i) - e_f(j)| \leq 1$ is not satisfied.

Case (ii) :$n \equiv 1 \ (mod\ 3)$

Let $n = 3t + 1$

Define$f(u_i) = 0, 1 \leq i \leq t + 1$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+1+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t+1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t$

Here the necessary conditions for cordiality $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1$ is not satisfied.

Case (iii) :$n \equiv 2\ (mod\ 3)$

Let $n = 3t + 2$

Define $f(u_i) = 0, 1 \leq i \leq t+1$

$f(u_{t+1+i}) = 1, 1 \leq i \leq t+1$

$f(u_{2t+2+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t+1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t+1$

Here $v_f(0) = 2t+1$, $v_f(1) = 2t+2$,$v_f(2) = 2t+1$,

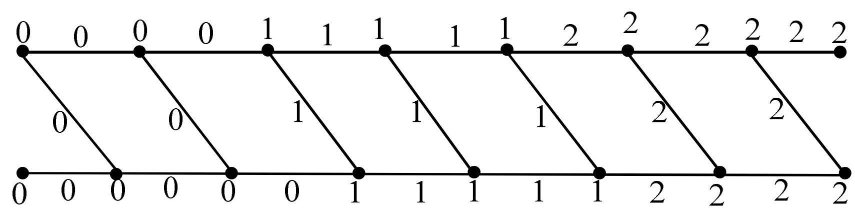
$e_f(0) = 3t+1$, $v_f(1) = 3t+1$,$v_f(2) = 3t+1$,

The condition for cordiality $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1$, $\forall\ i,j \in \{0,1,2\}$ are satisfied.

From all the above cases, the slanting ladder graph is a harmonic mean cordial graph for $n \equiv 2\ (mod\ 3)$and not for $n \equiv 0,1$ (mod 3)

Example 3.6

The harmonic mean cordial graph, SL_8 is given below.



Here $v_f(0) = 5$, $v_f(1) = 6$,$v_f(2) = 5$,

$e_f(0) = 7$, $v_f(1) = 7$,$v_f(2) = 7$,

Also $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1, \forall\ i,j \in \{0,1,2\}$

Therefore, SL_8 is a harmonic mean cordial graph.

Theorem 3.7

The triangular ladder graph, TL_n, $n \geq 2$ is a harmonic mean cordial graph for $n \equiv 2\ (mod\ 3)$.

Proof:

Let $G = TL_n$ be any triangular ladder graph with vertex set $V(G) = \{u_i, v_i : 1 \leq i \leq n\}$ and $E(G) = \{u_i u_{i+1}, v_i v_{i+1} 1 \leq i \leq n\}$

Define $f : V(TL_n) \rightarrow \{0,1,2\}$ as follows.

Case (i) : $n \equiv 0 \ (mod\ 3)$

Let $n = 3t$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(v_{2t+i}) = 2, 1 \leq i \leq t$

Here $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i, j \in \{0,1,2\}$ not satisfied.

Case (ii) : $n \equiv 1 \ (mod\ 3)$

Let $n = 3t + 1$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t + 1$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t$

Here $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i, j \in \{0,1,2\}$

Case (iii) : $n \equiv 2 \ (mod\ 3)$

Let $n = 3t + 2$

Define $f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t + 1$

$f(u_i) = 0, 1 \leq i \leq t + 1$

$f(u_{t+1+i}) = 1, 1 \leq i \leq t + 1$

$f(u_{2t+2+i}) = 2, 1 \leq i \leq t$

Then $v_f(0) = 2t + 1$, $v_f(1) = 2t + 2$, $v_f(2) = 2t + 1$,

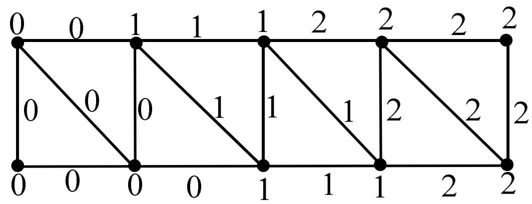
$e_f(0) = 4t + 2$, $v_f(1) = 4t + 1$, $v_f(2) = 4t + 2$

Also $|v_f(i) - v_f(j)| \leq 1$, $|e_f(i) - e_f(j)| \leq 1$ for all $i, j \in \{0,1,2\}$

From all the above cases, TL_n is a harmonic mean cordial graph

Example 3.8

The harmonic mean cordial graph, TL_5 is given below.



Here $v_f(0) \ = 3$, $v_f(1) = 4$,$v_f(2) = 3$

$e_f(0) \ = 6$, $e_f(1) = 5$,$e_f(2) = 6$

Also $|v_f(i) - v_f(j)| \leq 1$, $|e_f(i) - e_f(j)| \leq 1$ $\forall i, j \in \{0,1,2\}$

Theorem 3.9

The Open triangular ladder graph O(TL_n), $n \geq 2$ is a harmonic mean cordial graph for $n \equiv 1,2$ (mod 3) except $n \equiv 0$ (mod 3)

Proof:

Let $G \ = O(TL_n)$be a graph with vertex set $V(G) = \{u_i, v_i: 1 \leq i \leq n\}$ and

$E(G) = \{u_i u_{i+1}, v_i v_{i+1}: 1 \leq i \leq n\} \cup \{u_i v_i: 2 \leq i \leq n-1\} \cup \{v_i u_{i+1}: 1 \leq i \leq n-1\}$

Define $f : V\ (O(TL_n)) \ \rightarrow \{0,1,2\}$ as follows.

Case (i) :$n \ \equiv 0$ (mod 3)

Let $n \ = 3t$

Define $f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(v_{2t+i}) = 2, 1 \leq i \leq t$

Then $v_f(0) \ = 2t$, $v_f(1) = 2t$,$v_f(2) = 2t$,

$e_f(0) = 4t - 1$, $v_f(1) = 4t - 3$,$v_f(2) = 4t - 1$

Clearly $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1$ for all $\forall \ i, j \in \{$ 0,1,2$\}$

Case (ii) :$n \ \equiv 1$ (mod 3)

Let $n \ = 3t + 1$

Define$f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(u_{2t+1+i}) = 2, 1 \leq i \leq t$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t$

$f(u_{2t+i}) = 2, 1 \leq i \leq t$

$f(v_n) = 0,$

Then $v_f(0) \ = 2t + 1$, $v_f(1) = 2t + 1$,$v_f(2) = 2t,$

$e_f(0) = 4t$, $v_f(1) = 4t - 1$,$v_f(2) = 4t$

Clearly $|v_f(i) - v_f(j)| \leq 1$ and $|e_f(i) - e_f(j)| \leq 1 \ \forall \ i, j \in \{ 0,1,2\}$

Case (iii) :$n \ \equiv 2 \ (\ mod \ 3)$

Let $n \ = 3t + 2$

Define$f(u_i) = 0, 1 \leq i \leq t$

$f(u_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(u_{2t+1+i}) = 2, 1 \leq i \leq t$

$f(u_n) = 0$

$f(v_i) = 0, 1 \leq i \leq t$

$f(v_{t+i}) = 1, 1 \leq i \leq t + 1$

$f(v_{2t+1+i}) = 2, 1 \leq i \leq t + 1$

Then $v_f(0) \ = 2t + 1$, $v_f(1) = 2t + 2$,$v_f(2) = 2t + 1,$

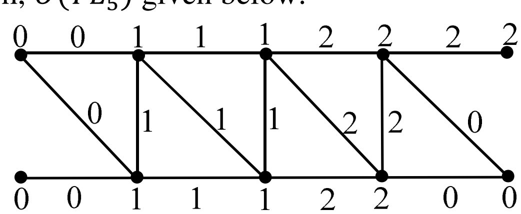
$e_f(0) = 4t + 1$, $v_f(1) = 4t + 1$,$v_f(2) = 4t + 1$

Also $|v_f(i) - v_f(j)| \leq 1$, $|e_f(i) - e_f(j)| \leq 1$ for all $i, j \in \{0,1,2\}$

Hence, $O(TL_n)$ is a harmonic mean cordial graph for $n \ \equiv 1{,}2 \ (\ mod \ 3)$ and not for$n \ \equiv 0 \ (\ mod \ 3)$

Example 3.10

The harmonic mean cordial graph, $O(TL_5)$ given below.



Here $v_f(0) \ = 3$, $v_f(1) = 4$,$v_f(2) = 3$

$e_f(0) \ = 5$, $e_f(1) = 5$,$e_f(2) = 5$

Also $|v_f(i) - v_f(j)| \leq 1$, $|e_f(i) - e_f(j)| \leq 1 \ \forall i, j \in \{0,1,2\}$

REFERENCES

1. *Bondy,J.A., Murthy,U.S.R., (1976). Graph Theory with Applications, Newyork, North Holland,*
2. *Cahit, I.,(1987). Cordial Graphs: A weaker Version of graceful and harmonious graphs ,Ars. Combinatoria, Vol 23, No. 3, pp. 201-207.*
3. *Gallian, J.A., (2011). A. Dynamic Survey of Graph Labeling, Vol.18*
4. *Harary, F., (1988). Graph Theory, Narosa Publishing House Reading, New Delhi.*
5. *Ponraj,R., Sivakumar,M.,(2012). Mean Cordial labeling of graphs, Open Journal of Discrete Mathematics, Vol.2, No.4,pp145.*
6. *Sandhya,S.S., Somasundaram, S., Ponraj,R.,(2012). Harmonic Mean Labeling of Some Cycle Related Graphs. Int. Journal of Math. Analysis, Vol.6, No. 40, pp 1997-2005.*
7. *Sandhya,S.S., Somadundaram,S., (2013). Harmonic Mean Labeling for Some Special Graphs, International Journal of Mathematics Research, Vol.5, No.1,pp.55-64.*
8. *DuraiBasker. A., (2020), Logarithmic Mean Labeling of Some Ladder Related Graphs, Applications and Applied Mathematics An International Journal(AAM), Vol. 15, Issue 1(Jan 2020), pp. 296 -313.*
9. *Sumathi. P., Rathi. A., (2018), Quotient Labeling of Some Ladder Graphs, American Journal of Engineering Research (AJER), Vol- 7, Issue- 12, pp. 38 - 42.*
10. *DuraiBasker. A., (2020), Logarithmic Mean Labeling of Some Ladder Related Graphs, Applications and Applied Mathematics An International Journal(AAM), Vol. 15, Issue 1(Jan 2020), pp. 296 -313.*
11. *Sumathi. P., Rathi. A., (2018), Quotient Labeling of Some Ladder Graphs, American Journal of Engineering Research (AJER), Vol- 7, Issue- 12, pp. 38 - 42.*

DOI No. : 10.5958/2394-9309.2022.00012.9
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

COEFFICIENT BOUNDS FOR A SUBCLASSOFM-FOLD SYMMETRIC BI-UNIVALENTFUNCTIONSCONCURRENT THE Q-ANALOGUE OF WANAS OPERATOR

S. Prathiba* and Thomas Rosy**
* & **Department of Mathematics, Madras Christian College, Tambaram, Chennai-600059,Tamil Nadu, India.
Email: prathimcc83@gmail.com*, thomas.rosy@gmail.com**

ABSTRACT

In this research article, we introduce subclasses $\mathfrak{S}_{\mathcal{E}_m}(\zeta,\mu,k,q;\delta)$and $\mathfrak{S}_{\mathcal{E}_m}(\zeta,\mu,k,q;\xi)$of holomorphic and m-fold symmetric bi-univalent functions involving the q-analogue of Wanas operator in the open unit disk and obtain estimates on the Taylor-Maclaurin coefficients $|a_{m+1}|$ and $|a_{2m+1}|$. Special cases of our results are indicated.

AMS Subject Classification: 30C45, 30C50.
Keywords: Analytic functions, m-fold symmetric bi-univalentfunctions, q-Wanas operator, coefficient estimates.

1. PRELIMNARIES

Let $\mathcal{A}$ indicate the class of all holomorphic functions of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n \tag{1}$$

in the open unit disk $\mathfrak{A}$= $\{z \in \mathbb{C}: |z| < 1\}$ normalized by the condition $f(0) = f'(0) =1$.Theclass of all functions in $\mathcal{A}$which are univalent in $\mathfrak{A}$are denoted by $\mathcal{S}$.

On account of Koebe one-quarter theorem [3],every univalent function $f \in \mathcal{S}$ has an inverse f^{-1}defined by

$$f^{-1}(f(z)) = z, \quad (z \in \mathfrak{A})$$

And $f(f^{-1}(\omega)) = \omega, \left(|\omega| < r_0(f);\ \ r_0(f) \geq \frac{1}{4}\right)$

where

$$g(\omega) = f^{-1}(\omega) = \omega - a_2\omega^2 + (2a_2^2 - a_3)\omega^3 - (5a_2^3 - 5a_2a_3 + a_4)\omega^4 + \cdots \tag{2}$$

A function $f \in \mathcal{A}$is said to be bi-univalent in $\mathfrak{A}$ iff$f$and $f^{-1}$are univalent in $\mathfrak{A}$. Denote by, $\mathcal{E}$ the class of all bi-univalent functions in $\mathfrak{A}$. Some illustrations of bi-univalent functions are$\frac{z}{1-z}, -\log(1-z),\ \ \frac{1}{2}\log\left(\frac{1+z}{1-z}\right), \ldots$

Lewin [8] inspected the class $\mathcal{E}$, and attained the bound $|a_2| \leq 1.51$. Brannan and Clunie[2] conjectured that $|a_2| \leq \sqrt{2}$. The coefficient estimate problem for $|a_n|$ $(n \in \mathbb{N}, n \geq 3)$ is still open [10].In recent years, the study of bi-univalent functions gained significance mainly due to the work of Srivastava et al [10]. Many researchers

(See[3,6]) recently explored several interesting subclasses of the class $\mathfrak{E}$ and found non-sharp estimates for the first two Taylor-Maclaurin coefficients.

For all$f \in \mathcal{S}$, the function $\mathfrak{h}(z) = \sqrt[m]{f(z^m)}$, $(z \in \mathfrak{U}, m \in \mathbb{N})$ is univalent and maps the unit disk $\mathfrak{U}$ into a region with m-fold symmetry. The normalized form ofm-fold symmetric bi-univalent functions (see [7,9]) is

$$f(z) = z + \sum_{k=1}^{\infty} a_{mk+1} z^{mk+1}, \quad z \in \mathfrak{U} \tag{3}$$

The class of m-fold symmetric univalent functions which are normalized by the above series expansion (3) is denoted by $\mathcal{S}_m$. The functions in $\mathcal{S}$ are one-fold symmetric.

Similar to the concept of m-fold symmetric univalent functions, we introduce the concept of m-fold symmetric bi-univalent functions. From (3), Srivastava et al [12]obtained the series expansion for $g = f^{-1}$ as

$$g(\omega) = f^{-1}(\omega) = (\omega - a_{m+1}\omega^{m+1} + [(m+1)a_{m+1}^2 - a_{2m+1}]\omega^{2m+1}$$
$$-\left[\frac{1}{2}(m+1)(3m+2)a_{m+1}^3 - (3m+2)a_{m+1}a_{2m+1} + a_{3m+1}\right]\omega^{3m+1} + \cdots \tag{4}$$

Denote $\mathfrak{E}_m$, the class of m-fold symmetric bi-univalent functions in$\mathfrak{U}$. For m=1, the formula (2) coincides with the formula (4) of the class $\mathfrak{E}$. Some illustrations of m-fold symmetric bi-univalent functions are $\frac{1}{2}\log\left(\frac{1+z^m}{1-z^m}\right)^{1/m}$, $(-\log(1-z^m))^{1/m}$, $\left(\frac{z^m}{1-z^m}\right)^{\frac{1}{m}}$ with the equivalent inverse functions $\left(\frac{e^{2\omega^m}-1}{e^{2\omega^m}+1}\right)^{\frac{1}{m}}$, $\left(\frac{\omega^m}{1+\omega^m}\right)^{\frac{1}{m}}$ and $\left(\frac{e^{\omega^m}-1}{e^{\omega^m}}\right)^{\frac{1}{m}}$ respectively. Recently, many researchers obtained bounds for various subclasses of m-fold symmetric bi-univalent functions (see [1,5,10,11]).

The aim of the present paper is to define subclasses $\mathfrak{S}_{\mathcal{E}_m}(\zeta, \mu, k, q; \delta)$ and $\mathfrak{S}_{\mathcal{E}_m}(\zeta, \mu, k, q; \xi)$ of $\mathfrak{E}_m$ involving the q-analogue of Wanas operator. Coefficient estimates for functions in these subclasses are obtained.

In 2019, Wanas [13]proposed an operator called Wanas operator. For$\mathcal{W}_{\nu,k}^{\varrho,\eta} : \mathcal{A} \to \mathcal{A}$defined as$\mathcal{W}_{\nu,k}^{\varrho,\eta} f(z) =$

$$z + \sum_{t=2}^{\infty} \chi_t(\eta, \varrho, \nu)^k a_t z^t, \tag{5}$$

where$\chi_t(\eta, \varrho, \nu) = \sum_{r=1}^{\eta} \binom{\eta}{r} (-1)^{r+1} \left(\frac{\varrho^r + t\nu^r}{\varrho^r + \nu^r}\right)$, (6)

$r, k \in \mathbb{N}_0, \nu \geq 0, \varrho \in \mathbb{R}$ and $\varrho + \nu > 0$.

For $f(z) \in \mathcal{A}$, we now define q-difference Wanas operator as

$$\mathcal{W}_{1,0,q}^{\varrho,1} f(z) = f(z)$$
$$\mathcal{W}_{1,1,q}^{\varrho,n} f(z) = z\mathcal{W}_q f(z)$$
$$\mathcal{W}_{1,k,q}^{\varrho,1} f(z) = z\mathcal{W}_q\left(\mathcal{W}_q^{k-1} f(z)\right)$$

$$\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z) = z + \sum_{t=2}^{\infty} [\chi_t(\eta, \varrho, \nu)]_q^k a_t z^t \tag{7}$$

where$\chi_t(\eta, \varrho, \nu)$ is given in (6), forr,$k \in \mathbb{N}_0, \nu \geq 0,\ \varrho \in \mathbb{R}, \varrho + \nu > 0,\ 0 < q < 1,$ $z \in \mathfrak{U}$.

Lemma 1. [3]If $\mathcal{U} \in \mathcal{P}$, then $|u_n| \leq 2$ for each $n \in \mathbb{N}$, where $\mathcal{P}$ is the class of all holomorphic functions $\mathcal{U}(z)$in $\mathfrak{U}$ for which $\Re e(\mathcal{U}(z)) > 0, (z \in \mathfrak{U})$, where $\mathcal{U}(z) = 1 + u_1 z + u_2 z^2 + \cdots, \ (z \in \mathfrak{A})$.

Definition 1. A function $f(z) \in \Sigma_m$ is said to be in the class $\mathfrak{S}_{\Sigma_m}(\zeta, \mu, k, q; \delta)$ if the following conditions are satisfied:

$$\left| arg\left(\frac{z(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z)\right) + \mu z} \right) \right| \le \frac{\pi\delta}{2} \tag{8}$$

and

$$\left| arg\left(\frac{\omega(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega)\right) + \mu \omega} \right) \right| \le \frac{\pi\delta}{2} \tag{9}$$

where $\mu \ge 1, 0 \le \zeta \le 1; 0 \le \delta \le 1; z, \omega \in \mathfrak{A}$ and $g = f^{-1}$ is given by (4). Also the q-Wanas operator $\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z)$ and $\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega)$ is given by

$$\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z) = z + \sum_{t=1}^{\infty} [\chi_{tm+1}(\eta, \varrho, \nu)]_q^k a_{tm+1} z^{tm+1} \tag{10}$$

and

$$\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega) = \omega + \sum_{t=1}^{\infty} [\chi_{tm+1}(\eta, \varrho, \nu)]_q^k b_{tm+1} \omega^{tm+1} \tag{11}$$

2. COEFFICIENT ESTIMATES OF THE FUNCTIONS $\mathfrak{S}_{\Sigma_m}(\zeta, \mu, k, q; \delta)$

Theorem1. A function $f(z) \in \mathfrak{S}_{\Sigma_m}(\zeta, \mu, k, q; \delta)$ $(\mu \ge 1, 0 \le \zeta \le 1; 0 \le \delta \le 1; z \in \mathfrak{A})$ be given by (3) then

$$|a_{m+1}| \le \frac{2\delta\sqrt{\zeta}}{\sqrt{m\left[2\delta[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1) - \left(2\delta - \frac{m(\delta-1)}{\zeta}\right)([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2\right]}} \tag{12}$$

and

$$|a_{2m+1}| \le \frac{\zeta\delta}{m[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k} + \frac{2(m+1)\zeta^2\delta^2}{m^2([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2}. \tag{13}$$

Proof: From the inequality (8) and (9),

$$\frac{z(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z)\right) + \mu z} = [s(z)]^\delta \tag{14}$$

and

$$\frac{\omega(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega)\right) + \mu \omega} = [t(\omega)]^\delta \tag{15}$$

where $g = f^{-1}$ and $s(z), t(\omega)$ *in* $\mathcal{P}$ have the following series representations:

$$s(z) = 1 + s_m z^m + s_{2m} z^{2m} + s_{3m} z^{3m} + \cdots \tag{16}$$

and

$$t(\omega) = 1 + t_m \omega^m + t_{2m} \omega^{2m} + t_{3m} \omega^{3m} + \cdots \tag{17}$$

Comparing the coefficients in (14) and (15) we get

$$\frac{m}{\zeta}[\chi_{m+1}(\eta, \varrho, \nu)]_q^k a_{m+1} = \delta s_m \quad (18)$$

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta, \varrho, \nu)]_q^k a_{2m+1} - \frac{m}{\zeta}([\chi_{m+1}(\eta, \varrho, \nu)]_q^k)^2 a_{m+1}^2 = \delta s_{2m} + \frac{\delta(\delta-1)}{2} s_m^2 \quad (19)$$

$$-\frac{m}{\zeta}[\chi_{m+1}(\eta, \varrho, \nu)]_q^k a_{m+1} = \delta t_m \tag{20}$$

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta, \varrho, \nu)]_q^k[(m+1)a_{m+1}^2 - a_{2m+1}] - \frac{m}{\zeta}\left([\chi_{m+1}(\eta, \varrho, \nu)]_q^k\right)^2 a_{m+1}^2$$

$$= \delta t_{2m} + \frac{\delta(\delta-1)}{2} t_m^2 \quad (21)$$

From (18) and (20) we obtain

$$s_m = -t_m \tag{22}$$

$$\frac{2m^2}{\zeta^2}\left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2 a_{m+1}^2 = \delta^2(s_m^2 + t_m^2) \tag{23}$$

Now from (19), (21) and (23) we get

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1)a_{m+1}^2 - \frac{2m}{\zeta}\left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2 a_{m+1}^2 = \delta(s_{2m} + t_{2m}) + \frac{\delta(\delta-1)}{2}(s_m^2 + t_m^2)$$

$$= \delta(s_{2m} + t_{2m}) + \frac{(\delta-1)m^2}{\delta\zeta^2}\left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2 a_{m+1}^2$$

Therefore, we have

$$a_{m+1}^2 = \frac{\zeta\delta^2(s_{2m}+t_{2m})}{m\left[2\delta[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1) - \left(2\delta - \frac{m(\delta-1)}{\zeta}\right)([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2\right]} \tag{24}$$

Using Lemma 1 for the coefficients s_{2m} and t_{2m},

$$|a_{m+1}| \leq \frac{2\delta\sqrt{\zeta}}{\sqrt{m\left[2\delta[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1) - \left(2\delta - \frac{m(\delta-1)}{\zeta}\right)([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2\right]}}.$$

Subtracting (21) from (19), and using (22), (23) we have

$$a_{2m+1} = \frac{\zeta\delta(s_{2m}-t_{2m})}{4m[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k} + \frac{(m+1)\zeta^2\delta^2(s_m^2+t_m^2)}{4m^2([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2} \tag{25}$$

Using Lemma 1 for the coefficients s_m, s_{2m}, t_m *and* t_{2m} in (25) we get,

$$|a_{2m+1}| \leq \frac{\zeta\delta}{m[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k} + \frac{2(m+1)\zeta^2\delta^2}{m^2([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2}.$$

For m =1 and $\eta = \nu = 1$ we have the one-fold symmetric bi-univalent functions.

Corollary 1 A function $f \in \mathfrak{S}_{\mathcal{E}_m}(\zeta,\mu,k,q;\delta)$ $(\mu \geq 1, 0 \leq \zeta \leq 1; 0 \leq \delta \leq 1; z \in \mathfrak{A})$ be given by (3) then

$$|a_2| \leq \frac{2\delta\sqrt{\zeta}}{\sqrt{4\delta\left[\frac{2\varrho+3}{\varrho+1}\right]_q^k - \left(2\delta - \frac{(\delta-1)}{\zeta}\right)([2]_q^k)^2}}$$

and

$$|a_3| \leq \frac{\zeta\delta}{\left[\frac{2\varrho+3}{\varrho+1}\right]_q^k} + \frac{4\zeta^2\delta^2}{([2]_q^k)^2}.$$

Definition 2. A function *f(z)* $\in \Sigma_m$ is said to be in the class $\mathfrak{S}_{\mathcal{E}_m}(\zeta,\mu,k,q;\xi)$ if the following conditions are satisfied:

$$\Re e\left\{\frac{z(\mathcal{H}_{\nu,k,q}^{\varrho,\eta}f(z))'}{\zeta\left(\mathcal{H}_{\nu,k,q}^{\varrho,\eta}f(z)\right)+\mu z}\right\} > \xi \tag{26}$$

$$\Re e\left\{\frac{\omega(\mathcal{H}_{\nu,k,q}^{\varrho,\eta}g(\omega))'}{\zeta\left(\mathcal{H}_{\nu,k,q}^{\varrho,\eta}g(\omega)\right)+\mu\omega}\right\} > \xi \tag{27}$$

where $\mu \geq 1, 0 \leq \zeta \leq 1, 0 \leq \xi \leq 1$ and $g = f^{-1}$ is given by (4).Also the q-Wanas operator $\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z)$ and $\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega)$is given by (10) and (11).

3. COEFFICIENT ESTIMATES OF THE FUNCTIONS $\mathfrak{S}_{\mathcal{E}_m}(\zeta, \mu, k, q; \xi)$

Theorem 2A function$f(z) \in \mathfrak{S}_{\mathcal{E}_m}(\zeta, \mu, k, q; \xi)$ $(\mu \geq 1, 0 \leq \zeta \leq 1; 0 \leq \xi \leq 1; z \in \mathfrak{U})$ be given by (3) then

$$|a_{m+1}| \leq \sqrt{\frac{2\zeta(1-\xi)}{m\left((m+1)[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k - ([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2\right)}} \tag{28}$$

and

$$|a_{2m+1}| \leq \frac{\zeta(1-\xi)}{m[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k} + \frac{2(m+1)\zeta^2(1-\xi)^2}{m^2([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2}. \tag{29}$$

Proof: It follows from (26) and (27) that there exist$\mathscr{s}(z), \mathscr{t}(\omega)$ in $\mathcal{P}$ such that

$$\frac{z(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} f(z)\right)+\mu z} = \xi + (1-\xi)\mathscr{s}(z) \tag{30}$$

$$\frac{\omega(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega))'}{\zeta\left(\mathcal{W}_{\nu,k,q}^{\varrho,\eta} g(\omega)\right)+\mu \omega} = \xi + (1-\xi)\mathscr{t}(\omega) \tag{31}$$

where $\mathscr{s}(z)$ and $\mathscr{t}(\omega)$ have the form (16) and (17), respectively.

Equating the coefficient of (30) & (31)

$$\frac{m}{\zeta}[\chi_{m+1}(\eta,\varrho,\nu)]_q^k a_{m+1} = (1-\xi)\mathscr{s}_m \tag{32}$$

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k a_{2m+1} - \frac{m}{\zeta}([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2 a_{m+1}^2 = (1-\xi)\mathscr{s}_{2m} \tag{33}$$

$$-\frac{m}{\zeta}[\chi_{m+1}(\eta,\varrho,\nu)]_q^k a_{m+1} = (1-\xi)\mathscr{t}_m \tag{34}$$

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k[(m+1)a_{m+1}^2 - a_{2m+1}] - \frac{m}{\zeta}\left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2 a_{m+1}^2$$

$$= (1-\xi)\mathscr{t}_{2m} \tag{35}$$

From (32) and (34), we get

$$\mathscr{s}_m = -\mathscr{t}_m \tag{36}$$

and

$$\frac{2m^2}{\zeta^2}\left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2 a_{m+1}^2 = (1-\xi)^2(\mathscr{s}_m^2 + \mathscr{t}_m^2) \tag{37}$$

Now from (33) and (35) we get

$$\frac{2m}{\zeta}\left[[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1) - \left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2\right] a_{m+1}^2 = (1-\xi)(\mathscr{s}_{2m} + \mathscr{t}_{2m})$$

$$a_{m+1}^2 = \frac{\zeta(1-\xi)(\mathscr{s}_{2m}+\mathscr{t}_{2m})}{2m\left[[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(m+1) - \left([\chi_{m+1}(\eta,\varrho,\nu)]_q^k\right)^2\right]} \tag{38}$$

Using Lemma 1 for the coefficients $\mathscr{s}_{2m}$ and $\mathscr{t}_{2m}$,we obtain

$$|a_{m+1}| \leq \sqrt{\frac{2\zeta(1-\xi)}{m\left((m+1)[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k - ([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2\right)}}.$$

Subtracting (35) from (33) we have,

$$\frac{2m}{\zeta}[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k(2a_{2m+1}-(m+1)a_{m+1}^2)=(1-\xi)(s_{2m}-t_{2m}) \quad (39)$$

Using Lemma 1 in (39),

$$|a_{2m+1}| \leq \frac{\zeta(1-\xi)}{m[\chi_{2m+1}(\eta,\varrho,\nu)]_q^k}+\frac{2(m+1)\zeta^2(1-\xi)^2}{m^2([\chi_{m+1}(\eta,\varrho,\nu)]_q^k)^2}$$

□

For m =1 and $\eta = \nu = 1$, we have the one-fold symmetric bi-univalent functions.

Corollary 2 A function $f(z) \in \mathfrak{S}_{\mathcal{E}_m}(\zeta,\mu,k,q;\xi)(\mu \geq 1, 0 \leq \zeta \leq 1; 0 \leq \xi \leq 1; z \in \mathfrak{A})$ be given by (3) then

$$|a_2| \leq \sqrt{\frac{2\zeta(1-\xi)}{2\left[\frac{2\varrho+3}{\varrho+1}\right]_q^k-([2]_q^k)^2}}$$

and

$$|a_3| \leq \frac{\zeta(1-\xi)}{\left[\frac{2\varrho+3}{\varrho+1}\right]_q^k}+\frac{4\zeta^2(1-\xi)^2}{([2]_q^k)^2}.$$

REFERENCES

1. *Altinkaya, S., Yalcin, S., (2015). Coefficient bounds for certain subclasses of m-fold symmetric bi-univalent functions, Journal of Mathematics, Art. ID 241683, 1-5.*
2. *Brannan, D.A., Clunie, J.G., (1980). Aspects of contemporary complex analysis, Academic Press Newyork.*
3. *Brannan, D.A., Taha, T.S., (1986). On some classes of bi-univalent functions, Studia Universitatis Babes-Bolyai, Mathematica, 31(2), 70-77.*
4. *Duren, P.L., (1983). Univalent Functions, Grundlehren der mathematischenWissenschaften 259, New York, Berlin, Hiedelberg, Tokyo, Springer-Verlag.*
5. *Eker, S.S., (2016). Coefficient bounds for subclasses of m-fold symmetric bi-univalent functions, Turk. J. Math., 40, 641-646.*
6. *Frasin, B.A., Aouf, M.K., (2004). New subclasses of bi-univalent functions, Appl. Math. Lett., 24, 1529-1573.*
7. *Koepf, W., (1989). Coefficients of symmetric functions of bounded boundary rotations, Proc. Amer. Math. Soc., 105, 324-329.*
8. *Lewin, M., (1967). On a coefficient problem for bi-univalent functions, Proc. Amer. Math. Soc, 18, 63-68.*
9. *Pommerenke, C., (1962). On the coefficients of close-to-convex functions, Michigan Math. J., 9, 259-269.*
10. *Srivastava, H.M., Mishra, A.K., Gochhayat, P., (2010). Certain subclasses of analytic and bi-univalent functions, Appl. Math. Lett., 23, 1188-1192.*
11. *Srivastava, H.M., Gaboury, S., and Ghanim, F., (2015). Coefficient estimates for some subclasses of m-fold symmetric bi-univalent functions, Acta Universitatis Apulensis, 41, 153-164.*
12. *Srivastava, H.M., Sivasubramanian, S., and Sivakumar, R., (2014). Initial coefficient bounds for a subclass of m-fold symmetric bi-univalent functions, Tbilisi. Math. J., 7(2), 1-10.*
13. *Wanas, A.K., (2019). New differential operator for holomorphic functions, Earthline J. Math. Sci., 527-537.*

DOI No. : 10.5958/2394-9309.2022.00013.0
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

AWARENESS ABOUT HEALTH INSURANCE POLICIES IN TAMILNADU- A STUDY WITH SPECIAL REFERENCE TO TAMBARAM TALUK

F. Sheela* Ph.D Research Scholar, **Dr. C.Venkatesh**** Assistant Professor
Department of Economics, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India.

ABSTRACT

In any pandemic situation made the entire world to realise the medical exigencies are unpredictable and can cause financial requisite which will be tough enough to handle the thought to have health insurance has emerged in the minds of ordinary peoples including socially and economically backward sections. Being insured under health insurance an individual person can protect the family on the whole.

Health insurance is very well established in many countries. But in India it is a new concept excepting for the prearranged sector employees. In India only about very least per cent of total health expenditure is funded by public/communal health insurance while 18 per cent is funded by government budget. This study focused upon the awareness about health insurance in Tambaram taluk. This study also provides suggestions for health insurance companies to deal with their limitations and to grab the opportunities more in the market

Keywords: Health insurance, Chi square test, Awareness, nature of employment, tambaram city, Sources of awareness etc.

INTRODUCTION

Health insurance is one of the considerable approaches that can help in boosting collective healthcare coverage through better healthcare utilization and financial fortification. Health insurance is also one of the important risk management systems to be used by the human beings to manage their health risks. It saves large amount of personal health expenditure. The involvement of health insurance in health financing is in increasing trend in emerging and developing economies.

How many accidents you need to comprehend that you need Health Cover? It takes just one visit to a sickbay to make us realize how defenseless we are, every passing every second. For the rich as well as poor, male as well as female and young as well as old, being diagnosed with an illness and having the need to be hospitalized can be a hard-hitting ordeal. Heart problems, diabetes, stroke, renal failure, cancer – the list of lifestyle diseases just seem to get longer and more common these days. Thankfully there are more speciality hospitals and specialist doctors – but all that comes at a cost. The super rich can afford such costs, but what about an average middle class person. For an illness that requires hospitalization/ surgery, costs can easily run into five digit bills. A Health insurance policy can cover such expenses to a hefty extent. Health Insurance is more significant these days compared to Old days Health is a human right, which has also been accepted in the constitution. Its accessibility and affordability has to be insured. While the well-to-do segment of the population both in rural & urban areas have suitability and affordability towards medical care, at the same time cannot be said about the people who belong to poor sector of the society. It is well known that more than 75% of the population utilizes private sectors for medical care unfortunately medical care becoming costlier day by day and it has become almost out of reach of the poor people. Today, there is need for injection of considerable resources in the health sectors to ensure affordability of medical care to all. Health insurance is an important option, which needs to be considered by the policy makers and planners. As mentioned earlier, the cost of Health Insurance depends on the sum assured, age, current health condition and your previous medical history. Higher the sum assured, higher the premium. Health insurance is important, one has to look at his/ her own lifestyle, health condition, age/ life stage, family history of illnesses and affordability. Keep in mind that most insurance companies limit the sum assured to a maximum of 5 lakhs. Also note that many health insurance policies ―provide additional benefits‖ such as daily allowance, ambulance charges, etc. for hospitalization. Not only are such ―benefits‖ superfluous, they tend to drive the premiums higher. So it is best to keep away from such plans and stick to something basic and simple.

Health insurance is a form of group insurance, where those pay premiums or taxes in order to help defend themselves from high or unexpected healthcare expenses. Health insurance works by estimating the overall "risk" of healthcare expenses and developing a routine finance structure (such as a monthly premium, or annual tax) that will ensure that money is available to pay for the healthcare benefits specified in the insurance agreement. The healthcare benefit is administered by a central organization, which is most often either a government agency, or a private or not-for-profit entity operating a health plan.

These objectives of this review are to identify various intrusions implemented in India to endorse awareness of health insurance and to provide evidence for the effectiveness of such interventions on the awareness and uptake of health insurance by the occupant Indian population.

STATEMENT OF THE PROBLEM

Health insurance in point of fact gives partial reimbursement to the people for expenditure on selected diseases. Although it is unneeded to say health insurance is an important apparatus in the modern world to save the individuals from the huge health shock, a very high percentage of people in Tambaram Taluk even from educated higher income groups are not covered under any health insurance policy and at the same time, although health care has become almost unaffordable for the poor people, it is surprising that the health insurance sector has not made much headway in India. In this study, the researcher has find out the various reasons for not aware for the health insurance so the interviewer analyzed for the detail studied about the awareness about health insurance policies in Tamilnadu-study with special reference to Tambaram Taluk.

OBJECTIVES OF THE STUDY

- To analyse the socio-economic condition of the health insurance policy holders in Tambaram Taluk.
- To find out the health insurance awareness level of the people and source of awareness of health insurance.

HYPOTHESIS OF THE STUDY

❖ There is no relationship between age and awareness level about the health insurance policies in Tambaram Taluk.

LIMITATIONS OF THE STUDY

In this study the researcher has selected only Tambaram Taluk. Due to time and cost constraints the researcher selects the tambaram not for whole of Tamilnadu. People only know about the life insurance not aware about the health insurance. Because most of the population in Tambaram Taluk outsiders of the city for come under the category of working.

REVIEW OF LITERATURE

The review of literature for health insurance in India is important as consumer behaviour changes with passage of time and in order to have knowledge about the various authors review findings and suggestions on the concerned topic. So, the review of literature for the study is as follows:

K. Selva Kumar and Dr. S. Vijay Kumar (2013) in their research article, titled on "Attitude of policy holders in the direction of administration of general insurance companies with orientation to Madurai region" The study reveals that 23% policy holders belongs to low level of attitude, 46% to medium level of attitude and 31% to high level of attitude. There is an important relationship between ages, sex, education, and marital status, type of family, community and level of their attitude headed for administration of services of public sector general insurance companies holds good.

R. Amsaveni and S. Gomathi (2013) in this study the researcher made an attempt to find out the satisfaction about medi claim policy holder, to analyse the reason for preferring medi claim policy to safe guard themselves and stay away from future risk, most of the respondents have taken personal scheme to employees. The major problems faced by the respondents are lack of timely communication and limited list of hospitals covered by the health insurance providers.

J. Jaypradha (2012) in this article titled on "Problems and prospects of health insurance in India" the researcher highlighted that the health insurance sector in India has registered 30% growth rate in 2008-09. The penetration of health insurance in India had risen to 4.8%, in 2008 from 1.2% in 1999-2000. The average medical expenditure of an Indian household is 6.7% of the annual income.

Ravikant Sharma (2011) in his paper titled, "A Comparison of Health Insurance Segment- India vs. China", the researcher seeks to compare both the economies India and China on health insurance aspect. Both economies have huge potential of health insurance and 45% of world's population lives in both the countries.

P. Jain et al., (2010) in his research paper titled "Problems faced by the Health Insurance Policyholders of Different Public and Private Health Insurance Companies for Settlements of their Claims" the researcher to find the problem faced by policyholders. The objectives were to study reason for refusal of claim, satisfaction level of customer and problems faced by them in getting their claim.

Ramesh Bhat and Falan Reuben (2001) in their article titled for, "Analysis of claim and reimbursements made under mediclaim policy of general insurance corporation of India" the researcher make an attempt to find out 621 claims and reimbursements data relating to policy beginning year 1997-98 and 1998-99 of Ahmadabad. They found that number of policies and premium collected have grown 30% during 1998- 00 and 50% during 1999-2000.

Research Methodology

This paper is based on exploratory research. The primary data was collected from people through a structured Interview schedule for bilingual language. The Secondary data was collected from different sources; Indian and international journals, health insurance bulletins, newspapers, books and websites etc.,

Research Type: - Exploratory

Sampling Technique: - Simple Random Sampling (Convenient sampling)

Sample Unit: - Respondents from Rajasthan Area (India)

Sample Size: - 230 (280 interview schedule were circulated but only 230 interview schedules were valid)

Tools for Data Collection: - Primary data (interview schedule)

Tools for Data Analysis: - Simple Percentage Test, Chi-square Test

DATA ANALYSIS AND DISCUSSION

In this paper the researcher analyses the socio-economic condition like age, income, type of family, status of the respondents, qualification etc., and the awareness levels of the health insurance and sources of health insurance.

Table 1: Socio-Economic Status of The Respondents

Sl. No.	Parameters	Variables	Frequency	Percentage
1.	Age	Below 25	24	10.43
		25-35	89	38.70
		35-45	77	33.47
		45-55	25	10.87
		Above 55	15	6.52
2.	Gender	Male	146	63.48
		Female	84	36.52
3.	Educational status	Illiterate	60	26.09
		Primary	32	13.91
		High school	35	15.22
		Intermediate	40	17.39
		Graduate and above	63	27.39
4.	Types of Family	Nuclear	210	91.30
		Joint	20	8.70
5.	Occupation type	Govt. Services	45	19.57
		Private Service	76	33.04
		Business	34	14.78
		House wives	36	15.65
		Students	39	16.96
6.	Locality	Urban	98	42.61
		Semi Urban	70	30.43
		Rural	62	26.96

Source: Collected Data

The above table 1 clearly exhibits that socio economic condition of the respondents, out of 230 respondents, Age of the respondents was recorded and it was observed that maximum respondents (38.70

per cent) belonged to the age group of 25 -35 years and minimum respondents (6.52 per cent) belonged to the age group of above 55 years. It was observed that 63.48 per cent of the respondents were male and only 37.52per cent were female in respondents. The data pertaining to education revealed that 26.09 per cent of the respondents were illiterate. Maximum number of respondents (27.39 per cent) was graduate or post graduate, 13.91 per cent respondents had qualification up to primary level only. The researcher has also revealed that 91.30 per cent respondents belonged to nuclear family and only 8.70 per cent belonged to joint family. It is clear from Table 1 that maximum respondent 33.04 per cent were belonged to private class, 19.57 per cent belonged to Government service, 18.33 percent belonged to government service, 11.67 per cent belonged to house wives and 16.96 per cent belonged to students. It is also clear from Table 1, that 42.61 per cent respondent lived in urban areas and only 26.96 per cent in rural area.

Table: 2 Awareness About Health Insurance

Sl No	Level of Awareness	No of respondents	Percentage
1	Yes	90	39.13
2	No	140	60.87
Total		230	100.00

Source: Collected data

The above analysis awareness about health insurance out of 230 respondents 60.87 per cent of the sample force were not aware for the health insurance and policies and the rest 39.13 per cent of the respondents only more awareness about the health insurance in Tambaram city.

Table: 3 Sources Awareness for Health Insurance

Sl No	Sources	No of respondents	Percentage
1	TV	22	24.44
2	Newspapers	8	8.89
3	Family members and Friends	12	13.33
4	Doctor	18	20.00
5	Insurance agent	15	16.67
6	Banks	5	5.56
7	Tax consultants	10	11.11
Total		**90**	**100**

Source: Collected Data

From the above table 3 clearly analyzed that source of awareness about health insurance, in this research only 90 respondents were more awareness and knowing about health insurance and insurance products. In these 24.44 per cent of the respondents knowing that Television; 20 per cent of the respondents know about doctors for checking for the health and pandemic situation the doctors are say to respondents what are the more benefits available in the health insurance; 16.67 per cent of the interviewee know that insurance agent followed by 13.33 per cent of the people know that family members and friends; only less no of respondents were know that banks because many of the private banks can also provide the insurance policies.

Table: 4 Reasons for Not Knowing Health Insurance

Sl No	Reasons	No of respondents	Percentage
1	No time to saw the newspaper and TV	26	18.57
2	Low level of education	15	10.71
3	Nature of employment	60	42.86
4	Inaccurate information	19	13.57
5	More health care cost	10	7.14
6	Unnecessary for me	10	7.14
Total		**140**	**100**

Source: Collected Data

The above table clearly preponderance that reasons for not knowing the health insurance; out of 140 respondents are not awareness for the health insurance, 42.86 per cent of the respondents say that nature of employment, 18.57 per cent of the respondents say that no time for saw the newspapers and Television etc., 13.57 sample force are say that inaccurate information in the various insurance companies and the remaining 7.14 per cent of the interviewees say that more health insurance premium and unnecessary for the family respectively.

VALIDATING OF HYPOTHESIS

To validate the research the researcher has analyse the null hypothesis that "There is no relationship between age and awareness of the health insurance". In this null hypothesis the researcher applied the chi square test for this research.

Calculated Chi Square Value	P value	Degrees of freedom	Significant level	Result
9.3536	9.488	5 % level	Significant	Accepted

Source: Computed Data

In this research 4 d.f at 5 % level of significance the table value is higher than the calculated value so the hypothesis is accepted. Hence concluded that there is no relationship between age of the respondents and awareness of health insurance.

Findings of the study

- Most (38.70%) of the respondents are fall under the age group of 25 to 35 years.
- Majority (63.48%) of the respondents are male remaining sample force are female.
- Very least (13.91 per cent) number of respondents are studied only primary level only.
- It is lucid that very least number of family are lived joint family because cultural changes etc.
- It is vivid that most of the respondents are working in private works.
- Majority 42.61 per cent of the interviewee were lived in urban area.
- 60.87 per cent of the not aware about health insurance policy and health insurance products.

- 24.44 per cent of sample force were knowing the health insurance through television (i.e. advertisement).
- 42.86 per cent of the respondents say that reason for not known the health insurance was nature of employment.

RECOMMENDATIONS AND CONCLUSION

Suggestions were asked from the respondents for health insurance companies to improve their current conditions. So many important points separately for government and private health insurance companies. Government companies were that they should improve their customer support services, introduce money back policies, develop short claim settlement process, develop more cashless facilities network and issue health insurance card like identity card to Indian citizens. Awareness campaigns and advertisements by the insurance companies should be held to enroll the rest of the families who are uninsured in spite of the good awareness. More enrolment and registration centres should be set up in each villages for easier accessibility and feasibility in enrolling for health insurance. The amount of premium should be customized to individual level by the government and private health insurance companies as majority of the families belonged to lower socio-economic classes. As lack of comprehensive coverage is one of the reason it should be brought to notice of insurance companies and dealt with it.

CONCLUSION

The researcher conclude that Health insurance is not a new concept and with the time more and more people are getting aware about it, but this awareness has not reach to that extent that people aware for it. Various socio- economic variables like types of family, education, income level, occupation etc drives people of Tambaram city to take the decision of taking health insurance.

REFERENCE

1. *Giedion U, Alfonso EA, Diaz Y. The impact of universal coverage schemes in the developing world: a review of the existing evidence. The World Bank, Washington (DC). [Online]. 2013 [cited 2017 Jan 15]; Available from: URL:http://hsrii.org/wpcontent/uploads/2014/07/IMPA CTofUHCSchemesin DevelopingCountries-A Review of Existing Evidence.pdf*
2. *Banerjee A, Criel B, Devdasan N, Guru S, Gangamurthy, Gautam RN et al. Training Manual on Health Insurance (with Facilitator's Guide). Institute of Public Health & Ministry of Health and Family Welfare. Government of India & WHO – Country Office for India 2011 May 27.*
3. *Indumathi K, Hajira SI, Gopi A, Subramanian M. Awareness of health insurance in a rural population of Bangalore, India. Int J Med Sci Public Health 2016;5(10):2162-7.*

4. *World Bank. Health expenditure indicators. [Online]. 2014 [cited 2017 Mar 10]; Available from:URL:http://data. worldbank.org/indicator/SH.XPD.PUBL.ZS?view=chart*
5. *Reddy KS, Selvaraj S, Rao KD, Chokshi M, Kumar P, Arora V et al. A critical assessment of the existing health models in India. Public Health Foundation of India,Planning Commission, Government of India [Online] 2011 [cited 2016 Sep 12]; Available from:URL:http://planning commission.nic.in/ reports/sereport/ser/ser_heal1305.pdf*
6. *India Fact Sheet. National Family Health Survey (NFHS)-4, Ministry of Health and Family Welfare Government of India. [Online]. 2015-2016 [Cited 2016 Aug 10]; Available from: URL: http://rchiips.org/NFHS/pdf/NFHS4/India. pdf*
7. *Sinha A,Jaiswal R, Pal BD, Shukla K,Sundar R, Natesh G. Pre-launch Survey Report of Insurance Awareness Campaign. National Council of Applied Economic Research, Insurance Regulatory Authority of India. [Online]. 2011 [Cited on 2017 Feb 16]; Available from: URL: file:/// C:/Users/User/Downloads/1387193211INSURANCE%20 AWARENESS.pdf*
8. *NSSO. Key Indicators of Social Consumption in India Health-NSS 71st Round. Ministry of Statistics and Programme Implementation, Government of India. [Online] 2015 [cited 2015 Sep 12]; Available from: URL: http://mospiold.nic.in/Mospi_New/site/inner.aspx?status=3&menu_id=31*
9. *Ramegowda C, Shivalingaiah AH. Awareness and Utilisation of Health Insurance Services by the Families Residing at Urban Field Practice Area of a Medical Col-lege in Bangalore. Natl J Community Med 2015;6(1):69-72.*
10. *Holyachi S, Gautham MS, Arvind BA, Arjunan I, Murthy NS, Pruthvish S et al. Health insurance: a study of awareness, coverage and factors influencing its subscription in a rural community. Indian Journal of Rural Health Care 2013 Jan-Mar;1(3):216-21.*
11. *Reshmi B, Nair NS, Sabu KM, Unnikrishan B. Awareness of health insurance in a south Indian population: A community based study. Health and Population Perspectives and issues 2008;30:177-88.*
12. *Nair D. Determinants of Enrollment in Comprehensive Health Insurance Scheme and Implementation Challenges: A Study in Kerala, South India. Health Sci J 2015;10:1.*
13. *Bawa SK, Ruchita. Awareness and willingness to pay for health insurance: an empirical study with reference to Punjab India. International Journal of Humanities and Social science 2011 June;1(7):100-8.*
14. *Madhukumar S, Sudeepa D, Gaikwad V. Awareness and perception regarding health insurence in banglore rural population. International Journal of Medicine and Public health 2012;2(2):18-22.*

DOI No. : 10.5958/2394-9309.2022.00014.2
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

FORECASTING TIME SERIES ANALYSIS FOR PADDY PRODUCTION

S. Stellamani [*], Dr. M. Vijaya Kumar [**], Dr. A. Philip Arokiadoss [***]
[*] Research Scholar , Department of Statistics Annamalai University ,Annamalai Nagar, chidambaram, Tamilnadu,India.
[**] Associate Professor , Department of Statistics Annamalai University ,Annamalai Nagar ,chidambaram,Tamilnadu,India.
[***] Assistant Professor , Department of Statistics St.Joseph's College,Trichy,Tamilnadu,India.

ABSTRACT

This paper investigates predictive performance of time series analysis method for paddy production like Autoregressive (AR) , Moving Average (MA) and Autoregressive Integrated Moving Average (ARIMA) process to select the appropriate model and predict in the future . In the present study had been carried out on the basis of time series production data of paddy crop pertaining to the period 1960 to 2020 has been collected and forecasting the production for 2021 to 2030 using . Based on ARIMA (p,d,q) and its components Autocorrelation Function (ACF) , Partial Autocorrelation Function (PACF),Bayasian Information Criterion (BIC) AND Box-Ljung Q statistics and residuals estimated (0,1,1) was selected.

Keywords: Health insurance, Chi square test, Awareness, nature of employment, tambaram city, Sources of awareness etc.

1. INTRODUCTION

Rice is one of the chief grains of India. Moreover, this country has the largest area under rice cultivation, as it is one of the principal food crops. It is, in fact, the dominant crop of the country. India is one of the leading producers of this crop. Rice is the basic food crop and being a tropical plant, it flourishes comfortably in hot and humid climate. Rice is mainly grown in rain-fed areas that receive heavy annual rainfall. That is why it is fundamentally a kharif crop in India. It demands temperature of around 25 degree Celsius and above, and rainfall of more than 100 cm. Rice is also grown through irrigation in those areas that receive comparatively less rainfall. Rice is the staple food of eastern and southern parts of India.

Rice can be cultivated by different methods based on the type of region. But in India, traditional methods are still in use for harvesting rice. The fields are initially ploughed and fertilizer is applied which typically consists of cow dung, and then the field is smoothed. The seeds are transplanted by hand and then through proper irrigation, the seeds are cultivated. Rice grows on a variety of soils like silts, loams and gravels. It can tolerate alkaline as well as acid soils . However, clayey loam is well suited to the raising of this crop. Actually the clayey soil can be easily converted into mud in which rice seedlings can be transplanted easily. Proper care has to be taken as this crop thrives if the soil remains wet and is under water during its growing years. Rice fields should be level and should have low mud walls for retaining water. In the plain areas, excess rainwater is allowed to inundate the rice fields and flow slowly. Rice raised in the well-watered lowland areas is known as lowland or wet rice. In the hilly areas, slopes are cut into terraces for the cultivation of rice. Thus, the rice grown in the hilly areas is known as dry or upland rice . The yield of upland rice per hectare is comparatively less than that of the wet rice.

The regions cultivating this crop in India are distinguished as the western coastal strip, the eastern coastal strip, covering all the primary deltas, Assam plains and surrounding low hills, foothills and Terai region- along the Himalayas and states like West Bengal ,Bihar , eastern Uttar Pradesh , eastern Madhya Pradesh , northern Andhra Pradesh and Odisha . India, being a land of eternal growing season, and the deltas of the Ganges-Brahmaputra (in West Bengal), Kaveri River, Krishna River , Godavari River , Indravathi River and Mahanadi River with a thick set-up of canal irrigation like Hirakud Dam and Indrāvati Dam , permits farmers to raise two, and in some pockets, even three crops a year. Irrigation has made even three crops a year possible. Irrigation has made it feasible even for Punjab and Haryana, known for their baked climate, to grow rice. They even export their excess to other states. Punjab and Haryana grow prized rice for export purposes. The hilly terraced fields from Kashmir to Assam are ideally suited for rice farming, with age-old hill irrigational conveniences. High yielding kinds, enhanced planting methods, promised irrigation water supply and mounting use of fertilizers have together led to beneficial and quick results. It is the rain fed-area that cuts down average yields per hectare.

In some states like West Bengal, Assam and Orissa two crops of rice are raised in a year. Winter season in northwestern India is extremely cold for rice. Rice is considered as the master crop of coastal India and in some regions of eastern India, where during the summer and monsoon seasons, both high temperature and heavy rainfall provide ideal conditions for the cultivation of rice. Almost all parts of India are suitable for raising rice during the summer season provided that the water is available. Thus, rice is also raised even in those parts of western Uttar Pradesh, Punjab and Haryana where low level areas are waterlogged during the summer monsoon rainy season.

Winter rice crop is a long duration crop and summer rice crop is a short duration crop. At some places in the eastern and southern parts of India, rice crop of short duration is followed by the rice crop of long duration. Winter rice crop is raised preferably in low-lying areas that remain flooded mainly during the rainy season. Autumn rice is raised in Uttar Pradesh, Maharashtra, Rajasthan, Madhya Pradesh, Punjab and Himachal Pradesh. Summer, autumn and winter rice crops are raised in West Bengal, Andhra Pradesh, Assam and Orissa. Summer rice crop is raised on a small scale and on a small area. However, winter rice crop is actually the leading rice crop accounting for a major portion of the total Hectare under rice in all seasons in the country. Moreover, in the last few years, several steps in order to augment yield per hectare were taken up very seriously at all levels. India ranks fourth in the production of wheat & second in production of rice in the world. Favorable Geographical Condition for Wheat Cultivation: In India, wheat is a winter crop. Wheat requires a moderately cool climate with moderate rain. In India, it is grown in winter. It needs temperature 10 degree C to 15 degree C for its cultivation. It thrives well in an average temperature of 16-degree C. Warm and sunny weather is essential at the time of ripening.

2. ARIMA Model:

A time series is a set of numbers that measures the status of some activity over time. It is the historical record of some activity, with measurements taken at equally spaced intervals with a consistency in the activity and the method of measurement.

2.1 Moving Average Process:

Moving average models were first considered by Slutsky (1927) and Wold (1938). The Moving Average Series can be written as

$Y_t = e_t - \theta_1 e_{t-1}\ e_t - \theta_2 e_{t-2}\ e_t - \ldots - e_t - \theta_q e_{t-q}$

We call such a series a moving average of order q and abbreviate the name to MA (q). Where, is the original series and is the series of errors

2.2 Auto-Regressive Process:

Yule (1926) carried out the original work on autoregressive processes. Autoregressive processes are as their name suggests regressions on themselves. Specifically, a p^{th}- order autoregressive process $\{Y_t\}$ satisfies the equation

$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \ldots + \theta_p Y_{t-p} + e_i$

The current value of the series Y_t is a linear combination of the most recent past values of itself plus an "innovation" term e_i that incorporates everything new in the series at time that is not explained by the past values. Thus, for every , we assume that is independent of

$Y_{t-1}, Y_{t-2}, Y_{t-3} \ldots Y_{t-q}$

2.3 Autoregressive Integrated Moving Average (ARMIA) model

The Box and Jenkins (1970) procedure is the milestone of the modern approach to time series analysis. Given an observed time series, the aim of the Box and Jenkins procedure is to build an ARIMA model. In particular, passing by opportune preliminary transformations of the data, the procedure focuses on Stationary processes. In this study, it is tried to fit the Box-Jenkins

Autoregressive Integrated Moving Average (ARIMA) model. This model is the generalized model of the non-stationary ARMA model denoted by ARMA (p, q) can be written as

$\theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \ldots + \theta_p Y_{t-p} + e_t - \theta_1 e_{t-1}\ e_t - \theta_2 e_{t-2}\ e_t - \ldots - e_t - \theta_q e_{t-q}$

Where, Y_t is the original series, for every t , we assume that e_t is independent of $Y_{t-1} + Y_{t-2} + Y_{t-3} + \ldots + Y_{t-q}$. A time series $\{Y_t\}$is said to follow an integrated autoregressive moving average (ARIMA) model if the d^{th} difference$W_t = \nabla^d Y_t$ is a stationary ARMA process. If$\{W_t\}$ follows an ARMA (p, q) model, we say that is an ARIMA (p, d, q) process. Fortunately, for practical purposes, we can usually take d = 1 or at most 2.

Consider then an ARIMA (p, 1, q) process. With $W_t = Y_t - Y_{t-1}$ we have

$W_t = \theta_1 W_{t-1} + \theta_1 W_{t-2} + \ldots + \theta_P W_{t-P} + e_t - \theta_1 e_{t-1}\ e_t - \theta_2 e_{t-2}\ e_t - \ldots - e_t - \theta_q e_{t-q}$

2.4 Box and Jenkins procedures

i.Preliminary analysis: create conditions such that the data at hand can be considered as the realization of a stationary stochastic process.

ii. Identification: specify the orders p, d, q of the ARIMA model so that it is clear the number of parameters to estimate. Recognizing the behavior of empirical autocorrelation functions plays an extremely important role.

iii. Estimate: efficient, consistent, sufficient estimate of the parameters of the ARIMA model (maximum likelihood estimator).

iv. Diagnostics: check if the model is a good one using tests on the parameters and residuals of the model. Note that also when the model is rejected, still this is a very useful step to obtain information to improve the model.

v. Usage of the model: if the model passes the diagnostics step, then it can be used to interpret a phenomenon, forecast.

2.6 Ljung-Box test

Ljung-Box Test can be used to check autocorrelation among the residuals. If a model fit well, the residuals should not be correlated and the correlation should be small. In this case the null hypothesis is

H_0: $\rho_1(e)= \rho_2(e)=\ldots= \rho_k(e)=0$ is tested with the Box-Ljung statistic

$Q^{*}=N(N+1)\sum_{i=1}^{n}(N-K)\ \rho_k^{2}(e)$

Where, N is the no of observation used to estimate the model. This statistic Q* approximately follows the chi-square distribution with (k-q)dfType equation here., where q is the no of parameter should be estimated in the model. If Q* is large(Significantly large from zero), it is said that the residuals autocorrelation are as a set are significantly different from zero and random shocks of estimated model are probably auto-correlated . So one should then consider reformulating the model.

3. RESULT AND DISCUSSION

Figure 1 reveals that the data used were non-stationary .Again non- stationarity in mean was corrected through first differencing of the data. The newly constructed variable Y_t could now be examined for stationary . Since , Y_t as stationary in mean , the next step was to identify the values of p and q . for this, the ACF and PACF of various orders of Y_t were computed and presented in table 1 and figure 2 .

The tentative ARIMA models are discussed with values differenced once (d-1) and the model which had the minimum normalized BIC was chosen. The various ARIMA models and the Corresponding normalized BIC values are given in Table 2. The value of normalized BIC of the chosen ARIMA was 17.232 .

4. MODEL ESTIMATION

Model parameters and fit statistics were estimated and the result of estimation are presented in table 3 and 4 . R^2 value was 0.963 . Hence the most suitable model for paddy production was ARIMA (0,1,1) as this model had the lowest Normalized BIC value , good R^2 and better model fit statistics (RMSE and MAPE) .

5. DIAGNOSTIC CHECKING

For model verification, various autocorrelation up to 16 lags were computed and the same along with their significance tested by Box – Ljung statistic are provided(in table 5). As the results indicate , none of these autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARIMA model is an appropriate model for forecasting paddy production in India.

The ACF and PACF of the residuals are given in figure 3 ,which also indicated the 'good fit' of the model Hence, the fitted ARIMA model for the paddy production data was

$$Y_t = \mu+\theta_1 Y_{t-1}+ \theta_2 Y_{t-2}+\ldots+ \theta_p Y_{t-p}+€_t$$

$$Y_t=1444.779+0.721€_{t-1}+0.081€_{t-1}+€_t$$

Table1:1ACF and PACF of paddy production

Lag	ACF value	Std error Df	Box-Ljung statistics Sig value	PACF value	Std error
1	0.925	0.125	54.855	0.925	0.128
2	0.882	0.124	105.494	0.176	0.128
3	0.830	0.123	151.135	-0.038	0.128
4	0.784	0.122	192.533	-0.003	0.128
5	0.750	0.121	231.101	0.076	0.128
6	0.700	0.120	265.326	-0.099	0.128
7	0.653	0.119	295.711	-0.047	0.128
8	0.604	0.117	322.155	-0.037	0.128
9	0.549	0.116	344.392	-0.079	0.128
10	0.504	0.115	363.553	0.013	0.128
11	0.462	0.114	379.957	0.020	0.128
12	0.428	0.113	394.348	0.037	0.128
13	0.384	0.112	406.152	-0.081	0.128
14	0.334	0.111	415.282	-0.081	0.128
15	0.287	0.109	422.186	-0.027	0.128
16	0.247	0..108	427.397	0.017	0.128

Table 2: BIC value of various ARIMA (p,d,q)

ARIMA(p,d,q)	BIC value
010	17.553
011	17.232
012	17.458
110	17.528
111	17.444
112	17.528
210	17.603
211	17.518
212	17.575
310	17.658
311	17.594
312	17.582

Table 3:Estimated ARIMA model

	Estimate	SE	T	Sig
Constant	1444.779	194.463	7.430	0.000
ARI	0.721	0.098	7.356	0.000

Table 4:

ARIMA (p,d,q)	Stationary R squared	R Squared	RMSE	MAPE	Max APE	MAP	Max AE	Normalized BIC
010	0.224	0.963	5155.254	6.798	30.913	3793.823	19785.596	17.553
011	0.370	0.963	5155.231	6.328	31.517	3793.602	19785.108	17.232
012	0.385	0.963	5155.289	6.386	31.623	3793.598	19785.368	17.458
110	0.398	0.963	5155.291	6.333	31.738	3793.567	19785.429	17.528
111	0.399	0.963	5155.309	6.372	31.746	3793.437	19785.486	17.444
112	0.399	0.963	5155.316	6.366	31.769	3793.482	19785.549	17.528
210	0.421	0.963	5155.389	6.364	31.822	3793.442	19786.265	17.603
211	0.424	0.963	5155.372	6.374	31.843	3793.438	19786.326	17.518
212	0.404	0.963	5155.392	6.377	31.876	3793.439	19786.343	17.575
310	0.423	0.963	5155.406	6.343	32.100	3793.562	19786.382	17.658
311	0.445	0.963	5155.496	6.398	32.256	3793.578	19786.392	17.594
312	0.452	0.963	5155.498	6.336	32.256	3793.581	19786.396	17.582

Figure 1: Time plot are paddy production

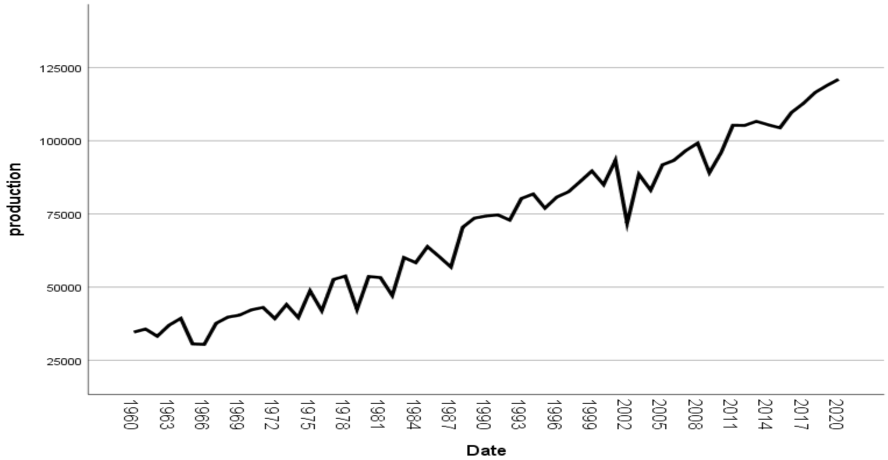
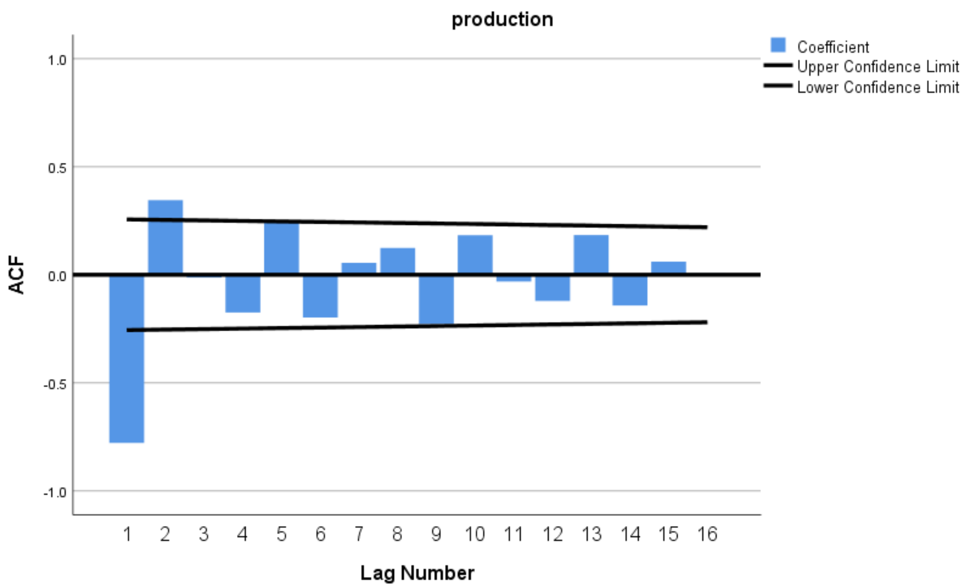


Figure 2:ACF and PACF of differenced data ACF



PACF

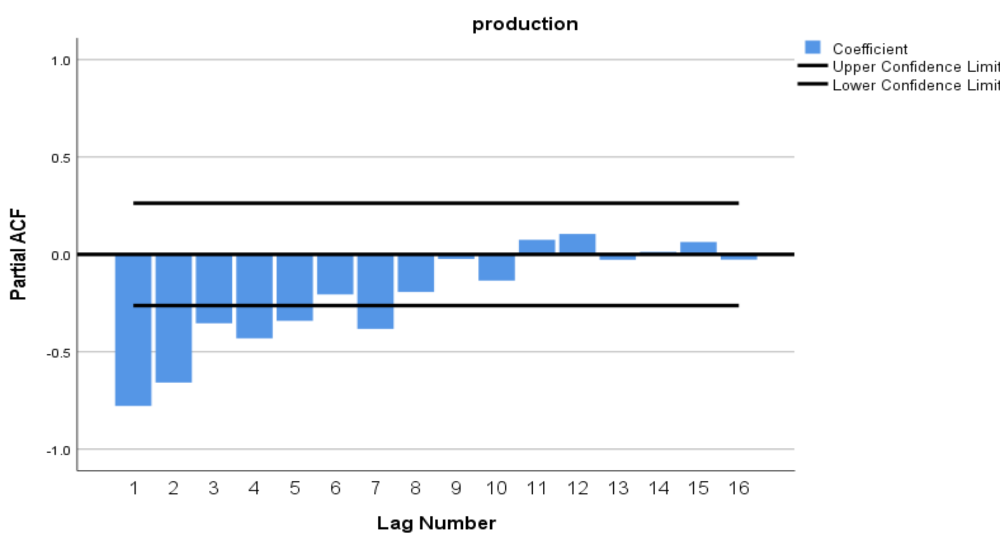


Figure 3: Residuals of ACF and PACF

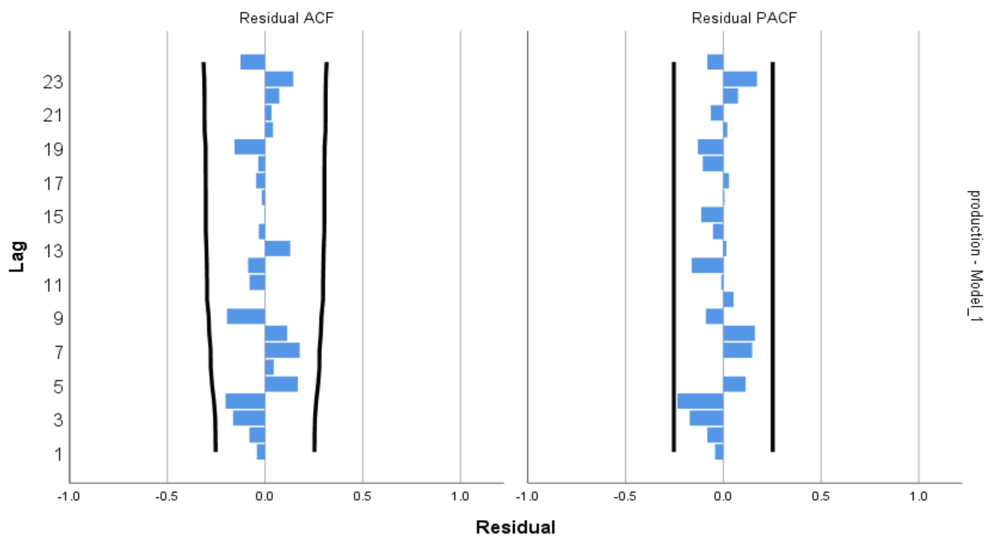
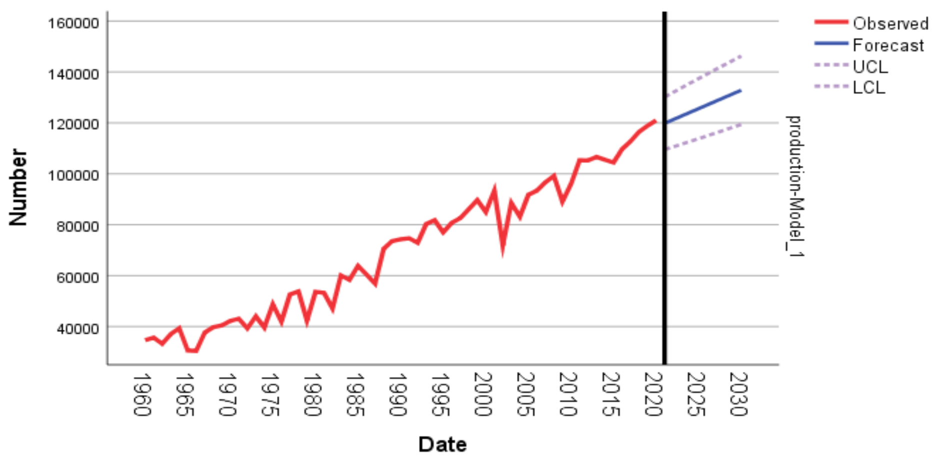


Figure 4: Actual and Estimated of paddy production



Conclusion

The most appropriate ARIMA model for paddy production forecasting of data was found to be ARIMA (0,1,1) from the temporal data , it can be found that forecasted production would increase to 135076 million tons in 2030 from 121000 million tons in 2020 in India . That is using time series data from 1960 to 2030 on paddy production, this study provides an evidence on future paddy production in the country ,which can be considered for future policy making and formulation strategies for augmenting and sustaining paddy production in India.

REFERENCE

1. *Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control. San Francisco, Holden- Day, California, USA.*
2. *Box, G.E.P., G.M. Jenkins (1976). Time Series Analysis, Forecasting and Control. San Francisco, Holden–Day, California, USA.*

3. *Brown, R.G and R.F.Meyer and D.A.D'Esopo (1961).The fundamental theorem of exponential smoothing, Operations Research, Vol.9 (5), pp.673-685.*
4. *Gardner, E.S. (1985).Exponential smoothing-The state of the art, J.Forecasting, vol.pp.1-28.*
5. *Holt C.C. (2004).Forecasting seasonal and trends by exponentially weighted moving averages, International Journal of Forecasting, vol.20 (1), pp.5-10.*
6. *Makridakis, S .,S.C.Wheelwright and R.J.Hyundman (1998). Forecasting: methods and applications, J.Wiley and Sons: New York.*
7. *Makridakis,S and M.Hibbon (1979).Accuracy of forecasting an empirical investigation, J.Roy. Statist. Soc.A,vol.41(2),pp.97-145.*
8. *Pankratz ,A (1983).Forecasting with univariate box Jenkins models concept and cases,J.Wiley and sons, New York.*
9. *Wold , H.O.A (1983).A study of the analysis of stationary time series, Uppasala:Almqvist and Wiksells.*
10. *M.Saranyadevi,Dr .A.Kachi Mohideen(2019).forecast of paddy production in tamilnadu using stochastic models,the international journal of analyticaland experimental model analysis vol. XI, IssueIX,September/2019*
11. *Yule ,G.U (1926). Why do we sometimes get nonsense- correlations between time series, Journal of Royal Statistical Society,Vol.89(1),pp.1-63.*

DOI No. : 10.5958/2394-9309.2022.00015.4
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

CUSTOMER'S ATTITUDE TOWARDS ONLINE BANKING SERVICES FOR PRE AND POST COVID - 19

Dr. N. Ramu* , Mr. P.Kumaravel ,Dr A. Mahendra Kumar*****
*Professor, Department of Commerce, Annamalai University, Annamalai Nagar, Chidambaram, India.
**Ph.D. Research scholar (Part-time), Department of Commerce, Annamalai University, Annamalai Nagar, Chidambaram, India.
***Assistant Professor, Department of Commerce, St. Thomas College of Arts and Science, Koyambedu, Chennai – 107.

ABSTRACT

The pandemic COVID-19 has severely affected the global economy. The strict lockdown measures have also changed the daily live, including consumer behavior in retail banking. In this context, the purpose of this paper is to investigate the customers attitude towards online banking service for pre and post COVID-19 banking, with a special focus on the nationalized banking sector only taken for 4 bank nationalized banks like Indian Bank, Indian Overseas bank, Canara Bank, and Karur Vysya Bank. To achieve our goal, we performed a survey among the Tambaram consumers in banking, using as research method the based-on interview schedule method. The final sample comprised 360 valid responses from the Tambaram city banking customers. The research brings a fresh insight on nationalised banking services consumption during the pandemic and validates a conceptual model regarding the online banking services acceptance.

Keywords: Customers' attitude, Pandemic situations, Covid 19, Online banking, testing of hypothesis in Chi square test etc

1. INTRODUCTION

The pandemic has impacted virtually all aspects of our lives. Some developments have been sudden and involuntary, such as social distancing, wearing masks, stopping public transport, restrictions on travel, etc. For others, it has merely accelerated the adoption of behavior's already gaining traction, such as the digitalization of shopping, banking and more.

All customers attitude have changing thecustomer's behaviour and it has strong location and time dependencies. Behaviour can differ significantly from one location to another depending on cultures, geographies, etc. The pandemic is making this dimension of consumer behaviour more complex; for example, since physical movement is restricted, consumers are migrating into virtual worlds at an unprecedented rate and are exposed to newer influences. This could require us to go beyond traditional methods of modeling their behavior.

Banks can play an immediate role in slowing the spread of COVID-19 by helping customers make better use of existing digital and remote channels. And banks can help limit the impact of the likely downturn by building new experiences to help their customers manage debt, adjust budgets, and make full use of new government programs.

The last time there was a global crisis, banks were widely perceived to be a big part of the problem. This time around, banks are central to the solution. Banks can play an immediate role in slowing the spread of COVID-19 by helping customers make better use of existing digital and remote channels. And banks can help limit the impact of the likely downturn by building new experiences to help their customers manage debt, adjust budgets, and make full use of new government programs. In normal times, customer experience in banking is about making customers happy—with the result that they are more loyal, use products more, and cost less to serve. In the context of COVID-19, superior customer experience means clarity and transparency.

The COVID-19 has proven to be a truly global pandemic, impacting people in just about every corner of the world. The COVID – 19 pandemic and the resulting lockdowns have caused significant disruption for people, communities and business changes in the banking sector. The banking sector plays a vital role development of one country's economy. The growth of banking sector depends upon the attitude of the banking customers in various aspects. Today the banking industry has been experiencing a totally unexpected paradigm shift, and in this age of advanced technology and modern resource of the banks. COVID-19 has radically impacted consumer behavior in world-wide. Banks are questioning if these changes will last once the global lockdowns end and the pandemic squanders. Before the pandemic situation Covid -19, the consumer expectations were changing, and through digital and technology transformation programs, banks were shifting the way they delivered products and services to consumers.Customer service has great significance in the banking industry. The banking system in India today has perhaps the largest outreach for delivery of financial services and is also serving as an important channel for delivery of financial services also.

2. STATEMENT OF THE PROBLEM

Customer service is the most important duty of the banking operations. Prompt and efficient service with smile will build up good public relations trim down complaints and increase business.In this research the customers are foremost and expect the various services from the banks. In the pandemic situation the customer's attitude has changed. The most important attitudes were changed for the pandemic situation. Consumers would like from a bank are high-quality customer service, low fees, security and fraud protection, and mobile and online access. If consumers could only have one feature in a bank, it'd be low fees. Competitive interest rates are important to over 90% of customers. In this research the customers find the customers attitude towards pre and post pandemic situation (COVID '19).

3. OBJECTIVES OF THE STUDY

- To find the services offered by the banks before and after Covid 19.
- To explore the problems faced by the banks before and after Covid 19
- To analyse the collection of loans and advances from the customers pandemic situations.

4. HYPOTHESIS OF THE STUDY

- There is no association between location and customers attitude in the pandemic situations.

5. LIMITATIONS OF THE STUDY

The researcher has faced many problems while collecting data from the banking customers. The customers not give the correct information because fear of fraudulent activities in the banking sector. Time has minimum only (i.e. Two months). The researcher has collected the data from Tambaram City and selected only four nationalized banks i.e. Indian Bank, Indian Overseas bank, Karur Vysya Bank and Canara Bank.

6. LITERATURE REVIEW

Banks started to invest in information technology to improve their product and services offered to the customer and also increase their productivity and efficiency of the business. The research shows that using a new kind of technology is affected by different theoretical model such as psychology and sociology.

Sinha and Mukherjee (2016) studied and explored regarding why off branch emanaging
an account in India is not very acknowledged when contrasted with cuttingedge nations. They incorporated develops from the innovation acknowledgmentdemonstrate (TAM), dispersions of advancement (DOI) model and trust hypothesismodels. Data gathered through review was investigated utilizing different relapsestrategy. They found that trust on innovation, trust on bank, saw convenience, sawhandiness, many-sided quality fundamentally influences customer to use off branch esavingmoney in India aside from saw hazard.

Szopinisi (2016) identified the factors applying an impact on the utilization of webbased keeping money in Poland. Observational material for this study was gottenfrom "Social Diagnosis" researches extend, worked by Board of Social Monitoringworking at the University of Finance and Management in Warsawin by 2015.Aftereffects of direct relapse investigation mirrored that the utilization of the Internet,exploiting other saving money items and trust in business banks decide the use of webbased managing an account. It was watched that home loans and charge cards twomanaging an account items have the greatest impact on the utilization of web basedkeeping money.

Hat and Oliveira (2016) suggested a model joining the DeLone& McLean ISachievement demonstrate alongside the Task Technology Fit (TTF) model to checkthe impact of m-depending on individual execution. They assembled data from 233people as the example through an online review survey. They distinguished thatutilization and client fulfilment are noteworthy points of reference of individualexecution, and the noteworthiness of the directing impacts of TTF over use toindividual execution. They likewise found that there is certain connection offramework quality, data quality, and administration quality on client fulfilment.

Tsionas and Mamatzakis (2016) proposed other criteria for estimation of specializedefficiency that considers conformity costs in factor inputs connected with changes inefficiency. They displayed that specialized efficiency depends on alteration costs infactor inputs. Assessing the proposed show had certain complexities which theyovercome by utilizing a non-parametric Local Linear Maximum Likelihood (LLML).They employed a complete worldwide keeping money test and evaluated bankelective profit efficiency over a plenty of nations with solid fluctuation in the hiddenalteration costs. The found that heterogeneity crosswise over nations confirmdemonstrates that change costs because of faculty costs are the most elevated amongcutting edge nations. They likewise saw that rising economies indicate solid potentialregarding efficiency post-budgetary emergency, for the most part because of lowerwork modification costs.

Veríssimo (2016) presented unpublished discoveries on m-keeping money application

use to distinguish the potential constraint that as of now confine its more extensiveselection. They employed fluffy set qualitative similar investigation (fsQCA) to testhow saw chance, similarity, saw value, saw usability, age and pay impact portablemanaging an account application utilize, and non-utilize. Investigation connected onexact data test of managing an account customers uncovered that versatile keepingmoney application utilize is in connection with high similarity, low saw chance, highsaw convenience, and high saw handiness. It was additionally found that anarrangement of low saw value, low similarity, low saw convenience, and a high sawhazard is an adequate condition for portable managing an account application nonutilize.

7. RESEARCH GAP

Research gap was identified after review of relevant literatures about customers attitude in banking for both global and national perspective. Many studies are carried out aboutcustomer and their perception only before COVID 19, but there is no *relevant* studyabout customer attitude towards banking in before and after Covid -19in banking in Tambaram city, so based on abovedefinition the research gap clearly identified and objectives of the study settled basedon the research gap.

8. RESEARCH METHODOLOGY

Sample of the study should represent population of the study. Sampling is one of the most methodological and imperative part of the research. Accurate sampling technique helps to correct estimation of replicas in the study. In probability analysis, the confidence level of test should be 95% which means dispersion is limited to ±2 of standard deviation of mean data. In scientific research, there are various ways to estimate sample size. Cochran‟s formula and Morgan‟s table are the most commonly used tools. In this research 480 respondents selected where as 360 numbers are sufficient number of observations based on Cochran‟ s formula and Morgan‟s table. This criterion provides more accuracy and confidence of tests and model estimations in this research. The geographical scope of this research is consumer attitude towards bankig before and after Covid 19 in Tambaram city. The researcher covered selected only four nationalized banks (Indian Bank, Indian Overseas Bank, Karuru Vysya bank and Canara bank) for the purpose of primary data collection. The researcher has selected random sample in each selected banks of Tambaram city and forthe present study the total sample was 480 people at random. In the focused group, through interview techniques, the sample of respondents were divided into Indian and non-Indian consumers and students, self-employed, businessmen/women and whoever uses the banking services in Tambaram city. The researcher has selected primary and secondary data. The primary data was collected through interview schedule secondary data were collected through journals, books, magazines and internet etc. The researcher has also using Chi square test for analysis of hypothesis of this study.

9. DATA ANALYSIS AND INTERPRETATION

Table -1 – Demographic profile of the Respondents

S. No	Demographic Profile	Factors	No. of Respondents	Percentage
1	Gender	Male	256	71.11
		Female	104	28.89
2	Age(in years)	Below 25	45	12.50
		25 – 35	188	52.22
		35 – 45	48	13.33
		45 – 55	57	15.83
		Above 55	22	6.11
3	Marital status	Married	288	80.00
		Unmarried	72	20.00
4	Occupation	Private employee	140	38.89
		Govt. Employee	42	11.67
		Business	120	33.33
		Farmer	58	16.11
5	Income level	Below Rs. 40,000	165	45.83
		Rs.40, 000 – Rs.50, 000	67	18.61
		Rs.50, 000 – Rs.60, 000	45	12.50
		Rs.60, 000 – Rs.70, 000	40	11.11
		Above Rs.70, 000	43	11.94
6	Qualification	Below Hr Sec	98	27.22
		UG	127	35.28
		PG	90	25.00
		Professional	45	12.50
7	Location	Urban	210	58.33
		Rural	85	23.61
		Semi urban	65	18.06

Source: Primary Data.

Above Table 1 clearly shows that demographic profile of the bank customers; out of 360 respondents 71.11 per cent of the customers are male and the remaining female because males are working category; 52.22 respondents are come under the age group of 25 to 35 years. In Tambaram most of the sample force are married followed by 38.89 per cent of the interviewees are work under the category of private employee (i.e. coolly, salaried people, Daily earnings people); 45.83 per cent of the customers are earned the monthly income of below Rs. 40,000 and the 35.28 per cent of the respondents qualified under UG level only very least per cent of the customers are professional category. The researcher further analyzed that location of the customers attitude towards online banking in the pandemic situations, 42.22 respondents are more difficult to using online banking for the semi urban customers; 35.56 per cent of the respondents were lived in urban area and the remaining customers are lives in rural area. The rural area and semi urban area customers were no using the inline banking for transfer of money and payment of loan and any bills.

Table 2 - Location of the banking customers

Sl. No	Location	No of Respondents				
		Indian Bank	Indian Overseas Bank	Canara Bank	Karur Vysya Bank	Total No of Customers (Percentage)
1	Rural	20	32	21	12	85 (23.61)
2	Urban	60	55	45	50	210 (58.33)
3	Semi Urban	10	3	24	28	65 (18.06)
Total		**90**	**90**	**90**	**90**	**360 (100.00)**

Source: Collected Data; ( ) indicates Percentage

Table 2 clearly exhibits that location of the banking customers, out of 360 respondents 58.33 per cent of the banking customers located in urban area, in these percentage most of the customers are maintaining their accounts in Indian bank followed by 23.61 per cent of the customers located in rural area in these rural area most of the customers maintaining their accounts in Indian overseas bank very least of the customers maintain their accounts in Karur Vysya Bank and the rest 18.06 per cent of the customers were located in semi urban area in these most of the semi urban area customers maintaining their bank accounts in Karur Vysya Bank.

Table 3 –Direct Usage banks for the customers before and after COVID-19

Sl No	Usage	Before Covid-19 (Percentage)	After Covid-19 (Percentage)
1	Used	318 (88.33)	80 (22.22)
2	Not used	42 (11.67)	280 (77.78)
Total		**360 (100)**	**360 (100)**

Source: Collected Data; () indicates percentage

From the above table 3 clearly analyzed that direct usage of banks for the customers before and after Covid 19 day to day activities; out of 360 respondents 88.33 per cent of the banking customers are used to banks before Covid -19 and the rest 11.67 per cent of the banking customers are not using the bank before Covid 19; After the pandemic situation i.e. Covid 19 most (77.78 %) of the customers are not using bank because fear of disease.

Table – 4 Uses of online banking for the pandemic situation

Sl no	Pandemic Situation	Uses	Not Used	No of respondents (Percentage)
1	Before Covid-19	70 (19.44)	290 (80.56)	360 (100)
2	After Covid-19	332 (92.22)	28 (7.78)	360 (100)

Source: Collected data; () Indicates Percentage

From the above table 4 clearly indicates that the customers using the online banking for the before and after pandemic situation. Out of 360 respondents 80.56 per cent of the customers are not using the online banking operations the customers go directly for the bank and made the online banking transactions. And the rest 19.44 per cent of the respondents were used the online banking transactions. After pandemic situation most of the customers were using the online banking transactions and the remaining 7.78 customers were not used the online banking transactions. For the pandemic situation many customers are learn about the digital adoption, mobility patterns, increased uses of mobile transaction for payment recharge, loan amount transfer of money etc.,

10. TESTING OF HYPOTHESIS

The researcher has applied and test the chi square test for the null hypothesis. H0 "There is no association between location and customers attitude for uses of online banking in the pandemic situations".

Before Pandemic situations

Sl no	Location of customers	Online banking		No of respondents (Percentage)
		Uses	Not Used	
1	Urban	65 (92.85%)	145 (50.00%)	210
2	Rural	1(1.43%)	84 (28.96%)	85
3	Semi urban	4 (5.71%)	61 (21.03%)	65
Total		**70 (100%)**	**290 (100%)**	**360**

Source: Collected data

After Pandemic situations

Sl no	Location of customers	Online banking		No of respondents (Percentage)
		Uses	Not Used	
1	Urban	200 (60.24%)	10 (35.71%)	210
2	Rural	78(23.49%)	7 (25.00%)	85
3	Semi urban	54 (16.27%)	11 (39.29%)	65
Total		**332 (100%)**	**28 (100%)**	**360**

Source: Collected data

Pandemic Situations	Calculated Chi Square Value	P value	Degrees of freedom	Significant level	Result
Before	44.07	5.991	5 % level	Not Significant	Rejected
After	10.235	5.991	5 % level	Not Significant	Rejected

Source: Computed Data

In this research 2 d.f at 5 % level of significance the table value is lesser than the calculated value for the before after pandemic situation so the hypothesis rejected. Hence concluded that "There is an association between location and customers attitude for uses of online banking in the pandemic situations".

Findings of the study

- 71.11 per cent of the customers are male and the remaining female because males are working category.
- 52.22 respondents are come under the age group of 25 to 35 years in the study area.
- In Tambaram most of the sample force are married followed by 38.89 per cent of the interviewees are work under the category of private employee (i.e. coolly, salaried people, Daily earnings people);
- It is vivid that Majority 45.83 per cent of the customers are earned the monthly income of below Rs. 40,000
- It is lucid that 35.28 per cent of the respondents qualified under UG level only very least per cent of the customers are professional category.
- 42.22 respondents are more difficult to using online banking for the semi urban customers; The rural area and semi urban area customers were no using the inline banking for transfer of money and payment of loan and any bills.
- Majority of the customers maintain the accounts in the Indian overseas and Indian bankin the rural area customers.

- 88.33 per cent of the banking customers are used to banks before Covid -19 and the rest 11.67 per cent of the banking customers are not using the bank before Covid 19; After the pandemic situation i.e. Covid 19 most (77.78 %) of the customers are not using bank because fear of diseases.

11. SUGGESTIONS AND CONCLUSION

The interviewer gives the suggestions are the customer attitude towards online banking and the banks communication regarding online transactions risk exposure related to mobile and internet banking services use (BCOT) have a direct. Consumers that have more confidence in using online banking showless interest in additional efforts of banks to communicate and minimize the risks related to online banking services. As the attitude of consumers towards online banking services becomes more and more positive in the pandemic situations. The government has increased the network server in the banks. Increasing the digital adoption and mobility pattern in the banks. The researcher concluded that the pandemic situations like COVID-19 has severely affected the global economy. The strict lockdown measures have also changed the daily live, including consumer behavior and attitude in banking sector. In this context, the purpose of this paper is to investigate the consumer attitude towards the pandemic situations. Customers are not knowing about the technology in rural and semi urban customers. So the banks are give awareness programmed and campaign has been conducted for using the technology.The experience of living through COVID-19 is changing the world in which we live and our behavior. Changes that provide positive experiences are likely to last longer, particularly those driven by convenience and well-being, such as digital adoption, value-based purchasing and increased health awareness.

REFERENCE

1. *Adapa, S., Roy, S.K., (2017). Consumers' post-adoption behaviour towards Internet banking: empirical evidence from Australia. Behaviour & Information Technology, 36(9), 970983. https://doi.org/10.1080/0144929X.2017.1319498.*
2. *Anneli Järvinen, R., (2014). Consumer trust in banking relationships in Europe. International Journal of Bank Marketing, 32(6), 551-566. https://doi.org/10.1108/IJBM-08-2013-0086.*
3. *Bratianu, C., Prelipcean, G., Bejinaru, R. (2020b). Exploring the latent variables which support SMEs to become learning organizations. Management & Marketing. Challenges for the Knowledge Society, 15(2), 154-171.*
4. *Buehler, K., Conjeaud, O., Giudici, V., Samandari, H., Serino, L., Vettori, M., Webanck, L., White, O., (2020). Leadership in the time of coronavirus: COVID-19 response and implications for banks, McKinesy&Company, Retrieved from https://www.mckinsey.com/industries/financial-services/our-insights/leadershipin-the-time-of-coronavirus-covid-19-response-and-implications-for-banks.*
5. *Cătoiu, I., Gârdan, D.A., Geangu, I.P., (2012). Customer loyalty - specific features of retail activity in Bucharest. Annales Universitatis Apulensis Series Oeconomica, 14(1), 278-285.*
6. *Mora, A.G., (2020). Patterns—and some implications—of Covid-19 financial sector policy interventions, World Bank Blogs, Retrieved from https://blogs.worldbank.org/psd/patterns-and-some-implications-covid-19financial-sector-policy-interventions.*

DOI No. : 10.5958/2394-9309.2022.00016.6
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

CONNECTED 2 – DOMINATING SETS AND CONNECTED 2 – DOMINATION POLYNOMIALS OF THE LOLLIPOP GRAPH $L_{m,1}$

Y.A. Shiny* and T. Anitha Baby**
*Research Scholar (Reg. No. 19213042092006) Department of Mathematics, Women's Christian College, Nagercoil, Tamil Nadu, India.
**Assistant Professor, Department of Mathematics, Women's Christian College, Nagercoil, Tamil Nadu, India.
E-mail : shinyjebalin@gmail.com, anithasteve@gmail.com

ABSTRACT
Let $G = (V, E)$ be a simple graph. Let $D_{c_2}(G, j)$ be the family of Connected 2−dominating sets in G with size j and $d_{c_2}(G, j) = |D_{c_2}(G, j)|$. Then the polynomial $D_{c_2}(G, x) = \sum_{j=\gamma_{c_2}(G)}^{|v(m)|} d_{c_2}(G, j)x^j$ is called the connected 2– domination polynomial of G. Let $D_{c_2}(L_{m,1}, j)$ be the family of Connected 2−dominating sets of the Lollipop graph $L_{m,1}$ with cardinality j and $d_{c_2}(L_{m,1}, j) = |D_{c_2}(L_{m,1}, j)|$. Then the Connected 2 − domination polynomial $D_{c_2}(L_{m,1}, x)$ of $L_{m,1}$ is defined as $D_{c_2}(L_{m,1}, x) = \sum_{j=\gamma_{c_2}(L_{m,1}, j)}^{|V(L_{m,1})|} d_{c_2}(L_{m,1}, j)x^j$ where $\gamma_{c_2}(L_{m,1})$ is the connected 2 – domination number of $L_{m,1}$. In this paper, we obtain a recursive formula for $d_{c_2}(L_{m,1}, j)$. Using this recursive formula we construct the Connected 2 − domination polynomial of $L_{m,1}$ as $D_{c_2}(L_{m,1}, x) = \sum_{j=3}^{m+1} d_{c_2}(L_{m,1}, j)x^j$, where $d_{c_2}(L_{m,1}, j)$ is the number of Connected 2−dominating sets of $L_{m,1}$ of cardinality j. Also some properties of this polynomial have been studied.

Keywords: Connected 2-dominating sets, Connected 2-domination number, Connected 2-domination polynomials, Lollipop graph.

1. INTRODUCTION:

Let $G = (V, E)$ be a simple graph of order, $|V| = m$. For any vertex $v \epsilon V$, the open neighbourhood of v is the set $N(v) = \{u \epsilon V / uv \epsilon E\}$ and the closed neighbourhood of V is the set $N[v] = N(v) \cup \{v\}$. For a set $S \subseteq V$, the open neighbourhood of S is $N(S) = U_{v \epsilon s} N(v)$ and the closed neighbourhood of S is $N(S) \cup S$.

A set $D \subseteq V$ is a dominating set of G, if $N[D] = V$ or equivalently, every vertex in $V - D$ is adjacent to atleast one vertex in D.

The domination number of a graph G is defined as the cardinality of a minimum dominating set D of vertices in G and is denoted by $\gamma(G)$.

A dominating set D of G is called a connected dominating set if the induced sub-graph $< D >$ is connected.The connected domination number of a graph G is defined as the cardinality of a minimum connected dominating set D of vertices in G and is denoted by $\gamma_c(G)$.

The Lollipop graph $L_{m,n}$ is the graph obtained by joining a complete graph k_m to a path graph P_n with a bridge and it is denoted by $L_{m,n}$.

As usual we use $\lfloor x \rfloor$ for the largest integer less than or equal to x and $\lceil x \rceil$ for the smallest integer greater than or equal to x. Also, we denote the set $\{1,2,\dots\dots\dots,m\}$ by $[m]$, throughout this paper.

2 . Connected 2 – Dominating sets of Lollipop Graph $L_{m,1}$

In this section, we state the connected 2 – domination number of the Lollipop graph $L_{m,1}$ and some of its properties.

Definition 2.1

The Lollipop graph $L_{m,1}$ is a graph consisting of a complete graph on m vertices and a path graph on 1 vertex connected with a bridge.

Definition 2.2

Let G be a simple graph of order m with no isolated vertices. A subset $D \subseteq V$ is a $2-$ dominating set of the graph G if every vertex $v \epsilon V - D$ is adjacent to atleast two vertices in D. A $2-$ dominating set is called a connected $2-$ dominating set if the induced subgraph $< D >$ is connected.

Definition 2.3

The cardinality of a minimum connected 2 – dominating set of G is called the connected 2 – domination number of G and is denoted by $\gamma_{C_2}(G)$.

Theorem 2.4

Let $L_{m,1}$ be the Lollipop graph with $m+1$ vertices.

Then $d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1}$ for all $m \geq 3$.

Proof:

Let $L_{m,1}$ be the Lollipop graph with $m+1$ vertices and $m \geq 3$.

Since $L_{m,1}$ contains $m+1$ vertices the number of subsets of $L_{m,1}$ with cardinality j is $\binom{m+1}{j}$. Every time $\binom{m}{j}$ number of subsets of $L_{m,1}$ with cardinality j and $\binom{m-1}{j-1}$ number of subsets of $L_{m,1}$ with cardinality $j-1$ are not connected 2 – dominating sets.

Hence $L_{m,1}$ have $\binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1}$ number of connected 2 – dominating sets of cardinality j.

Therefore, $d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1}$ for all $m \geq 3$.

Theorem 2.5

Let $L_{m,1}$ be the Lollipop graph with $m \geq 3$.Then

(i) $d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m-1}{j-2}$ for all $3 \leq j \leq m$.

(ii)$d_{c_2}(L_{m,1},\mathrm{j}) = \begin{cases} d_{c_2}(L_{m-1,1,}\, j) + d_{c_2}(L_{m-1,1,}j-1) \ if \ 3 < j \leq m \\ d_{c_2}(L_{m-1,1,}\, j) + 1 \ if \ j = 3 \end{cases}$

Proof:

(i) From Theorem 2.4, we have

$d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1}$ for all $m \geq 3$.

We know that, $\binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1} = \binom{m-1}{j-2}$

Therefore, $d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m-1}{j-2}$ for all $3 \leq j \leq m$.

(ii) From (i) $d_{c_2}(L_{m,1},\mathrm{j}) = \binom{m-1}{j-2}$

$d_{c_2}(L_{m-1,1},j) = \binom{m-2}{j-2}$ and

$d_{c_2}(L_{m-1,1}, j-1) = \binom{m-2}{j-3}$

We know that,

$$\binom{m-2}{j-2} + \binom{m-2}{j-3} = \binom{m-1}{j-2}$$

Therefore, $d_{c_2}(L_{m,1},\mathrm{j}) = d_{c_2}(L_{m-1,1},j) + d_{c_2}(L_{m-1,1},j-1)$

When $j = 3$,

$d_{c_2}(L_{m,1},3) = \binom{m-1}{3-2} = \binom{m-1}{1} = m-1$

Consider, $d_{c_2}(L_{m-1,1},3) + 1 = \binom{m-1-1}{1} + 1$

$$= \binom{m-2}{1} + 1$$

$$= m - 2 + 1$$

$$= m - 1$$

That is, $d_{c_2}(L_{m-1,}3) + 1 = d_{c_2}(L_{m,1,}3)$

Therefore, $d_{c_2}(L_{m,1},\mathrm{j}) = d_{c_2}(L_{m-1,1}, j) + 1 \; if \; j = 3$

3. Connected 2 – Domination Polynomials of Lollipop Graph $L_{m,1}$.

Definition 3.1

Let $D_{c_2}(L_{m,1},\mathrm{j})$ be the family of connected 2 – dominating sets of the Lollipop graph $L_{m,1}$ with cardinality j and $d_{c_2}(L_{m,1},\mathrm{j}) = |D_{c_2}(L_{m,1},\mathrm{j})|$. Then the connected 2 – domination polynomial $D_{c_2}(L_{m,1},x)$ of $L_{m,1}$ is defined as $D_{c_2}(L_{m,1},x) = \sum_{j=\gamma_{c_2}(L_{m,1},\mathrm{j})}^{m+1} d_{c_2}(L_{m,1},\mathrm{j})x^j$, where $\gamma_{c_2}(L_{m,1})$ is the connected 2 – domination number of $L_{m,1}$.

Remark 3.2

$\gamma_{c_2}(L_{m,1}) = 3$.

Proof:

Let $L_{m,1}$ be the Lollipop graph with $m+1$ vertices and $m \geq 3$ .Label the vertices of $L_{m,1}$ by $\{v_1, v_{2,} v_3, \dots \dots, v_m, v_{m+1}\}$ where $\{v_1, v_2, v_{3,} \dots \dots, v_m\}$ are the vertices of degree $m-1$ and v_{m+1} is a vertex of degree 1. Since v_{m+1} is adjacent to v_m only, every connected 2–dominating set must contain the vertices v_m, v_{m+1} and one more vertex from $\{v_1, v_2, \dots \dots, v_{m-1}\}$.

Hence, $\gamma_{c_2}(L_{m,1}) = 3$.

Theorem 3.3

Let $L_{m,1}$ be the Lollipop graph with $m+1$ vertices.Then the connected $2-$domination polynomial of $L_{m,1}$ is

$D_{c_2}(L_{m,1,}x) = (1+x)D_{c_2}(L_{m-1,1,}x) + x^3$ with $D_{c_2}(L_{3,1,}x) = 2x^3 + x^4$.

Proof:

We have, $D_{c_2}(L_{m,1,}x) = \sum_{j=3}^{m+1} d_{c_2}(L_{m,1,}\mathrm{j})x^j$

$D_{c_2}(L_{m,1,}x) = d_{c_2}(L_{m,1,}3)x^3 + \sum_{j=4}^{m+1}(L_{m,1,}\mathrm{j})x^j$

$D_{c_2}(L_{m,1,}x) = \begin{bmatrix} m-1 \\ 1 \end{bmatrix} x^3 + \sum_{j=4}^{m+1}\left[d_{c_2}(L_{m-1,1,}j) + d_{c_2}(L_{m-1,1,}j-1)\right] x^j$

by Theorem 2.5(ii)

$D_{c_2}(L_{m,1,}x) = (m-1)x^3 + \sum_{j=4}^{m+1} d_{c_2}(L_{m-1,1,}j)x^j$

` $+ \sum_{j=4}^{m+1} d_{c_2}(L_{m-1,1,}j-1)x^j$ ________(1)

Consider, $\sum_{j=4}^{m+1} d_{c2}(L_{m-1,1,}j)x^j = \left[\sum_{j=3}^{m+1} d_{c2}(L_{m-1,1,}j)x^j\right] - d_{c_2}(L_{m-1,1,}3)x^3$

$= D_{c_2}(L_{m-1,1,}x) - \binom{m-2}{1} x^3$

$= D_{c_2}(L_{m-1,1,}x) - (m-2)x^3$

Consider, $\sum_{j=4}^{m+1} d_{c2}(L_{m-1,1,}j-1)x^j = x\left[\sum_{j=4}^{m+1} d_{c2}\ (L_{m-1,1,}j-1)x^{j-1}\right]$

$= x\sum_{j=4}^{m+1} d_{c2}(L_{m-1,1},j-1)x^{j-1}$.

$= x\, D_{c_2}(L_{m-1,1,}x)$.

Hence from (1) we get

$D_{c_2}(L_{m,1,}x) = (m-1)x^3 + D_{c_2}(L_{m-1,1,}x) - (m-2)x^3 + xD_{c_2}(L_{m-1,1,}x)$

$D_{c_2}(L_{m,1,}x) = (1+x)D_{c_2}(L_{m-1,1,}x) + x^3$ with $D_{c_2}(L_{3,1,}x) = 2x^3 + x^4$.

Example 3.4

Let $L_{8,1}$ be the Lollipop graph $L_{m,1}$ with order 9 as given in Figure 2.1

$\boldsymbol{L_{8,1}}$ **:**

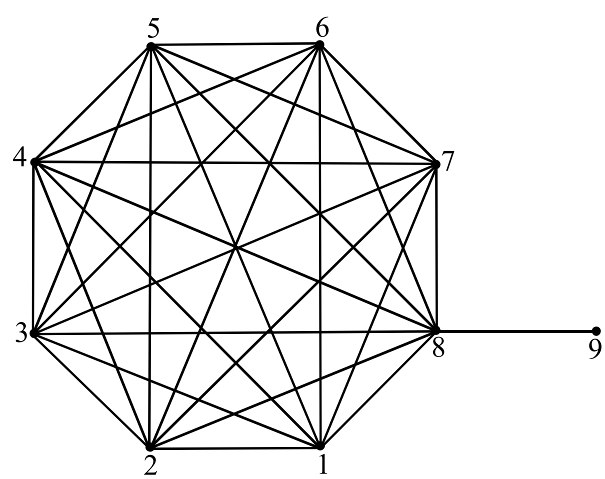


Figure 2.1

$D_{c_2}(L_{7,1,}x) = 6x^3 + 15x^4 + 20x^5 + 15x^6 + 6x^7 + x^8$

By Theorem 3.3,we have,

$D_{c_2}(L_{8,1,}x) = (1+x)D_{c_2}(L_{7,1,}x) + x^3$

$= (1 + x)(6x^3+15x^4+20x^5+15x^6+6x^7+x^8)$
$= 7x^3+21x^4+35x^5+35x^6+21x^7+7x^8+x^9.$

Theorem 3.5

Let $L_{m,1}$ be the Lollipop graph with $m \geq 3$.Then

$(i) D_{c_2}(L_{m,1,}x) = \sum_{j=3}^{m+1} \binom{m+1}{j} x^j - \sum_{j=3}^{m+1} \binom{m}{j} x^j - \sum_{j=3}^{m+1} \binom{m-1}{j-1} x^j.$

(ii) $D_{c_2}(L_{m,1,}x) = \sum_{j=3}^{m+1} \binom{m-1}{j-2} x^j.$

Proof:

(i) Since, $d_{c_2}(L_{m,1,}x) = \binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1}$, from the definition of Connected 2− domination Polynomial, we have the result.

(ii) Since, $\binom{m+1}{j} - \binom{m}{j} - \binom{m-1}{j-1} = \binom{m-1}{j-2}$, from the definition of Connected 2− domination Polynomial, we have the result.

We obtain $d_{c_2}(L_{m,1,}j)$ for 3≤ m≤ 15 and 3≤ j≤ 16 as shown in Table 1.

Table 1 $d_{c_2}(L_{m,1,}j)$, the number of Connected 2− dominating sets of $L_{m,1}$ with cardinality j.

m \ j	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3	0	2	1												
4	0	3	3	1											
5	0	4	6	4	1										
6	0	5	10	10	5	1									
7	0	6	15	20	15	6	1								
8	0	7	21	35	35	21	7	1							
9	0	8	28	56	70	56	28	8	1						
10	0	9	36	84	126	126	84	36	9	1					
11	0	10	45	120	210	252	210	120	45	10	1				
12	0	11	55	165	330	462	462	330	165	55	11	1			
13	0	12	66	220	495	792	924	792	495	220	66	12	1		
14	0	13	78	286	715	1287	1716	1716	1287	715	286	78	13	1	
15	0	14	91	364	100	2002	3003	3432	3003	2002	1001	364	91	14	1

In the following Theorem, we obtain some properties of $d_{c2}(L_{m,1,}j)$.

Theorem: 3.5

The following properties hold for the coefficients of $D_{c2}(L_{m,1,}x)$ for all m.

(i) $d_{c2}(L_{m,1,}m+1) = 1$, for every $m \geq 3$.

(ii) $d_{c2}(L_{m,1}\ ,m) = m-1$, for every $m \geq 3$.

(iii) $d_{c2}(L_{m,1}\ ,m-1) = \frac{1}{2}(m^2-3m+2)$, for every $m \geq 4$.

(iv) $d_{c2}(L_{m,1}\ , m-2) = \frac{1}{6}(m^3 - 6m^2 + 11m - 6)$, for every $m \geq 5$.

(v) $d_{c2}\left(L_{m,1}\ , m-3\right) = \frac{1}{24}(m^4 - 10m^3 + 35m^2 - 50m + 24)$, for every $m \geq 6$.

(vi)$\left(L_{m,1}\ , 3\right) = \ m-1$, for every $m \geq 3$.

Proof:

i) Since,$d_{c2}\left(L_{m,1}\ , m+1\right) = [m+1]$, we have the result.

ii)To Prove,$d_{c2}\left(L_{m,1}\ , m\right) = m-1$, for every $m \geq 3$, we apply induction on m.

When $m = 3$,

L.H.S =$d_{c2}\left(L_{3,1}\ , 3\right) = 2$ (from the table) and

R.H.S = 3−1= 2

Therefore, the result is true for m = 3.

Now, suppose that the result is true for all natural numbers less than ‘ m’ and prove it form.

By Theorem 2.5,(ii) we have

$d_{c2}\left(L_{m,1}\ , j\right) = \ dc_2\left(L_{m-1,1}\ , m\right) + d_{c2}(L_{m-1,1}\ , m-1)$.

$= 1 + m - 2$.

$= m - 1$.

Hence the result is true for all $m \geq 3$.

(iii) To prove, $dc_2\left(L_{m,1}\ , m-1\right) = \frac{1}{2}(m^2 - 3m + 2,)$ for every $m \geq 4$, we apply induction on m.

When $m = 4$,

L.H.S= $dc_2\left(L_{4,1}, 4-1\right) = dc_2\left(L_{4,1}, 3\right) = \ 3$ (from the table) and

R.H.S= $\frac{1}{2}[m^2 - 3m + 2] = \frac{1}{2}[4^2 - 3\ \times 4 + 2] = 3$.

Therefore, the result is true for $m = 4$.

Now, suppose that the result is true for all numbers less than ‘mm’ and we prove it for .

By Theorem 2.5, (ii) we have,

$$dc_2\left(L_{m,1}\ , m-1\right) = dc_2\left(L_{m-1,1}\ , m-1\right) + dc_2\left(L_{m-1,1}\ , m-2\right)$$

$= m - 2 + \frac{1}{2}[(m-1)^2 - 3(m-1) + 2]$.

$= \frac{1}{2}[m^2 - 2m + 1 - 3m + 3 + 2 + 2m - 4]$.

$= \frac{1}{2}[m^2 - 3m + 2]$.

Hence the result is true for all $m \geq 4$.

(iv) To prove, $dc_2\left(L_{m,1,}\ m-2\right) = \frac{1}{6}(m^3 - 6m^2 + 11m - 6)$, for every $m \geq 5$, we apply induction on m.

When $m \geq 5$,

L.H.S = $dc_2\left(L_{5,1,}3\right) = 4$ (from the table) and

R.H.S = $\frac{1}{6}[5^3 - 6 \times\ 5^2 + 11\ \times\ 5 - 6] = 4$

Therefore, the result is true for m =5.

Now, Suppose that the result is true for all numbers less than ‘mm’ and we prove it for m.

By Theorem 2.5,(ii) we have

$dc_2(L_{m,1}, m-2) = dc_2(L_{m-1,1}, m-2) + dc_2(L_{m-1,1}, m-3)$.
$= \frac{1}{2}[(m-1)^2 - 3(m-1) + 2] + \frac{1}{6}[(m-1)^3 - 6(m-1)^2 + 11(m-1) - 6]$.
$= \frac{1}{2}[m^2 - 2m + 1 - 3m + 3 + 2] + \frac{1}{6}[(m^3 - 3m^2 + 3m - 1] - 6(m^2 - 2m + 1) + 11m - 11 - 6]$.
$= \frac{1}{6}[(m^3 - 3m^2 + 3m - 1 - 6m^2 + 12m - 6 + 11m - 11 - 6 + 3m^2 - 15m + 18]$.
$= \frac{1}{6}[m^3 - 6m^2 + 11m - 6]$.
Hence, the result is true for all $m \geq 5$.
(v) To prove, $d_{c2}(L_{m,1}, m-3) = \frac{1}{24}(m^4 - 10m^3 + 35m^2 - 50m + 24)$, for every $m \geq 6$, we apply induction on m.
When $m = 6$,
L.H.S = $dc_2(L_{6,1,} 6-3)$ = $dc_2(L_{6,1,} 3)$ = 5(from the table) and
R.H.S = $\frac{1}{24}[6^4 - 10 \times 6^3 + 35 \times 6^2 - 50 \times 6 + 24]$
$= \frac{1}{24}[1296 - 2160 + 1260 - 300 + 24]$
$= \frac{1}{24}[120]$
= 5.
Therefore ,the result is true for $m = 6$.
Now, Suppose that the result is true for all numbers less than 'm' and we prove it for m .
By Theorem 2.5, (ii) we have,
$dc_2(L_{m,1}, m-3) = dc_2(L_{m-1,1}, m-2) + dc_2(L_{m-1,1}, m-4)$.
$= \frac{1}{6}[(m-1)^3 - 6(m-1)^2 + 11(m-1) - 6] +$
$\frac{1}{24}[(m-1)^4 - 10(m-1)^3 + 35(m-1)^2 - 50(m-1) + 24]$.
$= \frac{1}{6}[(m^3 - 3m^2 + 3m - 1) - 6(m^2 - 2m + 1) + 11m - 66 - 6] +$
$\frac{1}{24}[(m^4 - 4m^3 + 6m^2 - 4m + 1) - 10(m^3 - 3m^2 + 3m - 1) +$
$35(m^2 - 2m + 1) - 50m + 50 + 24]$.
$= \frac{1}{24}[m^4 - 14m^2 + 77m^2 - 154m + 120 + 4m^3 - 36m^2 + 104m - 96]$
$= \frac{1}{24}[m^4 - 10m^3 + 35m^2 - 50m + 24]$.
Therefore, the result is true for all $m \geq 6$.
(vi) To Prove, $dc_2(L_{m,1}, 3) = m - 1$, for every $m \geq 3$,we apply induction on m.
When m =3,
L.H.S = $dc_2(L_{3,1,} 3)$ = 2(from the table) and
R.H.S = 3−1=2.
Therefore , the result is true for $m = 3$.
Now, Suppose that the result is true for all numbers less than 'm' and we prove it for m .
By Theorem 2.2, We have, $dc_2(L_{m,1}, 3) = dc_2(L_{m-1,1}, 3)$+1.
$= m$ −2+1.
$= m$ −1.
Therefore, the result is true for all $m \geq 3$.

CONCLUSION:

In this paper, the Connected 2− domination polynomials of Lollipop graph $L_{m,1}$ has been derived by identifying its Connected 2− dominating sets.It also help us to characterize the connected Connected 2− dominating sets of cardinality j.We can generalize this study to any of Lollipop graph $L_{m,n}$ and some interesting properties can be obtained.

REFERENCES:

1. *Alikhani.S,peng.Y.H."Introduction to Domination Polynomial of a Graph," ArsCombinatoria ,Available as ar xiv:0905.2251|v| [math.co]14 May 2009.*
2. *Sahib.Sh.Kahat.Abdul Jalil M.Khalaf and RoslanHasni,"Dominating sets and Domination Polynomials of Wheels".Asian Journal of Applied Sciences(ISSN:2321− 0893),volume 02− Issue 03,June 2014.*
3. *Sahib ShayyalKahat,Abdul Jalil M. Khalaf and RoslanHasni,"Dominating sets and Domination Polynomials of Stars".Australian Journal of Basics and Applied Science,8(6) June 2014,pp 383−386.*
4. *A.Vijayan,T.AnithaBaby,G.Edwin,Connected Total Dominating Sets and Connected Total Domination Polynomials of Stars and Wheels,IOSR Journal of Mathematics, VolumeII,pp 112*−121.
5. *A.Vijayan,T.AnithaBaby,G.Edwin,Connected Total Dominating Sets and Connected Total Domination Polynomials of Fan Graphs* $F_{2,n}$ *.International Journal of Math. Sci & Engg.Appls (IJMSEA),Volume 10,pp 135*−146.

DOI No. : 10.5958/2394-9309.2022.00017.8
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

THE UPPER DETOUR COTOTAL DOMINATION NUMBER OF A GRAPH

S. L. Sumi*, V. Mary Gleeta, J. Befija Minnie*****
*Research Scholar, Register No.20123042092007, Department of Mathematics, Holy Cross College, Nagercoil - 629 004, India.
**Assistant Professor, Department of Mathematics, T.D.M.N.S College, T. Kallikulam – 627 113, India.
***Assistant Professor, Department of Mathematics, Holy Cross College, Nagercoil - 629 004, India.
E-mail : sumikrish123@gmail.com

ABSTRACT

A detour cototal dominating set cototal dominating set of G if no proper subset of S is a detourcototal dominating set of G. The upper detour cototal domination number$\gamma^{+}_{dct}(G)$ of G is the maximum cardinality of a minimal detour cototal dominating set of G. Some general properties satisfied by this concept are studied. For a connected graph G of order n with upper detour cototal domination number$n$ is characterized. It is shown that for every two positive integers a and b,with$2 \leq a \leq b$, there exists a connected graph G with $\gamma_{dct}(G) = a$ and $\gamma^{+}_{dct}(G) = b$.

Keywords: detour set, cototal domination, detour cototal domination, upper detour cototal domination.
AMS Subject Classification: 05C12,05C69.

1. INTRODUCTION

For a graph $G = (V, E)$, we mean a finite, undirected graph without loops ormultiple edges. The order and size of G are denoted by m and n respectively. We consider connected graph with at least two vertices. For basic definitions and terminologies, we refer to West [2,11].

For vertices u and v in a connected graph G, the detour distance $D(u, v)$ isthe path of the longest $u - v$ path in G. A $u - v$ path of length $D(u, v)$ is called a $u - v$detour. It is known that the detour distance is a metric on the vertex set V(G).

.These concepts were studied by Chartrand et. al [3,4,5].

A vertex x is said to lie on a $u - v$ detour P if x is a vertex of $u - v$ detour path Pincluding the vertices u and v. A set $S \subseteq V(G)$ is called a detour set if every vertex$v$ in G lies on a detour joining a pair of vertices of S. The detour number $dn(G)$ of$G$ is the minimum order of a detour set and any detour set of order $dn(G)$ is calleda minimum detour set of G. These concepts were studied by G. Chartrand et. al [6,7,8].

A dominating set of a graph $G = (V, E)$ is a cototal dominating set if everyvertex $v \in V \setminus D$ is not an isolated vertex in the induced subgraph $< V \setminus D >$. The cototal domination number $\gamma_{ct}(G)$ $of\ G\ \ is\ the$ minimum cardinality of a cototal dominating sets. The cototal domination number ofa graph was studied in [9,10]. A set $S \subseteq V$ is called a detour cototal dominating set if$S$ is both a detour and a cotoal dominating set of G. The detour cototal dominationnumber γ_{dct}(G) is the minimum size of a detour cototal dominating set of G [ ]. Thefollowing Theorem is used in sequel.

Theorem 1.1. [8] Each end vertex of G belongs to every detour cototal dominating set of$G$.

2. THE UPPER DETOUR CO TOTAL DOMINATION NUMBER OF A GRAPH

Definition 2.1. A detour cototal dominating set S in a connected graph G is calleda minimal detour cototal dominating set of G if no proper subset of S is a detourcototal dominating set of G. The maximum cardinality of a minimal detour cototaldominating set of G is the upper detour cototal domination number of G and isdenoted by $\gamma_{dct}^{+}(G)$.

Example 2.2. For the graph G given in Figure 2.1, $S_1 = \{v_1, v_4, v_7, v_9\}$, $S_2 = \{v_1, v_4, v_7, v_{10}\}$, $S_3 = \{v_2, v_4, v_7, v_9\}$ and $S_4 = \{v_2, v_4, v_7, v_{10}\}$ are the only four minimaldetour cototal dominating sets of G, so that $\gamma_{dct}(G)$= 4. Also $S_5 = \{v_1, v_4, v_6, v_9, v_{11}, v_{14}\}$, $S_6 = \{v_1, v_4, v_6, v_{10}, v_{11}, v_{14}\}$, $S_7 = \{v_2, v_4, v_6, v_9, v_{11}, v_{14}\}$, $S_8 = \{v_2, v_4, v_6, v_{10}, v_{11}, v_{14}\}$, $S_9 = \{v_1, v_5, v_7, v_9, v_{12}, v_{13}\}$, $S_{10} = \{v_1, v_5, v_9, v_{10}, v_{12}, v_{13}\}$, $S_{11}\{v_2, v_5, v_7, v_9, v_{12}, v_{13}\}$ and $S_{12} = \{v_2, v_5, v_7, v_{10}, v_{11}, v_{13}\}$ are some minimal detour cototal dominating sets of G, so that $\gamma_{dct}^{+}(G) \geq 6$. It can be easily verified that there is no minimal detourcototal dominating set of G with cardinality more than seven. Therefore $\gamma_{dct}^{+}(G) = 6$.

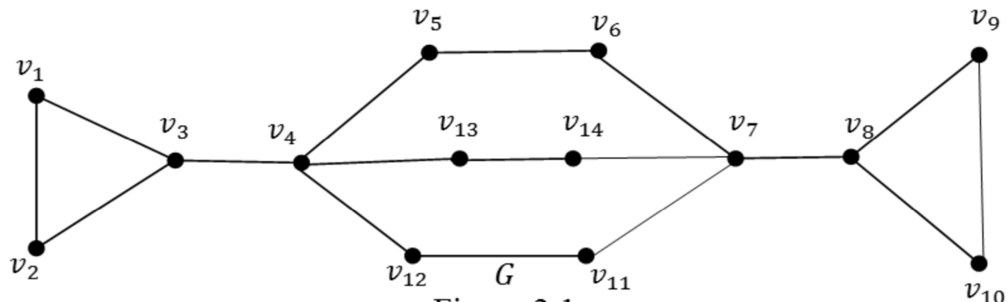


Figure 2.1

Theorem 2.3. Let G be a connected graph of order n, $2 \leq \gamma_{dct}(G) \leq \gamma_{dct}^{+}(G) \leq n$.

Proof: Since every detour cototal dominating set of G needs at least two vertices.Therefore $\gamma_{dct}(G) \geq 2$. Since every minimum detour cototal dominating set of G is a minimaldetour cototal dominating set of G. It follows that $\gamma_{dct}(G) \leq \gamma_{dct}^{+}(G)$. Also,since $V(G)$ is a detour cototal dominating set of G, we have $\gamma_{dct}^{+}(G) \leq n$.

Therefore $2 \leq \gamma_{dct}(G) \leq \gamma_{dct}^{+}(G) \leq n$. ∎

Remark 2.4. The bounds in Theorem 2.3 is strict. For $G = C_4, \gamma_{dct}(G) = 2$. For adouble star G,$\gamma_{dct}^{+}(G) = 2 = n$. Also the bounds in Theorem 2.3 can be strict. For thegraph G given in Figure 2.1, $\gamma_{dct}(G) = 4$, $\gamma_{dct}^{+}(G) = 6$and $n = 12$. Thus $2 < \gamma_{dct}(G) < \gamma_{dct}^{+}(G) < n$.

Theorem 2.5. For any connected graph G of order $n \geq 2$. Then $\gamma_{dct}(G) = n$ if and only if $\gamma_{dct}^{+}(G) = n$.

Proof: If $\gamma_{dct}(G) = n$, then by Theorem 2.4, $\gamma_{dct}^{+}(G) = n$. Conversely let$\gamma_{dct}^{+}(G) = n$.Then $S = V(G)$ is the unique minimal detour cototal dominating set and so it is aminimum detour cototal dominating set of G. Hence $\gamma_{dct}(G) = n$. ∎

Observation 2.6. (i)For the complete graph $G = K_n\ n \geq 2, \gamma^{+}_{dct}(G) = 2$.
(ii) For the path $G = P_n(n \geq 5), \gamma^{+}_{dct}(G) = \left\lceil\frac{n}{2}\right\rceil$.
(iii) For the cycle $G = C_n\ (n \geq 6), \gamma^{+}_{dct}(G) = \left\lceil\frac{n}{2}\right\rceil$.
(iv) For any non-trivial tree T with k-end vertices. $\gamma^{+}_{dct}(G)=\begin{cases} k & if\ T\ is\ not\ a\ star \\ k+1 & if\ T\ is\ star \end{cases}$

Theorem 2.7. Let G be the complete bipartite graph $K_{r,s}$, then
$\gamma^{+}_{dct}(G) = \begin{cases} n & if\ 1 \leq r \leq s \\ 2 & if\ \ 4 \leq r \leq s \end{cases}$

Proof: If $r = 1$, then $S = V\ (G)$ is the unique detour cototal dominating set of Gso that $\gamma^{+}_{dct}(G)= n$. Let $2 \leq r \leq s$, Let $U = \{u_1, u_2, \ldots, u_r\}$ and $V = \{v_1, v_2, \ldots,\ v_s\}$be the two bipartite sets of G. Let $u \in U$ and $v \in V$. Then $S = \{u, v\}$ is a detourcototal dominating set of G. we prove that $\gamma^{+}_{dct}(G)$= 2. On the contrary supposethat $\gamma^{+}_{dct}(G) \geq 3$. Then there exists a detour cototal dominating set S' of G such that$|S'| \geq 3$. Then $S' \subseteq U$ or $S' \subseteq V$. Hence neither $S'$is a detour set of G nor a cototaldominating set of G, which is a contradiction. Therefore $\gamma^{+}_{dct}(G)$= 2.■

Theorem 2.8. For the triangular snake graph $G = T\ S_n, \gamma^{+}_{dct}(G) = \left\lfloor\frac{n}{2}\right\rfloor$.

Proof: Let P_n: $v_1,\ v_2, \ldots, v_n$ be a path on n vertices. $G = TS_n$ be the graph obtained from $P_{n-1}$by adding the new vertices $w_1, w_2, \ldots, w_{\lfloor\frac{n}{2}\rfloor}$ and introducing the edges $viwi$ and $v_{i-1}w_i$for $1 \leq i \leq \left\lfloor\frac{n}{2}\right\rfloor$. Let $W = \{w_1, w_2, \ldots, w_{\lfloor\frac{n}{2}\rfloor}\}$. Then W is a detour cototaldominating set of G. We prove that W is a minimal detour cototal dominating set of$G$. On the contrary suppose that W is not a minimal detour cototal dominating setof G. Then there exists a detour cototal dominating set W_1 such that $W_1 \subset W$. Let$x \in W_1$ such that $x \notin W$. Let y, z be two vertices in P_{n-1} which are adjacent to x.Then either y or z is not an element of W_1, which is a contradiction. Therefore W isa minimal detour cototal dominating set of G and so $\gamma^{+}_{dct}(G) \geq \left\lfloor\frac{n}{2}\right\rfloor$. We prove that$\gamma^{+}_{dct}(G)= \left\lfloor\frac{n}{2}\right\rfloor$. On the contrary suppose that $\gamma^{+}_{dct}(G) \geq \left\lfloor\frac{n}{2}\right\rfloor + 1$. Then there existsa detour cototal dominating set M of G such that $|M| \geq \left\lfloor\frac{n}{2}\right\rfloor + 1$.Since $W\ M$,there exists two vertices $v_i'$and $v_j' \in P_{n-1}$such that $v_i', v_j' \in M$. If $v_i'$and v_j'areadjacent then $\langle V \setminus M\rangle$ contains isolated vertices. If $v_i'$and $v_j'$are not adjacent, theneither $M = \{v_i'\}$ or $M = \{v_j'\}$ is a detour cototal dominating set of G, which is acontradiction. Therefore $\gamma^{+}_{dct}(G)= \left\lfloor\frac{n}{2}\right\rfloor$. ■

Proof: Let $V\ (K_2 \circ P_n))$: $\{v_1,\ v_2, \ldots, v_n\} \cup \{u_1, u_2, \ldots, u_n\}$. Then $S = \{v_1,\ v_2, \ldots, v_n\}$ and $M = \{u_1, u_2, \ldots, u_n\}$ is a detour cototal dominating sets γ_{dct}-sets of G. We Prove that $\gamma_{dct}(G) = n$. On the contrary suppose that $\gamma^{+}_{dct}(G) \geq n + 1$. Then there existsa minimal detour cototal dominating set S' of G such that $|S'| \geq n + 1$. Then$S' \subsetneq S \cup M$. Since $S \nsubseteq S'$ and $M \nsubseteq S'$, it follows that either $\langle V - S'\rangle$ containsisolated vertices or S' is not a dominating set of G, which is a contradiction. Therefore$\gamma^{+}_{dct}(G) = n$.■

Theorem 2.10. For the crown graph $G, \gamma^{+}_{dct}(G) = 3$.

Proof: Let $V\,(C_n) = \{v_1, v_2, \ldots, v_n\}$. Let $S_i = \{v_i, v_{i+1}, v_{i+3}\}$. Then Siis a minimaldetour cototal dominating set of G so that $\gamma^+_{dct}(G) \geq 3$. We prove that $\gamma^+_{dct}(G) = 3$.On the contrary suppose that $\gamma^+_{dct}(G) \geq 4$. Then there exists a minimal detour cototal
dominating set S$'$ of G such that $|S'| \geq 4$. Hence it follows that there exists at leastone detour cototal dominating set S'', such that $S'' \subset S'$, which is a contradiction.Therefore $\gamma^+_{dct}(G) = 3$. ■

Theorem 2.11. For the helm graph $G = H_r, \gamma^+_{dct}(G) = r + 1$.

Proof: Let x be the central vertex of G and z be the set of r end vertex of G. ByTheorem 1.1, Z is a subset of every detour cototal dominating set of G. Since $x$is not dominated by any vertex of Z, Z is not a cototal dominating set of G and so$\gamma_{dct}(G) \geq r + 1$. Let $Z' = Z \cup \{x\}$. Then $I[Z'] = V$ and $V - Z'$ has no isolated vertices. Therefore $Z'$is the unique detour cototal dominating set of G so that$\gamma^+_{dct}(G)$= $r + 1$. ■

Theorem 2.12. For the banana tree graph $G = B_{r,s,}\gamma^+_{dct}(G) = r(s - 1) - 1$.

Proof: Let x be the central vertex of G and y be the adjacent vertex of x and Z bethe set of cut vertices of G such that $x \notin Z$. Then $S = V\,(G) - \{x, y\} - Z$ is a minimaldetour cototal dominating set of G and so $\gamma^+_{dct}(G)$≥ $r(s - 1) - 1$. We prove that$\gamma^+_{dct}(G)$= $r(s-1) - 1$. On the contrary suppose that $\gamma^+_{dct}(G)$≥ $r(s-1) - 1$. Then thereexists a minimal detour cototal dominating set S' of G such that $|S'| \geq r(s - 1)$. Let$u \in S'$such that $u \notin S$. If $u \in Z$. Then $S' - \{u\}$ is a detour cototal dominating setof G, which is a contradiction. If u is adjacent to x, then $\langle V - S' \rangle$ contains isolatedvertex, which is a contradiction. If $u = x$, then $S' = \{u\}$ is the detour cototaldominating set of G, which is a contradiction. Therefore $\gamma^+_{dct}(G) = r(s - 1) - 1$. ■

Theorem 2.13. For the wheel graph $G = K_1 + C_{n-1} (n \geq 5), \gamma^+_{dct}(G) = \left\lceil \frac{n-1}{2} \right\rceil$.

Proof: Let x be the central vertex of G and C_{n-1}be $v_1, v_2, \ldots, v_{n-1}, v_1$. Then $S_i = \{x, v_i\}\ (1 \leq i \leq n - 1)$ is a γ_{dct}-set of G.

Case(i) Let $n - 1$ be even. Then $M = \{v_1, v_3, v_5, \ldots, v_{n-2}\}$ is a minimal detour cototal
dominating set of G and so $\gamma^+_{dct}(G)$≥$\left\lceil \frac{n-1}{2} \right\rceil$. We prove that $\gamma^+_{dct}(G)$= $\left\lceil \frac{n-1}{2} \right\rceil$.On the contrary, suppose that $\gamma^+_{dct}(G) \geq \left\lceil \frac{n-1}{2} \right\rceil + 1$. Then there exists a detourcototal dominating set M'such that $|M'| \geq \left\lceil \frac{n-1}{2} \right\rceil + 1$. Since $S_i\ (1 \leq i \leq n - 1)$ isa detour cototal dominating set of $Gx \notin M'$. Let $u \in M'$such that $u \notin M$. Then there exists $v_i' \in M'$such that $uv_i' \in E(G)$. Then $M'' = M' - \{u\}$ is a detour cototaldominating set G, which is a contradiction. Therefore $\gamma^+_{dct}(G)$= $\left\lceil \frac{n-1}{2} \right\rceil$. ■

Case(ii) Let$n - 1$ be odd. By similar argument as in case (i), we prove that $\gamma^+_{dct}(G) = \left\lceil \frac{n-1}{2} \right\rceil$.

Theorem 2.14. For the Fan graph $G = K_1 + P_{n-1}, \gamma^+_{dct}(G) = \left\lceil \frac{n-1}{2} \right\rceil$.

Proof: Let x be the central vertex of G and $V(P_{n-1}) = \{v_1, v_2, \ldots, v_{n-1}\}$.

Case(i) Let $n - 1$ be even. Then $S_1 = \{x, v_1\}$, $S_2 = \{x, v_{n-1}\}, S_3 = \{x, v_{\frac{n-1}{2}}\}S_4 =$

$\{x, u, v\}$ are some detour cototal dominating set of G, where $u \in \{v_2, v_3, \ldots, v_{\frac{n-1}{2}-1}\}$ and $v \in \{v_{\frac{n-1}{2}+1}, v_{\frac{n-1}{2}+2}, \ldots, v_{n-2}\}$. Let $S = \{v_1, v_3, \ldots, v_{n-2}\}$. Then S is aminimal detour cototal dominating set of G and so $\gamma_{dct}^{+}(G) \geq \left\lceil\frac{n-1}{2}\right\rceil$. We provethat $\gamma_{dct}^{+}(G) \geq \left\lceil\frac{n-1}{2}\right\rceil$.
On the contrary, suppose that $\gamma_{dct}^{+}(G) \geq \left\lceil\frac{n-1}{2}\right\rceil + 1$.Then there exists a detour cototal dominating set M'such that $|M'| \geq \left\lceil\frac{n-1}{2}\right\rceil + 1$.Since $S_i(1 \leq i \leq n-1)$ is a detour cototal dominating set of $G, x \notin M'$. Let$u \in M'$such that $u \notin M$. Then there exists $v_i' \in M'$such that $uv_i' \in E(G)$. Then$M'' = M' - \{u\}$ is a detour cototal dominating set of G, which is a contradiction.Therefore $\gamma_{dct}^{+}(G) = \left\lceil\frac{n-1}{2}\right\rceil$.

Case(ii) Let $n-1$ be odd. By similar argument as in case (i), we prove that $\gamma_{dct}^{+}(G) = \left\lceil\frac{n-1}{2}\right\rceil$. ■

Theorem 2.15. For every pair of a, b with $3 \leq a \leq b$, there exists a connected graph G such that $\gamma_{dct}(G) = a$ and $\gamma_{dct}^{+}(G) = b$.

Proof: Let $V(\overline{K}_2) : \{u, v\}$.Let $P_i: u_i, v_i$ $(1 \leq i \leq b-a+1)$ be a copy of path on two vertices. Let H be a graph obtained from K_2 and P_i $(1 \leq i \leq b-a+1)$by joining u with u_i $(1 \leq i \leq b-a+1)$. Let G be the graph obtained from H byintroducing new vertices $z_1, z_2, \ldots, z_{a-2}$ and introducing the edges vzi $(1 \leq i \leq a-2)$. The graph G is shown in Figure 2.2.Type equation here.

First we prove that $\gamma_{dct}(G) = a$. Let $Z = \{z_1, z_2, \ldots, z_{a-2}\}$ be the set of endvertices of G. Then by Theorem 1.1, Z is a subset of every detour cototal dominatingset of G. Since Z is not a detour set of G, $\gamma_{dct}(G) \geq a-1$. It is easily observed thatthere is no detour cototal dominating set of cardinality $a-1$ and so $\gamma_{dct}(G) \geq a$. Let $S = Z \cup \{u, v\}$. Then S is a detour cototal dominating set of G so that $\gamma_{dct}(G) = a$.

Next we prove that $\gamma_{dct}^{+}(G) = b$. Let $S' = Z \cup \{u, u_1, u_2, \ldots, u_{b-a+1}\}$. Then S'isa detour cototal dominating set of G. We prove that $S'$is a minimal detour cototaldominating set of G. On the contrary suppose that $S'$is not a minimal detour cototaldominating set of G. Then there exist a detour cototal dominating set of G. We provethat
$S'$is a minimal detour cototal dominating set of G. On the contrary, supposethat $S'$is not a minimal detour cototal dominating set of G. Then there exists adetour cototal dominating set of S_1 such that $S_1 \subset S'$. By Theorem 1.1, $Z \subset S_1$. Let$x \in S'$such that $x \notin S_1$. Therefore $x \neq z_i$for some i $(1 \leq i \leq a-2)$ then at least oneu_i $(1 \leq i \leq b-a+1)$ not belong to S'. Hence S_1 is not a detour cototal dominatingset of G, which is a contradiction. If $x = u_i$for some i $(1 \leq i \leq b-a+1)$. Then$x$ is not dominated by any element of S_1, which is a contradiction. Therefore S'isa minimal detour cototal dominating of G. Therefore $\gamma_{dct}^{+}(G) = b$. On the contrary$\gamma_{dct}^{+}(G) \geq b+1$. Then there exists a detour cototal dominating set S'' such that$|S''| \geq b+1$. Hence it follows that there exists at least one detour cototal dominatingset

S''' such that$S''' \subset S''$, which is acontradiction. Therefore $\gamma^{+}_{dct}(G) = b$.■

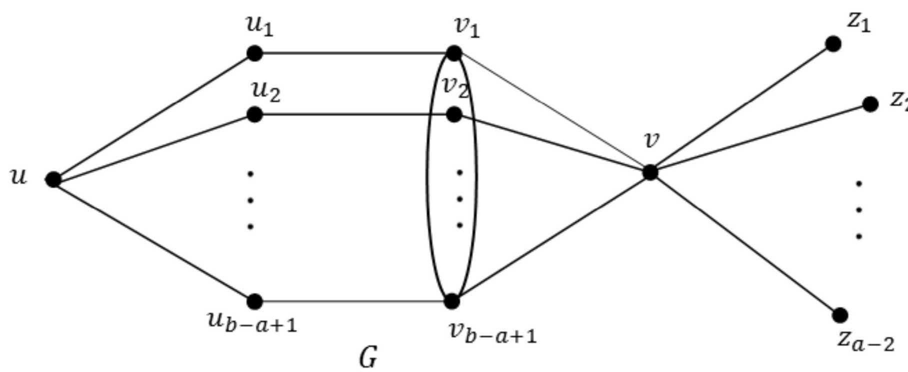


Figure 2.2

REFERENCES

1. *B.Basavanagoud and S.M. Hosamani, Complete Cototal Domination Number of a Graph, Journal of Scientific Research, 3(3),(2011), 547-555.*
2. *F.Buckley and F.Harary, Distance in Graphs, Addition- Wesley, Redwood City,CA, 1990.*
3. *G.Chartrand,G.L.Johns and P.Zhang, The detour number of a graph ,Utilitas Mathematica, 64(2008),97-113.*
4. *G. Chartrand,T.W. Haynes,M.A.Henning and P.Zhang,Detour Domination in Graphs ,Ars. Combinatoria,(2004),149-160,*
5. *G. Chartrand, H.Escuadro and P.Zhang,Detour Distance in graphs,J.Com-bin,computer,53(2005),75-94.*
6. *T.W. Haynes, P.J. Slater and S.T. Hedetniemi, Fundamentals of domination in graphs,Marcel Dekker,NewYork,1998.*
7. *J.John and N.Arianayagam,The detour domination number of a graph,Discrete Mathematics, Algorithms and Applications,09(01),(2017),1750006,(1-7).*
8. *J.John and N.Arianayagam,The Upper detour domination number of a graph, ,International Journal of Engineering Science Advanced Computing and Bio-Technology,8(1),(2017),24*
9. *V.R. Kulli, B.Janakiram and Radha Rajamani Iyer, The Cototal domination number of a graph, Journal of Discrete Mathematical Sciences and Cryptography 2(2),(1999), 179-184.*
10. *K. Uma Samundesvari and J. Maria Regila Baby, Complete cototal domination number of certain graphs, International Journal of Recent Technology and Engineering (IJRTE), (2019), 8,(1C2), 1107-1109.*
11. *D.M.West, Introduction to graph theory, 2nd edition, Upper Saddle River, NJ: Prentice-Hall, 2001.*
12. *S.L.Sumi,V.Mary Gleeta,J.Befija Minnie,The Geodetic Cototal Domination Number of a Graph ,ICDM2021,ISBN:978-93-91077-53-2*
13. *S.L.Sumi,V.Mary Gleeta,J.Befija Minnie, The Detour Cototal Domination Number of a Graph(communicated)*

DOI No. : 10.5958/2394-9309.2022.00018.X
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Online) : 2394-9309

STOCHASTIC MODELING FOR GOSSYPIUM ARBOREUM PRODUCTION IN INDIA

T. Jai Sankar and P. Pushpa
Department of Statistics, Bharathidasan University, Tiruchirappalli, Tamilnadu, INDIA
E-mail : tjaisankar@gmail.com *and* pushbkr@gmail.com

ABSTRACT

This paper describes an empirical study of stochastic modeling for time series data of Gossypium arboreum (Cotton) productions in India. The Box Jenkins ARIMA methodology has been applied for forecasting. The diagnostic checking has shown that ARIMA (2,1,2) is appropriate. The selected model have to be considered for the forecasts from 2019 to 2023. These forecasts would be useful for the policy makers to take necessary action for the future requirements of suitable measures in this regard. Based on the chosen model, it could be predicted that G. arboreum production would increase to 31.63 million tons in 2023 from 28.71 million tons in 2018.

Key Words: ARIMA, Gossypium arboreum Production, Forecasting, ARIMA, BIC.

1. INTRODUCTION

Cotton (*Gossypium arboreum*) has performed the part of an important role in human history ever since it was first used, which anthropologists date back to prehistoric times. Cotton has been grown for fiber, food, and even fuel for more than 5,500 years. The most widely used fiber in the world today, *G. arboreum* is soft, extremely versatile, and sturdy. *G. arboreum* is an important commercial crop of India and plays a key role in the national economy. Agriculture production of *G. arboreum* is a soft, fluffy staple fiber that protective case, or grows in a boll, around the seeds of the cotton plants (Figure 1). *G. arboreum* is used to make a number of textile products. Outputs of *G. arboreum* is primarily made into yarns and threads (Figure 2) to be used in the manufacture of clothing that constitutes to 60 per cent of the cotton consumption. It is applied to make home furnishings like cushions, mattresses etc. In 2019, India is the world's largest producer of *G. arboreum*, with annual production about 18.53 million tonnes; most of this production is consumed by their respective textile industries. The five major exporters of *G. arboreum* in 2019 are India, the United States, China, Brazil, and Pakistan. Therefore the present study was undertaken to analyze the stochastic forecasting for *G. arboreum* production in India.

Figure 1. Agriculture of *G. arboreum* production

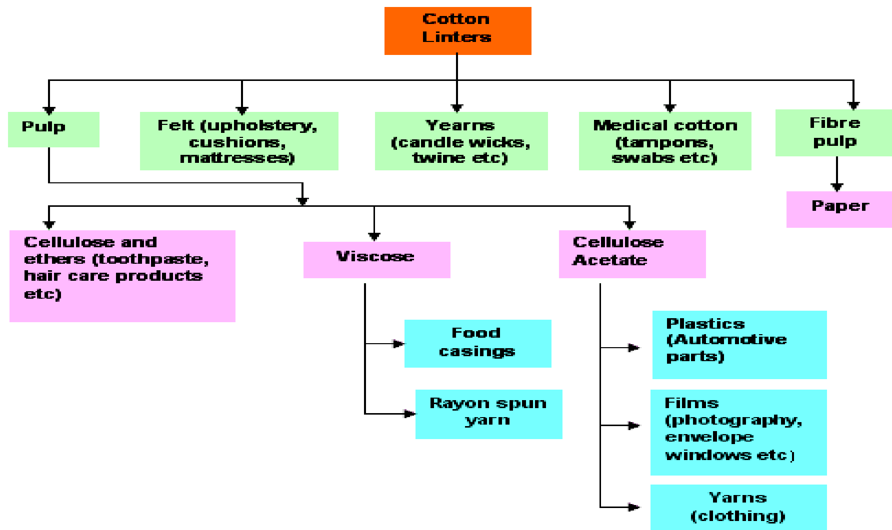


Figure 2. Outputs of *G. arboreum* production

2. MATERIAL AND METHODS

The aim of the study was to perform stochastic modelling for *G. arboreum* production in India, various forecasting techniques were considered for use. Autoregresive Integrated Moving Average (ARIMA) model, introduced by Box and Jenkins (1976), was frequently applied for discovering the pattern and predicting the future values of the time series data. Box and Pierce, (1970) considered the distribution of residual autocorrelations in ARIMA. Akaike (1970) applied the stationary time series by an AR(p), where p is finite and bounded by the same integer. Moving Average (MA) models were applied by Slutzky (1973). Applying ARIMA model Hossian et al. (2006) predicted to forecast three different varieties of pulse prices namely motor, mash and mung in Bangladesh with monthly data from 1998 to 2000; Khin et al. (2008) predicted and forecasted the natural rubber price in world market; Masuda and Goldsmith (2009) forecasted world Soybean productions; Assis et al. (2010) forecasted and pridicted cocoa bean prices in Malaysia; Wankhade et al. (2010) forecasted pigeon pea production with annual data from 1950-51 to 2007-08 in India; Shukla and Jharkharia (2011) forecasted wholesale vegetable market in Ahmedabad; The study of Debnath et al. (2013) revealed that area, production and yield of cotton in India would increase from 2016-17 to 2020-21. Borkar et al. (2016) showed that ARIMA (2,1,1) is the appropriate model for forecasting the cotton production in India. Similar studies have been conducted by Poyyamozhi et al. (2017) exposed that ARIMA (0,1,0) is the best model for forecasting cotton production in India, the analysis exposed that ARIMA (0,1,0) is the best model for forecasting cotton production. Vijaya Wali et al. (2017) revealed that, ARIMA (0,1,0) and ARIMA (1,1,1) were the best model for predicting and forecasting area and production of cotton in India.

Stochastic time-series ARIMA models were widely used in time series data which are having the properties (Alan Pankratz, 1983) of parsimonious, invertible, stationary, significant estimated coefficients and statistically independent and normally distributed residuals. When a time series is non-stationary, it can be made stationary by taking first differences of the series i.e., creating a new time series of successive differences (Y_t-Y_{t-1}). If the first differences do not convert the series form to stationary form, then first differences can be created. This is called 2nd order differencing. A distinction is made between a 2nd differences (Y_t-Y_{t-2}).

The time series when differenced follows both AR and MA models and is known as ARIMA model. Hence, ARIMA model was included in this study, which required a sufficiently large data set and considered four steps: identification, estimation, diagnostic checking and forecasting. Model parameters were estimated to fit the ARIMA models.

Autoregressive AR(p)	$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t$
Moving Average MA(q)	$Y_t = \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q \varepsilon_{t-q} + \varepsilon_t$
ARIMA (p, d, q)	$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q \varepsilon_{t-q} + \varepsilon_t$
where Y_t is *G. arboreum* production, ε_t 's are independently and normally distributed with zero mean and constant variance σ^2 for t = 1,2,..., n; d is the fraction differenced while interpreting AR and MA and ϕs and θs are coefficients to be estimated.	

***Trend Fitting* :** The Box-Ljung Q statistics was applied to transform the non-stationary data into stationarity data and also to verify the adequacy for the residuals. For estimating the adequacy of AR, MA and ARIMA processes, various reliability test statistics like R^2, Stationary R^2, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and BIC as suggested by Schwartz (1978) were applied. The reliability statistics RMSE, MAPE, BIC and Q statistics were calculated as below:

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \right]^{1/2} ; \; MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(Y_i - \hat{Y}_i)}{Y_i} \right| \; and$$

$$\text{BIC(p,q)} = \ln v^*(p,q) + (p+q) \, [\ln(n) / n]$$

where p and q are the order of AR and MA processes respectively and n is the number of observations in the time series and v* is the estimate of white noise variance σ^2.

$$Q = \frac{n(n+2) \sum_{i=1}^{k} rk^2}{(n-k)}$$

where n is the number of residuals and rk is the residuals autocorrelation at lag k.

In this study, the data on *G. arboreum* production in India were collected from the Annual Report (2018),Directorate of Economics and Statistics, Department of Agriculture, Cooperation and Farmers Welfare, Ministry of Agriculture and Farmers Welfare, Government of India for the period from 1950 to 2018 (Table 1) and were used to fit the ARIMA model to forecast and predict the future production.

RESULTS AND DISCUSSION

Model Identification : ARIMA model was designed after assessing that transforming variable under forecasting was a stationary series. The stationary series was the set of values that is varied over time around a constant mean and constant variance. The most common method to check the stationarity is to explain the data through graph and hence is done in Figure 3.

Year	Production	*Year*	Production	*Year*	Production	*Year*	Production
1950	3.04	*1968*	5.45	*1986*	6.91	*2004*	16.43
1951	3.28	*1969*	5.56	*1987*	6.38	*2005*	18.50
1952	3.34	*1970*	4.76	*1988*	8.74	*2006*	22.63
1953	4.13	*1971*	6.95	*1989*	11.42	*2007*	25.88
1954	4.45	*1972*	5.74	*1990*	9.84	*2008*	22.28
1955	4.18	*1973*	6.31	*1991*	9.71	*2009*	24.02
1956	4.92	*1974*	7.16	*1992*	11.40	*2010*	33.00
1957	4.96	*1975*	5.95	*1993*	10.74	*2011*	35.20
1958	4.88	*1976*	5.84	*1994*	11.89	*2012*	34.22
1959	3.68	*1977*	7.24	*1995*	12.86	*2013*	35.9
1960	5.60	*1978*	7.96	*1996*	14.23	*2014*	34.8
1961	4.85	*1979*	7.65	*1997*	10.85	*2015*	30.01
1962	5.54	*1980*	7.01	*1998*	12.29	*2016*	32.58
1963	5.75	*1981*	7.88	*1999*	11.53	*2017*	32.81
1964	6.01	*1982*	7.53	*2000*	9.52	*2018*	28.71
1965	4.85	*1983*	6.39	*2001*	10.00		
1966	5.27	*1984*	8.51	*2002*	8.62		
1967	5.78	*1985*	8.73	*2003*	13.73		

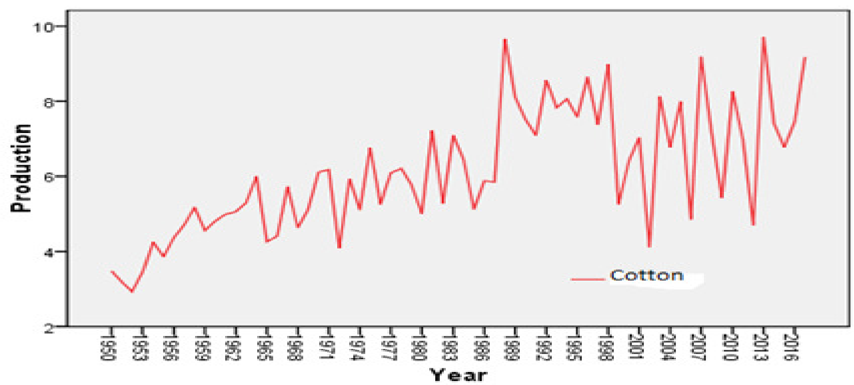


Figure 3. Time plot of *G. arboreum* production

Table 2. ACF and PACF of *G. arboreum* production

Lag	AC	Std. Error [a]	Box-Ljung Statistic	PAC	Std. Error	Lag	AC	Std. Error [a]	Box-Ljung Statistic	PAC	Std. Error
	Value	Df	Sig.[b]	Value	Df		Value	Df	Sig.[b]	Value	Df
1	0.95	0.12	64.69	0.95	0.12	*17*	0.11	0.10	351.62	0.06	0.12
2	0.88	0.12	121.51	-0.16	0.12	*18*	0.10	0.10	352.68	-0.06	0.12
3	0.82	0.12	171.99	0.07	0.12	*19*	0.09	0.10	353.44	-0.02	0.12
4	0.76	0.12	215.67	-0.12	0.12	*20*	0.07	0.10	353.93	-0.10	0.12
5	0.68	0.11	250.96	-0.20	0.12	*21*	0.05	0.10	354.23	-0.06	0.12
6	0.60	0.11	279.09	0.04	0.12	*22*	0.04	0.10	354.39	0.05	0.12
7	0.52	0.11	300.75	-0.11	0.12	*23*	0.01	0.10	354.40	-0.15	0.12
8	0.43	0.11	315.61	-0.18	0.12	*24*	-0.02	0.10	354.43	0.11	0.12
9	0.35	0.11	325.81	0.17	0.12	*25*	-0.03	0.09	354.56	0.06	0.12
10	0.31	0.11	333.54	0.15	0.12	*26*	-0.05	0.09	354.89	-0.03	0.12
11	0.26	0.11	339.41	0.02	0.12	*27*	-0.07	0.09	355.53	0.06	0.12
12	0.21	0.11	343.19	-0.11	0.12	*28*	-0.09	0.09	356.40	-0.06	0.12
13	0.17	0.11	345.75	0.06	0.12	*29*	-0.10	0.09	357.60	-0.07	0.12
14	0.15	0.11	347.72	0.03	0.12	*30*	-0.12	0.09	359.41	-0.03	0.12
15	0.13	0.10	349.28	0.02	0.12	*31*	-0.13	0.09	361.61	0.01	0.12
16	0.11	0.10	350.46	-0.06	0.12	*32*	-0.13	0.09	363.83	0.03	0.12

[a]The underlying process assumed is independence (white noise).
[b]Based upon the asymptotic chi-square approximation.

Figure 3 reveals that the data used were non-stationary. Again, non-stationarity in mean was corrected through first differencing of the data. The newly constructed variable Y_t could now be examined for stationarity. Since, Y_t was stationary in mean, the next step was to identify the values of p and q. For this, the ACF and PACF of various orders of Y_t were computed and presented in Table 2 and Figure 4.

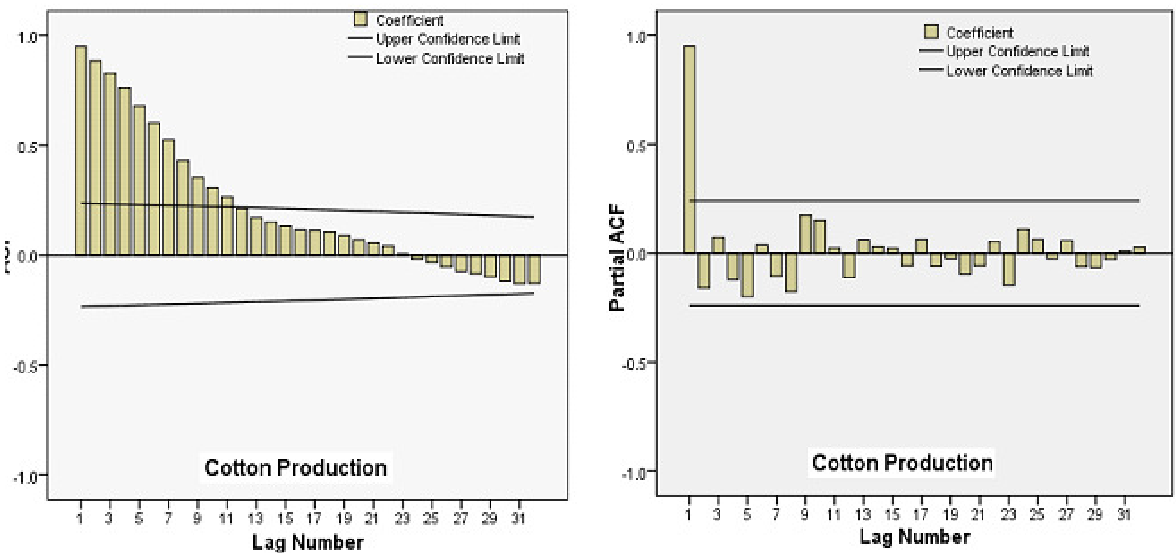


Figure 4. ACF and PACF of differenced data

The various ARIMA models are expressed with values differenced once (d=1) and the model which had the minimum normalized BIC was taken. The various ARIMA models and the corresponding normalized BIC values are given in Table 3. The value of normalized BIC of the chosen ARIMA was 1.618.

Table 3. BIC values of various ARIMA (p, d, q)

ARIMA (p,d,q)	0,1,0	0,1,1	0,1,2	1,1,0	1,1,1	1,1,2	2,1,0	2,1,1	**2,1,2**	3,1,0	3,1,1	3,1,2
BIC Values	1.673	1.650	1.657	1.659	1.602	1.654	1.661	1.653	**1.618**	1.642	1.7321	1.658

***Model Estimation*:** Model parameters and fit statistics were estimated and the results of estimation are presented in Tables 4 and 5. Hence, the most suitable model for *G. arboreum* production was ARIMA (2,1,2), as this model had the lowest normalized BIC value, good R^2 and better model fit statistics (RMSE and MAPE).

Table 4. Estimated ARIMA model

	Estimate	SE	t	Sig.
Constant	-20.175	24.445	-0.825	0.412
AR1	-0.938	0.116	-8.075	0.000
AR2	-0.947	0.115	-8.251	0.000
MA1	-1.141	0.958	-1.191	0.238
MA2	-0.995	1.663	-0.598	0.552

Table 5. Estimated ARIMA model fit statistics

Performances of different ARIMA (p,d,q) models of *G. arboreum* production in India								
ARIMA (p,d,q)	Stationary R^2	R^2	RMSE	MAPE	MaxAPE	MAE	MaxAE	Normalized BIC
0,1,0	0.01	0.954	2.063	12.843	35.821	1.416	8.339	1.673
0,1,1	0.01	0.954	2.079	12.874	35.939	1.417	8.322	1.65
0,1,2	0.084	0.958	2.016	13.535	37.522	1.437	6.959	1.65
1,1,0	0.01	0.954	2.079	12.86	35.887	1.417	8.33	1.65
1,1,1	0.127	0.96	1.968	13.632	44.392	1.427	6.981	1.602
1,1,2	0.152	0.961	1.954	13.525	46.976	1.414	6.795	1.65
2,1,0	0.074	0.957	2.027	12.862	35.209	1.413	7.174	1.661
2,1,1	0.15	0.961	1.956	13.572	47.243	1.415	6.833	1.653
2,1,2	***0.241***	***0.965***	***1.864***	***12.69***	***41.197***	***1.283***	***5.63***	***1.618***
3,1,0	0.159	0.961	1.946	13	41.515	1.366	6.368	1.642
3,1,1	0.159	0.961	1.962	13.028	41.234	1.367	6.361	1.72
3,1,2	0.207	0.966	1.843	12.115	36.735	1.311	6.793	1.658

***Diagnostic Checking*:** The model verification is concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be extracted to improve the chosen ARIMA, which has been completed through examining the autocorrelations and partial autocorrelations of the residuals of various orders. For this purpose, various autocorrelations up to 32 lags were computed and the same along with their significance tested by Box-Ljung statistic are provided in Table 6. As the results indicate, no one of these autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARIMA model was an appropriate model for forecasting *G. arboreum* production in India.

Table 6. Residual of ACF and PACF

Lag	ACF		PACF		*Lag*	ACF		PACF	
	Mean	**SE**	**Mean**	**SE**		**Mean**	**SE**	**Mean**	**SE**
1	-0.053	0.121	-0.053	0.121	***17***	0.045	0.142	0.029	0.121
2	-0.015	0.122	-0.018	0.121	***18***	0.144	0.142	-0.003	0.121
3	0.129	0.122	0.128	0.121	***19***	-0.057	0.144	-0.002	0.121
4	-0.033	0.124	-0.020	0.121	***20***	-0.124	0.145	-0.221	0.121
5	-0.047	0.124	-0.047	0.121	***21***	0.086	0.146	-0.043	0.121
6	-0.039	0.124	-0.062	0.121	***22***	0.144	0.147	0.032	0.121
7	0.141	0.124	0.146	0.121	***23***	-0.117	0.149	-0.068	0.121
8	-0.196	0.127	-0.180	0.121	***24***	-0.110	0.150	-0.216	0.121
9	-0.146	0.131	-0.159	0.121	***25***	0.101	0.152	0.052	0.121
10	-0.133	0.133	-0.208	0.121	***26***	-0.123	0.153	-0.039	0.121
11	0.040	0.135	0.087	0.121	***27***	0.042	0.154	0.222	0.121
12	-0.212	0.135	-0.203	0.121	***28***	0.015	0.154	-0.082	0.121
13	-0.110	0.140	-0.124	0.121	***29***	-0.034	0.154	-0.055	0.121
14	0.040	0.141	-0.085	0.121	***30***	-0.011	0.154	-0.016	0.121
15	0.044	0.142	0.160	0.121	***31***	-0.116	0.154	-0.043	0.121
16	0.050	0.142	0.063	0.121	***32***	0.152	0.156	-0.057	0.121

The ACF and PACF of the residuals are given in Figure 5, which also indicated the 'good fit' of the model. Hence, the fitted ARIMA model for *G. arboreum* production data was

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .... - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

$$Y_t = -20.175 - 0.938Y_{t-1} - 0.947Y_{t-2} + 1.141\varepsilon_{t-1} + 0.995\varepsilon_{t-2} + \varepsilon_t$$

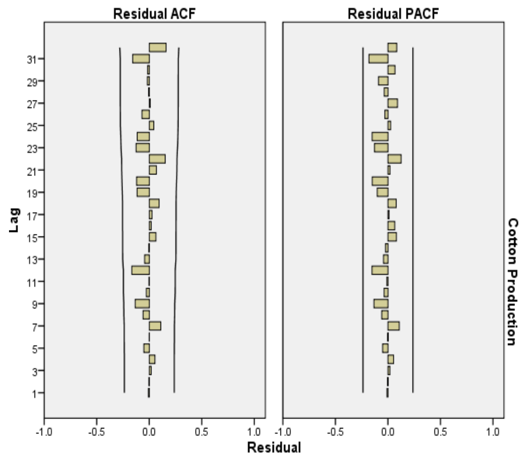


Figure 5. Residuals of ACF and PACF

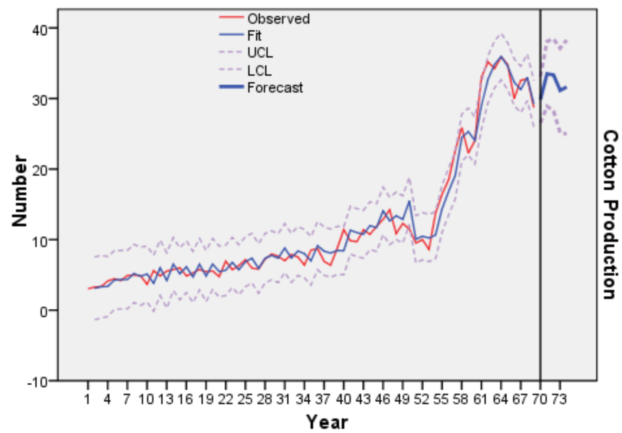


Figure 6. Actual and estimate of production

Forecasting: Based on the model fitted, forecasted *G. arboreum* production (in million tons) for the year 2019 through 2023 respectively given by 29.83, 33.50, 33.33, 31.15 and 31.63 are given in Table 7. To evaluate the forecasting ability of the fitted ARIMA model, the counts of the sample period forecasts' accuracy were also calculated. This measure signified that the forecasting inaccuracy was low. Figure 6 shows the actual and forecasted value of *G. arboreum* production (with 95% confidence limit) in the country. The upper control limit (UCL) and lower control limit (LCL) values of the forecasted *G. arboreum* production in India is provided in the same Table 7.

Table 7. Forecast of *G. arboreum* production

Year	Predicted	UCL	LCL
2019	29.83	33.13	26.53
2020	33.50	38.10	28.90
2021	33.33	38.45	28.22
2022	31.15	37.12	25.18
2023	31.63	38.28	24.97

CONCLUSION

The most appropriate ARIMA model for *G. arboreum* production forecasting of data was found to be ARIMA (2,1,2). From the temporal data, it can be found that forecasted production would increase to 31.63 million tons in 2023 from 28.71 million tons in 2018 in India for using time series data from 1950 to 2018 on *G. arboreum* production, this study contributes an evidence on future *G. arboreum* production in the country, which can be considered for future policy making and formulating strategies for augmenting and sustaining *G. arboreum* production in India.

REFERENCES

1. *Akaike, H. (1970). Statistical Predictor Identification. Annals of Institute of Statistical Mathematics.**22**: 203-270.*
2. *Alan Pankratz, (1983). Forecasting with Univariate Box - Jenkins Models: Concepts and Cases. John Wiley & Sons, New York.*
3. *Assis, K., Amran, A., Remali, Y. and Affendy, H. (2010). A comparison of univariate time series methods for forecasting cocoa bean prices. Trends Agric. Econ., **3**: 207–215.*
4. *Borkar, Prema and Tayade, P.M. (2016). Forecasting of Cotton Production in India Using ARIMA Model. International Journal of Research in Economics and Social Sciences, **6(5)**: 1-7.*
5. *Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. San Francisco, Holden-Day, California, USA.*
6. *Box, G.E.P. and Pierce, D.A. (1970). Distribution of Residual Autocorrelations in ARIMA Models. J. American Stat. Assoc., **65**: 1509-1526.*
7. *Debnath, M.K., Kartic Bera and Mishra, P. (2013). Forecasting Area, Production and Yield of Cotton in India Using ARIMA Model, Research and Reviews: Journal of Space & Technology, **2(1)**: 16-20.*
8. *Hossain, M.Z., Samad, Q.A. and Ali, M.Z. (2006). ARIMA model and forecasting with three types of pulse prices in Bangladesh: A case study. Int. J. Social Econo., **33**: 344–353*
9. *Khin, A.A., Eddie, C.F.C, Shamsundin, M.N. and Mohamed, Z.A (2008). Natural Rubber Price Forecasting in the World Market, Agricultural Sustainability Through Participate Global Extension, University Putra Malaysia, Kuala Lumpur, Malaysia.*
10. *Masuda, T. and Goldsmith, P.D. (2009). World soybean production: area harvested, yield, and long-term projections, Int. Food Agribus. Manag. Rev., **12**: 143–162.*
11. *Poyyamozhil, S. and Kachi Mohideen, A. (2017). Forecasting of Cotton Production in India Using ARIMA Model. Asia Pacific Journal of Research, **XLVIII (I)**: 70-74.*
12. *Shukla, M. and Jharkharia, S. (2011). Applicability of ARIMA models in wholesale vegetable market: An investigation. Proceedings of the International Conference on Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia.*
13. *Slutzky, E. (1973). The summation of random causes as the source of cyclic processes. Econometrica, **5**:105-146.*
14. *Vijaya B. Wali, Devendra Beeraladinni and Lokesh, H. (2017). Forecasting of Area and Production of Cotton in India: An Application of ARIMA Model. Int. J. Pure App. Biosci. **5(5)**: 341-347.*
15. *Wankhade, R., Mahalle, S., Gajbhiye, S. and Bodade, V.M. (2010). Use of the ARIMA model for forecasting pigeon pea production in India. Int. Rev. Bus. Finance, **2**:97–102.*

DOI No. : 10.5958/2394-9309.2022.00019.1
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

FACTORS AFFECTING GREEN MARKETING TOWARDS FMCG PRODUCTS IN TAMBARAM CITY

Mrs. N. Selvi -Ph. D. Research Scholar*, Dr. N. Premila - Associate Professor
*Department of Commerce, PG Research Department, Queens Marry's College, Chennai -600 004

ABSTRACT

Now a days, Whole World is recognizing the prerequisite of the Green Marketing, Environmental Marketing and Biological marketing which gives the same meaning to the research area. Although environmental matters influence on all most of all activities of our life, there were only few academic disciplines have discussed green issues in Indian market. Especially from Tamilnadu business discipline, FMCG sector is a noticeably large sector in the economy which has to open their eyes on eco friendliness.

The findings of the study show that Green Brand Image, Green Satisfaction, Green Loyalty of the customers affects their Green Brand Equity. The empathetic on the changing feasting pattern of Tambaram City consumers and the value of accessory, they gave to the environment delivers useful understandings that are particularly pertinent to an improved considerate of factors affecting Green Purchasing Behaviour in the fast-moving consumer goods in the Tambaram city.

(Keywords: Green products , factors analysis, Chi square Test mean standard deviation, Tambaram city, FMCG , Consumers etc.,)

1. INTRODUCTION

In the modern epoch of globalisation, it has become a prodigious challenge to keep the customers as well as consumers in fold and even keep our natural atmosphere safe. The growing social concern for the environment has emerged as a key perception in marketing- i.e, Green Marketing. It incorporates a broad range of activities, including product variation, changes to the production process, packaging changes, as well as modifying promotion. The Fast-Moving Consumer Goods (FMCG) sector is one of the growing industries and they have increasingly recognised the various competitive recompences and opportunities to be gained from eco-sustainability and green marketing. However, a better understanding of consumer behaviour is necessary especially in the FMCG sector.

The mounting social and governing concerns for the environment lead an increasing number of companies to contemplate green issues as a major source of planned change. In specific, this trend has major and multifaceted implications on the technological approach of a company and on its product innovations. Even though it is increased eco-awareness of Tambaram customers during the past few eras, there are some fences to the dispersals of more ecologically concerned with consumption and production styles.

Therefore, companies are increasingly recognizing the importance of green marketing concepts and concentrates for the Green products.

The green marketing concept has been newly enormously studied due to its impact on day-to-day procurement decisions. This concept can also be perceived as: environmental marketing or ecological marketing (Henion and Kinnear 1976).

"Green products" or "environmental products" are often accompanying for having certain appearances which were developed with green process, or in other words with as less impact on the environment as possible. Protecting the environment, reducing energy and resources describe some of their topographies. The processes associated tend to eliminate the use of toxic products, pollution and waste which are threat to the environment. The green products have to increase the productivity and use of expected resources, must involve biological production model and also reduce the quantity of materials used in its processes - dematerialization (Singh and Pandey 2012).

2. OBJECTIVES OF THE STUDY

- To investigate consumer preference towards green products in the FMCG sector
- To analyses the factors affecting green purchasing behaviour.

3. HYPOTHESIS OF THE STUDY

- There is no association between location and awareness of the green products for the FMCG products.

4. LIMITATIONS OF THE STUDY

In this study, the researcher faced more difficulties while collecting data from the customers for green products users. Duration of data collection was Very short duration. The respondents are not give the accurate data for fear.

5. LITERATURE REVIEW

Green marketing doesn't only refer to the promotion or advertising of products with environmental characteristics. We tend to associate this concept to terms such as: recycling, environmentally friendly, or for example refillable. However, regarding Henion and Kinnear (1976) this concept is wider, as it also involves product modification, changes in the production process, brand adaptation (e.g. logo, packaging), and alteration and improvement of the advertising approaches. It also implies the consequences marketing activities have on pollution, environment degradation and on energy consumption. It provides a different perspective of the general concept of marketing, as it is no longer focused on particular societies concerns, but in global ones. It consists of being able to satisfy customers' needs having the minimum impact and harm in the environment.

In marketing has become increasingly significant to the modern market. Companies have to re-think about all the activities which involve their products, whether it is the process or advertising for example, in order to reach environmentally conscious consumers, those consumers who are focused on their actions and the impact they have in the world, setting apart their materialistic side. And at the same time expand the mind set of consumers who are still not aware of their actions impact and the meaning of green-friendliness products. The current tendency is for consumers paying more attention to the companies' practices as well as the product characteristics whether they are sustainable or not. Marketing practices such as adapting their brand image with "visual images most associated with the environment" or products made with recycled materials. However, implementing green marketing involves the analysis of several conditions as consumer awareness, costs and profit issues, awareness of the topic and competitive pressures (Singh and Pandey 2012).

In order to have a successful green marketing strategy, the organization must be genuine in terms of it stands for, or in other words, whether companies are accurate, comply and act according the companies' policies. Organizations must also act accordingly to what they claim during their marketing campaigns. Empower customers is also an important key to achieve success as it gives customers the power to intervene and make them part of the environmental actions and assure customers recognize the coming benefits from being green. The acknowledgment of customers' preferences and characteristics is an essential asset allows the anticipation of their need. In this particular case in order to sell green products, organizations must know whether customers are aware and conscious of their actions and the consequences they might have in the environment. Focus on transparency as long as companies have been acting according to their procedures and their claims regarding their position as environmentally friendly and so, avoid all the skepticism coming from adopting a greener strategy. And finally, set prices according to the defined target (Singh and Pandey 2012).

One of the main issues of green marketing is satisfying customers' needs, providing them with alternatives which do not contaminate or are considered harmful to the environment due to the scarcity of resources which relies on the basic definition of Economics - the study of how people use their limited resources to try to satisfy unlimited needs (McTaggart, Findlay and Parkin 1992). So, it is important to find different alternatives with limited resources to satisfy these unlimited needs of both individuals and industry, and at the same time correspond with the company's goals.

Regarding Ottman (2011), nowadays, companies focus on product development and ways to align it with sustainability. Actions in producing, distributing, usage and recycle processes must have the minimal impact on the environment. It's important to consider the resources involved in each process, which resources are required during the product lifetime and whether the practices used are ethical. The companies' role in society has more duties and consequently is becoming more important with their active participation (Keller 1987, Shearer 1990). They believe they have moral obligation (Davis 1992) in their policies and practices to have a positive impact on the environment (Azzone, Giovanni and Manzini 1994). So companies using green marketing approaches are gaining competitive advantage over companies which don't engage in responsible practices. This current trend emerges in order to better satisfy customers' needs (Polonsky 1994).

On the other hand, the increasing pollution and the damages from global warming are some of the examples of human destruction which are devastating the environment. So, we have been observing how social responsibility has been assuming an important part in today's society. Therefore, not only marketeers but also consumers are taking actions and being more concerned about it, changing their behavior patterns.

This concept explores customers' mind set and how they stand about the environmental issues, becoming more of a "central core value" rather than being too explicit and so that influence consumption and marketing decisions.

6. RESEARCH METHODOLOGY

The research also explored the impact of green marketing practices of FMCG in Tambaram city. Qualitative approach was adopted for the study by using a interview schedule and the total sample composed of 320 respondents from Tambaram City. Stratified random sampling was used to collect data. The researcher has collected the primary and secondary data. Primary sources of data are reflective of the answers of participants used from the data collection process. Responses gathered through the use of interview schedule are able to address the objectives of the study and the questions that were raised from the onset of this research. As such, they are considered as the main data to be processed and evaluated. Some sources of data, those that are considered as secondary, were also gathered from the literature that was reviewed within this study. The literature presents significant information with regards to the objectives of the study and is also particularly helpful in the interpretation of findings and results. The researcher also applying chi square analysis for this research.

7. DATA ANALYSIS AND DISCUSSION

The researcher analyse data for various segments Demographic profile of the respondents like age, Sex, marital status, qualification, income level, place,awareness about green products and factors affecting the green products for the FMCG and recommendation about the green products for FMCG etc.

Table -1- Demographic Profile of Green Product Consumers

Sl no	*Demographic variable*	*Classification*	*Location*		*No of Green Consumers*	*Percentage*
			Rural	*Urban*		
1	*Gender*	*Male*	*180*	*65*	*245*	*76.56*
		Female	*46*	*29*	*75*	*23.44*
2	*Age (in years)*	*Below 35*	*130*	*30*	*160*	*50.00*
		35 – 45	*45*	*30*	*75*	*23.44*
		45 – 55	*25*	*20*	*45*	*14.06*
		Above 55	*26*	*14*	*40*	*12.50*
3	*Marital status*	*Married*	*196*	*74*	*270*	*84.38*
		Unmarried	*30*	*20*	*50*	*15.62*
4	*Educational qualification*	*Below Hr.Sec*	*50*	*10*	*60*	*18.75*
		UG	*81*	*44*	*125*	*39.06*
		PG	*70*	*20*	*90*	*28.13*
		Diploma and Professional	*25*	*20*	*45*	*14.06*
5	*Status*	*Business*	*25*	*5*	*30*	*9.38*
		Agriculture	*10*	*30*	*40*	*12.50*
		Govt Employee	*26*	*9*	*35*	*10.94*
		Private Employee	*120*	*35*	*155*	*48.44*
		Student	*45*	*15*	*60*	*18.75*
6	*Income (Rs. In Months)*	*Below Rs.25000*	*126*	*64*	*190*	*59.38*
		Rs.25000-Rs.35000	*28*	*12*	*40*	*12.50*
		Rs. 35000- Rs 45000	*40*	*10*	*50*	*15.63*
		Above Rs 45000	*32*	*8*	*40*	*12.50*

Source: Collected Data

From the above analysis, the researcher analysed the demographic status of the green products user among the FMCG products in Tambaram City. Out of 320 respondents most 76.56 per cent of the customers are male in the rural place; followed by 50 per cent of the green product users were fall under the age below 25 years; majority of the customers are married under this study because the students are very lesser amount of numbers; 39.06 per cent of the respondents are educated in the Under Graduate level; out of 320 respondents 48.44 per cent of the customers status are private employees in this study area and the income level of the green product users are earned below Rs 25000 per month. In this study the researcher further analysed that rural and urban consumer in the green product in FMCG. In this study most of the respondents are rural people in Tambaram City.

Table 2- Awareness about Green Products for the FMCG

Sl No	*Awareness*	*Rural*	*Urban*	*Total Respondents (Percentage)*
1	*Yes*	*200 (88.50%)*	*60 (63.83%)*	*260 (81.25)*
2	*No*	*26 (11.50%)*	*34 (36.17)*	*60 (18.75)*
Total		***226 (100%)***	***94 (100%)***	***320 (100)***

Source: Collected Data '()' indicates Percentage

Table 2 clearly exhibits that the awareness about green product users in FMCG goods, out of 320 respondents, 81.25 per cent of the interviewee are more awareness about the green products in FMCG products, out of 260 respondents 200 respondents (88.50 per cent) of the sample force are rural customers and the remaining 11. 50 per cent of customers are not aware about green products. 18.75 customers are not aware about rural and urban.

Table 3 - Factors affecting Green Products in FMCG

Sl No	*Factors*	*Rural*	*Urban*	*Total Respondents (Percentage)*
1	*Apprehension for health and atmosphere*	*70*	*20*	*90 (28.13)*
2	*Eco buying insolence*	*14*	*8*	*22 (6.88)*
3	*Eco certification*	*22*	*6*	*28 (8.75)*
4	*Social consciousness and value*	*30*	*12*	*42 (13.13)*
5	*Nonappearance of marketing*	*16*	*4*	*20 (6.25)*
6	*Advancement of Green Products*	*18*	*6*	*24 (7.50)*
7	*Unresponsive defiance*	*18*	*12*	*30 (9.38)*
8	*Life style*	*16*	*9*	*25 (7.81)*
9	*Brand awareness*	*17*	*15*	*32 (10.00)*
10	*Packaging*	*5*	*2*	*7 (2.19)*
Total		***226***	***94***	***320 (100)***

Source: Collected Data; '( )' indicates Percentage

Table 3 clearly indicates that various factors affecting the Green products for FMCG; out of 320 respondents 28.13 per centage of the consumers are affected that apprehension for health and atmosphere followed by 13.13 per cent of the consumers affected by the social consciousness and value; 10 per cent of the consumers are affected that brand awareness for the green products in the FMCG; 9 per cent of the sample force affected that unresponsive definance and eco certification respectively; very least number of sample force are affected by the packaging of the green products in FMCG goods.

The importance of each factor as per respondents rating the researcher calculated the mean values and standard deviation of the factors analysis.

Table 4 –Importance for Factors affecting green products for FMCG

Sl No	Factors	Mean	Standard Deviation
1	Apprehension for health and atmosphere	3.88	0.82
2	Eco buying insolence	3.21	0.90
3	Eco certification	3.27	0.99
4	Social consciousness and value	3.61	1.19
5	Nonappearance of marketing	2.99	0.86
6	Advancement of Green Products	3.23	1.02
7	Unresponsivedefiance	3.29	1.03
8	Life style	3.25	1.04
9	Brand awareness	3.41	0.89
10	Packaging	2.53	1.24

Source: Computed Data

Result for table 4 mean and standard deviation in the factor analysis for loaded variables apprehension for health and atmosphere first followed by brand awareness is second; unresponsive defiance is third factor followed by eco certification is fourth factor and the very least factor is packaging.

8. VALIDATING FOR HYPOTHESIS

To validate the research the researcher hasanalyse the null hypothesis that "There is no association between location and awareness of the green products for the FMCG products". In this hypothesis the researcher applies the chi square test for this research.

Calculated Chi Square Value	P value	Degrees of freedom	Significant level	Result
24.9988	3.84	5 % level	Not significant	Rejected

Source: Computed

In this research 1 d.f at 5 % level of significance the table value is lesser than the calculated value so the hypothesis is rejected. Hence concluded that there is an association between location of the green products users and awareness of the green marketing in FMCG products.

9. FINDINGS OF THE STUDY

Out of 320 respondents most 76.56 per cent of the customers are male in the rural place; followed by 50 per cent of the green product users were fall under the age below 25 years; majority of the customers are married under this study because the students are very lesser amount of numbers; 39.06 per cent of the respondents are educated in the Under Graduate level; out of 320 respondents 48.44 per cent of the customers status are private employees in this study area and the income level of the green product users are earned below Rs 25000 per month. In this study the researcher further analysed that rural and urban consumer in the green product in FMCG. In this study most of the respondents are rural people in Tambaram City. Most (81.25 %) of the respondents awareness about the green marketing products in FMCG products and 28.13 % of the respondents affected for the factors and loaded factors mean and standard deviation first factor is apprehension for health and atmosphere and also hypothesis testing the null hypothesis rejected and alternate hypothesis that there is an association between location and awareness about the green marketing of FMCG.

10. RECOMMENDATION AND CONCLUSION

Though consumers are enthusiastic to purchase green products, many business establishments still in behind the need of the eco-friendly society. The authors suggest business organizations to follow various strategies in order to get benefits from the environmentally friendly tactic as green marketing offers business inducements and growth occasions while it may involve start-up costs, it will save money in the extended term. Therefore, in the product strategy, marketers can identify customers' environmental needs and advance products to address this issue, produce more environmentally responsible packages. (Recycle, biodegradable, reuse), and ensure that products meet or exceed the superiority expectations of customers.

The researcher concludes that consumer is waking up to the virtues of green products. But it is still a new concept for the majority. The new green activities need to reach the masses and that will take a lot of time and effort. The government, the organization, the masses and the consumers have to join their hands together in fetching the ecological balance.

Green marketing is a tool used by many companies to upsurge their inexpensive advantage as people nowadays concerned about their health as well as on environment. It is not only an environment guard tool but also a marketing strategy. companies can create more awareness to customers regarding green marketing and increase their sales as well as protect the environment.

REFERENCE:

1. *Anderson, W.T. Jr, Henion, K.E. II, and Cox, E.P. (1974), "Socially vs. ecologically concerned consumers", American Marketing Association Combined Conference Proceedings, Vol. 36 (Spring and Fall), pp. 304-11.*
2. *Antil, J.A. (1984), "Socially responsible consumers: profile and implications for public policy", Journal of Macro marketing, Fall, pp. 18-39.*
3. *Clare D'Souza, Mehdi Taghian, Peter Lamb, Roman Peretiatko (2004),"Green decisions: demographics and consumer understandingof environmental labels",International Journal of Consumer Studies.*

4. *Coddington, W. (1993), Environmental Marketing: Positive Strategies for Reaching the Green Consumer, McGraw-Hill, New York, NY.*
5. *Kinnear, T.C., Taylor, J.R. and Ahmed, S.A. (1974), "Ecologically concerned consumers: who are they?", Journal of Marketing, Vol. 38, April, pp. 20-24.*
6. *Ottman, J.A. (1993), Green Marketing: Challenges and Opportunities for the New Marketing Age, NTC Business Books, Lincolnwood, IL.*
7. *McIntosh, A. (1991), ``The impact of environmental issues on marketing and politics in the 1990s'', Journal of the Market Research Society, Vol. 33 No. 3, pp. 205-17.*
8. *Aditya Maheshwari andDr. Gunjan Malhotra (2011), "Green marketing: A study on Indian youth", International journal of management and strategy, Vol.2, No.3.*
9. *Ajay Kumar Singh and Monkia Bansal (2012), "Green marketing: A study of consumer attitude and environmental concern", Indian journal of Commerce, Vol.65, No.2.*
10. *Bahram Kheiry and Arezoo Nakhaei (2012), "Consumers' green purchase decision: An examination of environmental beliefs, environmental literacy and demographics", International Journal of Marketing and Technology, Vol.2, Issue 9*
11. *Dipti Shankar Barge and Dinkar Khashaba More (2014), "Awareness of customers about eco-friendly products- A study of Satara district", Golden Research Thoughts, Vol.4, Issue 4*
12. *Ishaswini and Saroj Kumar Data (2011), "Pro-environmental concern influencing green buying: a study on Indian consumers", International Journal of Business and Management, Vol.6, No.6.*

DOI No. : 10.5958/2394-9309.2022.00020.8
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

CLASSIFICATION, PREDICTION OF COIVD-19 TOTAL VACCINATION VERSUS GENDER IN INDIAN STATES AND UNION TERRITORIES USING MACHINE LEARNING ALGORITHMS AND MULTIPLE REGRESSION ANALYSIS

K. Gokila*, R. Lakshmi Priya, Manimannan G*****
*, ** Assistant Professor, Department of Statistics, Dr. Ambedkar Govt. Arts College, Chennai
***Assistant Professor, Department of Statistics, Apollo Arts and Science College, Chennai

ABSTRACT
Aim: In this research paper, the main objective of this study is topredict and identify the differences between Total vaccinated (with different vaccination) people vs. gender (Male, Female and Transgender) using Machine Learning (ML) method of Multiple Regression Analysis (MRA).
Setting and Designs: The independent variable X (Total vaccinated population) and Y (Gender) are dependent variables. Initially the data set issplit into two categories;namely tested and training data of total vaccinated population of entire states and UTs of the country. The python script writes and imports the entire database and split database into train and test category in the ratio of 30:70. The 30 percent of data is using for training and remaining 70 percentage of data is used for testing. The predicted results of MRA with original databaseare identified with the differences between actual and predicted data of entire datasets.
Methods and Materials:The secondary sources of data were collected from Ministry of Health and Family Welfare Department, India. Four main parameters were used in this research paper.They are Total vaccinated population, Male, Female and Transgender population.
Machine Learning Methods:The statistical function displays the result of descriptive statistics. In this research, the vaccination of two doses with total population and gender were fitted very effectively and R-square value indicates 0.999999. The scatter diagram shows the linear relationship of results of Actual and predicted value of vaccinated population.
Results and Discussion: MRA algorithm is very accurate and makes good predictions.Finally the data visualized MRA in the form of graphical representation, this result shows that the actual and predicted MRA results achieved hundred percent between total population and predicted population. In visualization also it is clearly displayed through MRA results line and simple bar diagrams.
Conclusion: The more vaccinated states are Uttar Pradesh followed by Rajasthan, Karnataka, West Bengal, Tamilnadu, Kerala, Telengana and Odisha, etc. The more vaccinated Union Territories are Goa followed by Puduchery, Andaman Nicobar Lakshadweep, Ladakh and Dadra and Nagar Haveli and Daman and Diu, etc.
Keywords: Machine Learning (ML), Multiple Regression Analysis (MRA), Training and Testing, Prediction, COVID-19, Vaccinated Population and Visualization.

1. INTRODUCTION

The ability of this family of viruses readily undergoes genetic recombination not only within groups, but also between groups; it makes them readily vulnerable to natural selection and changing its environment of virulence [1]. The most striking feature however, is itstravel to freely cross from one species to another species. The group of Human Corona Virus 229 E(HCoV 229 E) belongs to the group 1 family thought to be responsible for the epidemic of common cold [7]. The virus transmission from bats to humans is thought to be the initial transmission process for HCoV 229 E, which had happened within the last two eras. However, the two dramatic events-Severe Acute Respiratory Syndrome Corona Virus (SARS-CoV), originated from bats and got transmitted from civet cats and Middle East Respiratory Syndrome Corona Virus (MERS-CoV), originated from bats and got transmitted from camels in 2003 and 2012 respectively[2].

India is the second largest populated country in the world. The COVID-19 vaccination is administered around 64,05,28,644people in the month of August 31, 2021. Nearing half the population is vaccinated with first and second dose. The different vaccinations are COVISHIELD, COVAXIN, SPUTNIK V and Adverse Events Following Immunization (AEFI). The COVISHIELD vaccine is produced by Serum Institute of India private limited. This vaccine is administered to the age group 18 to upper age. Most of the Indian population has been vaccinated with COVISHIELD and COVAXIN is the second largest vaccine administered to Indian population. This vaccine is produced by Bharat Biotech, Hyderabad, India. Latter stage SPUTNIK V and AEFI was imported from Russia. Initially, the Indian vaccinescosts varied and in latter stage Indian government supplied for free of cost to entire population.

2. BACKGROUND OF THE STUDY

This section discuss about COVID-19 that has taken the world to surprise that almost all walks of life has been impacted due to it. Different medical specialties have been diverted towards caring for COVID-19 patients. One such situation arises for radiologists, who currently undergo down time in their medical professional. Fessell et al. [3] discusses about how to make the down time useful and meaningful to our lives. They discuss about the importance of self-care, like applying emergency masks in airline flight before applying to your dependents, and how stress is created due to current news outlets. Re- thinking about those books that are bought before but not read and those parts of your home that are not well maintained are certain thought proving entities as of now [3].

Reduction in acute coronary syndrome in current times[4] is attributed towards reduction in air pollution. The literature towards counseling general public to be wary of air pollution did not gain popularity but the current prevalent usage of N95 masks to prevent the spread of COVID-19 has addressed multiple issues [4]. Prior research has shown impact on usage of masks to protect humans from air pollutants like PM2.5 [5], indicating the size and spread of such pollutants. Comparison of the presence of air pollutants with SARS-CoV-2 in nasal epithelium and upper respiratory tract [6] suggests masks could be one factor to control the spread of this deadly virus.

In the recent years, Data scientists all over the world are demanding in making sense out of the available data and predict the near future. To identify trend pattern, parameter selection, forecasting techniques are being applied in and out to come to a conclusion [7]. Gupta, Pal and Kumar [10] in their research paper 'Trend analysis and forecasting of COVID-19 break in India' used exploratory data analysis to report the situation in the time period of January to March in India. In this paper, they used time series forecasting models to predict future trends. A well-known machine learning model-ARIMA model prediction was used and inferred that a huge surge in the number of likely COVID -19 positive cases was predicted in

April and May. The regular that was forecasted was a detection of approximately 7000 patients in a total span of 30 days in April. However in reality the figures were higher [8]. Recent research paper termed SEIR and Regression Model based COVID-19 outbreak predictions in India' by Pandey and Chowdhury [9] from department of CSE and IT of North cap University, India in association with Defence Research and Development Organization (DRDO),Indiacovered data from January 30 to March 30, 2020.
They used regression models for forecasting. Another very relevant paper named "Analysis of Spatial Spread Relationships of Coronavirus (COVID-19) Pandemic in the World using Self Organizing Maps (2020)" by Julio, Monica, Sanchez, Castillo in the paper of 'Chaos, Solitonsand Fractals' uses clustering methods to analyze countries on the basis of most affected patients and how they are reacting to it.
Recent research paper described the results of various machine learning classification models and cross validation accuracy and are 88%, 97%, 91% and 91% respectively. The State Classificationand Union Territories was named as Very Low Affected (VLA), Low Affected (LA), Moderately Affected (MA), Highly Affected (HA) and Very Highly Affected (VHA) States and Union Territories of India by COVID-19 cases. Maharashtra is correctly classified as Very High Affected States, Delhi, Uttar Pradesh and West Bengal falls in Moderately Affected States, Assam, Bihar, Chattisgarh, Haryana, Gujarat, Madhya Pradesh, Odisha, Punjab, Rajasthan and Telangana falls in Low Affected States and Tamilnadu, Kerala Andhra Pradesh and Karnataka forms a group of highly affected States. Remaining States and Union Territories falls in Very Low affected by Covid-19 Cases carried out by Manimannan G. et. al. [11].This analysis was aimed at predicting a trend related to vaccination of first and second dose of COVID-19 in India.

3. MATERIAL AND METHODS

The secondary sources of data were collected from Ministry of Health and Family Welfare Department, India [12] during the period of January 16, 2021 to 06, July, 2021. The states and union territories of vaccinated database consist of four parameters, total vaccinated population, male population, female population and transgender population. The database is split in two categories in the ratio of 30: 70. Thirty percent data is training and seventy percent data is testing in ML method of Multiple Regression Analysis (MRA).

4. MACHINE LEARNING METHOD

In general, a regression equation can have one dependent variable and many independent variables. The Multiple Regression equation in most general form can be written as follows;

$$Y = \beta_1 X_1^a + \beta_1 X_2^b + \cdots + \lambda_1 X_1^c X_2^d + \lambda_2 X_2^e X_2^f + \cdots$$

Where Y is dependent variable (Total vaccinated population);$X_1, X_2, \ldots$ are independent variables;$\beta_1, \beta_2, \ldots$ are regression coefficients and $a, b, \ldots$ are exponents of the model. A regression equation cannot have more than one independent variable. For a regression equation to be linear all exponents have to be unity, that is, $a = b =, \ldots = 1$; and each λ value, a cross-product coefficient, has to be zero. Solving the nonlinear problem is not any more difficult in these days and solving the linear problem due to availability of data science. In linear regression, when there is only one independent variable the regression analysis is called simple linear regression and if more than one independent variables then it is called Multiple Linear Regression Analysis [13].The proposed ML algorithm is used to develop a multiple linear regression analysis.

4.0.1 Data Preprocessing Algorithm

Step 1: Import Database using different ML libraries to python windows.
Step 2: To drop the dependent variable and get the independent variables.
Step 3: To find descriptive statistics
Step 4: The database split into independent variable (total vaccinated population) and dependent variables male, female and transgender.
Step 5: To define the variables Y and X.
Step 6: Print the data of X and Y and display the result and visualize the data of X and Y.

4.0.2 Training Algorithm

Step 1. Import ***sklearn.model_selection.train_test_split***library to split the database for testing and training the data.
Step 2. To fit linear regression model for Independent and dependent variables
Step 3: To predict the independent variable based on dependent variables.
Step 4: To identify the r2 score.

4.0.3 Prediction and Evaluation Algorithm

Step 1: To find the evaluation value of Mean Absolute Error (MAE), Mean Squared Error MSE and Root Mean Square Error (RMSE)
Step 2: To visualize the predicted score of entire database, it is linearly dependent.
Step 3: To display the Actual vaccinated population, predicted population and their difference in table 1.3 for testing states and union territories.
Step1 4: To visualize step 3 of the algorithm using line and simple bar charts.

4.1 Mean Absolute Error

Mean absolute error is the average of the errors between predicted and actual values where all errors are considered positive. If we don't change the sign, the positive and negative errors will cancel each other giving us a much lower value of error measures. Statistically MAE is given by the following expression.

$$MAE = \frac{|a_1 - C_1| + |a_2 - C_2| + \cdots + |a_n - C_n|}{n}$$

where, C_1 is the numerical value of prediction for the i^{th}test instance, a_i is the actual value of the i^{th} test instance, and n is the total number of test instances.

4.2 Root Mean Square Error

There is another way to get rid of the negative sign in errors instead of working with the absolute value. The new way is to square all the errors. All negative terms become positive when squared. The mean squared error isthe average of the squared difference between each computed value and its corresponding actual value. The root mean squared error is simply the square root of the mean square error. The root mean squared error brings back the squared errors in the same dimension as the actual and predicted values. The expression of RMSE in statistical notation is:

$$RMSE = \sqrt{\frac{\langle a_1 - C_1\rangle^2 + \langle a_2 - C_2\rangle^2 + \cdots + \langle a_n - C_n\rangle^2}{n}}$$

4.3 Mean Squared Error

Mean Squared Error of an estimator measures the average of error squares i.e. the average squared difference between the estimated values and true value. It is a risk function, corresponding to the expected value of the squared error loss. It is always non – negative and values close to zero are better. The MSE is the second moment of the error (about the origin) and thus incorporates both the variance of the estimator and its bias.

$$MSE = \frac{1}{N}\sum_{1=1}^{n} \langle Y_i - \hat{\grave{Y}}_\imath\rangle^2$$

Here N is the total number of observations/rows with in the dataset. The $Sigma$ symbol denotes the difference between actual and predicted values taken on every i^{th} value ranging from1 $to\ n$ **.**

4.4 $\boldsymbol{R^2}$ Score

Coefficient of determination also called as $\boldsymbol{R^2}$ score which is used to evaluate the performance of a linear regression model. It is the amount of variation in the output dependent attribute which is predictable from the input independent variables. It is used to check how well-observed results are reproduced by the model, depending on the ratio of total deviation of results described by the model.

$$\boldsymbol{R^2 = 1 - \frac{SS_{res}}{SS_{tot}}}$$

Where**,** is $\boldsymbol{SS_{res}}$ the sum of squares of the residual errors. $\boldsymbol{SS_{tot}}$t is the total sum of the errors.This can be implemented using ML **sklearn's** (Machine Learning Scikit-Learn)

5. RESULT AND DISCUSSION

The ML source code MRA generated using python script using Jupiter notebook and run the source code. In this research paper, the results of MLRA algorithm using ML algorithm is discussed in the following section.

The data preprocessing, the original database is spilt into testing and training for entire database at randomly. The separated independent variable named as Y (Total administered) and dependent variable as X (Male, Female and Transgender). Based on this X and Y MRA algorithm is executed and results of various portion of python script are achieved.

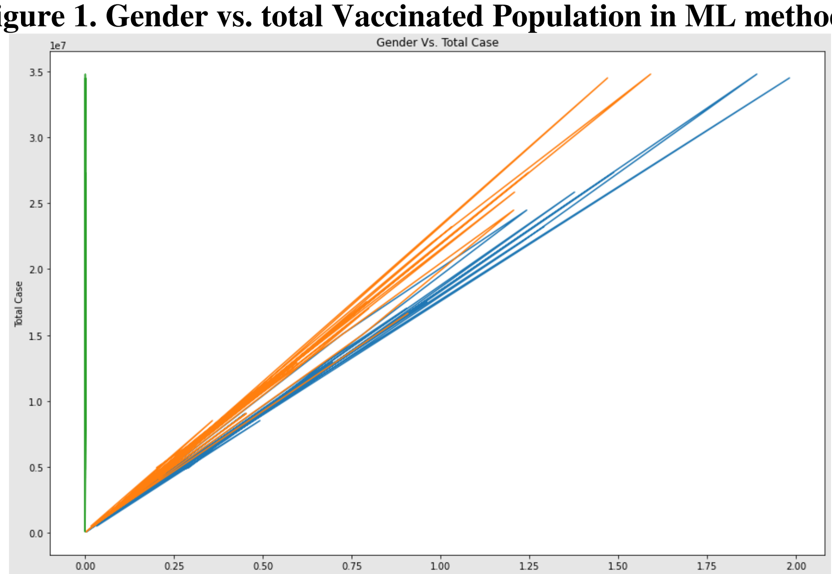
The vaccination summary statistics results show the Total Administered population, Male, Female and Transgender population. In 35 States and Union Territories of India on an average the vaccinated total population is 10171330, Male population is 5458707, Female population is 4710886 and transgender population is 1775. The Standard Deviation, minimum, maximum, 25^{th} percentile, 50^{th} percentile and 75^{th} percentiles are in Table 1 andFigure 1, shows the results of Gender vs. Total vaccinated population in particular period of database.

Table 1. Descriptive Statistics of Machine Learning (ML) methods

	Total	Male	Female	Transgender
count	35	35	35	35
mean	10171330	5458707	4710866	1755
std	10455210	5696222	4788877	1968
min	56755	31263	25489	3
25%	708814	382086	323920	147
50%	7444490	4091291	3351946	1125
75%	16825570	8296749	7946128	2637
max	34802540	19811050	15905160	9117

The below visualization results shows that, the states Maharashtra, Uttar Pradesh (30000000 and above), Karnataka, Rajasthan, Madhya Pradesh, Gujarat, Rajasthan and West Bengal followed by other states and UTs is vaccinated with COVISHIELD. The COVAXIN vaccination is administered in the states of Maharashtra, Uttar Pradesh, Karnataka, Rajasthan, Karnataka, followed by other states and UTs.

Figure 1. Gender vs. total Vaccinated Population in ML methods



5.1 Evaluation

ML methods of Multiple Regression Analysis (MRA) algorithm, shows the results ofMean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error MSE and R-square Values in Table 2. The R-square range is from 0 to 1. The R square value 0indicates that it is not suitable model for any given database and 1 indicates best fitted model for any database. In this research, the vaccination of two doses with total population and gender are fitted best model, and R-square value indicates 0.999999. The error rate result is showed in table 2. The scatter diagram shows the linear relationship of results of Actual and predicted value of vaccinated population.This means that, the MRAalgorithm is very accurate and makes good predictions.

Table 2. The evaluation of MLRA model for Total Population vs. Gender

Mean Absolute Error:	0.01201003
Mean Squared Error:	0.10521001
Root Mean Squared Error:	0.10542300
R2 Score:	0.99999999

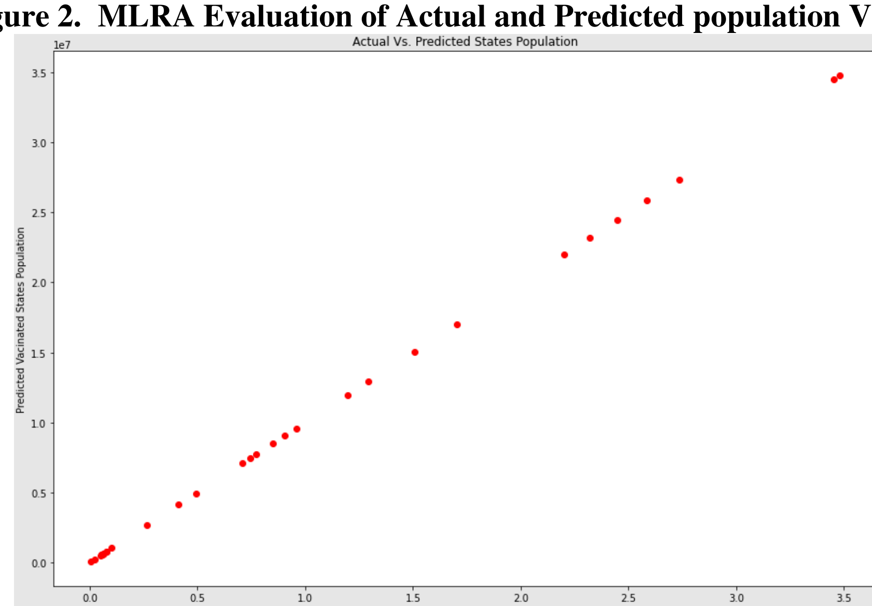
Mean Absolute Error (MAE) takes the **average** of this error from every sample in a dataset and gives the output. The output of MAE value is 0.012.

The Mean Squared Error (MSE) is calculated by taking the average of square of difference between original and predicted values of the data.The MSE value is 0.1052.

The Root Mean Squared Error (RMSE) is the standard deviation of the errors which occur when a prediction is made on a dataset. This is the same as MSE (Mean Squared Error) but the root of the value is considered while determining the accuracy of the model. The Value of RMSE is 0.1054.

The above table$R^2 = 0.99999$, it can be referred that 99.9% of changeability of the dependent output attribute can be explained by the model while the remaining 0.1 % of the variability is still unaccounted for R^2 indicates the proportion of data points which lie within the line created by the regression equation. A higher value of R^2 is desirable as it indicates better results. This result achieved hundred percent between actual and predicted values and 0.1% deviated between actual and predicted values (Table 3).

Figure 2. MLRA Evaluation of Actual and Predicted population Value



5.2 Prediction and Visualization

The\actual vaccinated population, Predicted population and differences are executed and displayed in Table 3, Figure 3 and 4. This results shows that the actual and predicted are linearly dependent with minor differences of all Indian States and Union Territories (UTs). The more vaccinated states are Uttar Pradesh followed by Rajasthan, Karnataka, West Bengal, Tamilnadu, Kerala, Telengana and Odisha, etc. The more vaccinated Union Territories are Goa followed byPuduchery, Andaman Nicobar Lakshadweep, Ladakh and Dadra and Nagar Haveli and Daman and Diu, etc.

Table 3. States and UTS Actual, Predicted and Difference of both dose administered

States and UTs	Actual	Predicted	Differences
Tamil Nadu	17017222	17017630	-405.71
Manipur	785691	7.86E+05	-639.527
Ladakh	237455	2.38E+05	-743.288
Sikkim	527470	5.28E+05	-631.957
Mizoram	620536	6.21E+05	-706.222
Kerala	15062316	1.51E+07	-147.765
Haryana	9581886	9.58E+06	-389.82
Assam	7712614	7.71E+06	-694.02
Himachal Pradesh	4115822	4.12E+06	-427.84
Rajasthan	25848234	2.58E+07	-36.4227
Puducherry	548593	5.49E+05	-685.568
West Bengal	23205213	2.32E+07	-869.098
Uttar Pradesh	34513734	3.45E+07	3128.807
Punjab	7095644	7.10E+06	-1008.46
Telangana	11966330	1.20E+07	-142.656
Goa	1021435	1.02E+06	-778.561
Jharkhand	7444490	7.45E+06	-552.23
Chhattisgarh	9044519	9.05E+06	-894.183
Lakshadweep	56755	5.75E+04	-737.709
Karnataka	24477242	2.45E+07	142.9219
Tripura	2642602	2.64E+06	-723.587
Odisha	12923503	1.29E+07	-204.795
Arunachal Pradesh	654032	6.55E+05	-631.021
Jammu and Kashmir	4928467	4.93E+06	-619.226
Dadra and Nagar Haveli and Daman and Diu	490259	4.91E+05	-713.075

Figure 3. States and UTS Actual, Predicted and Difference of both dose administered in Line Chart

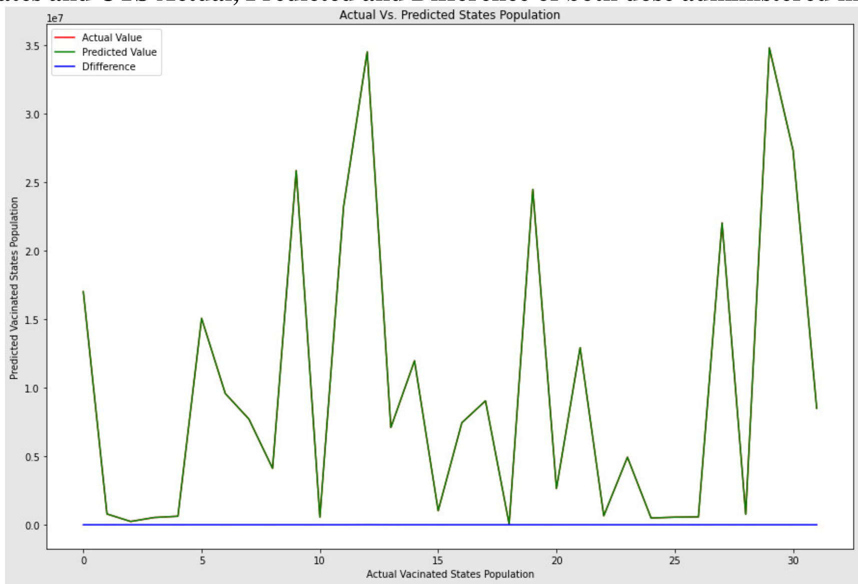
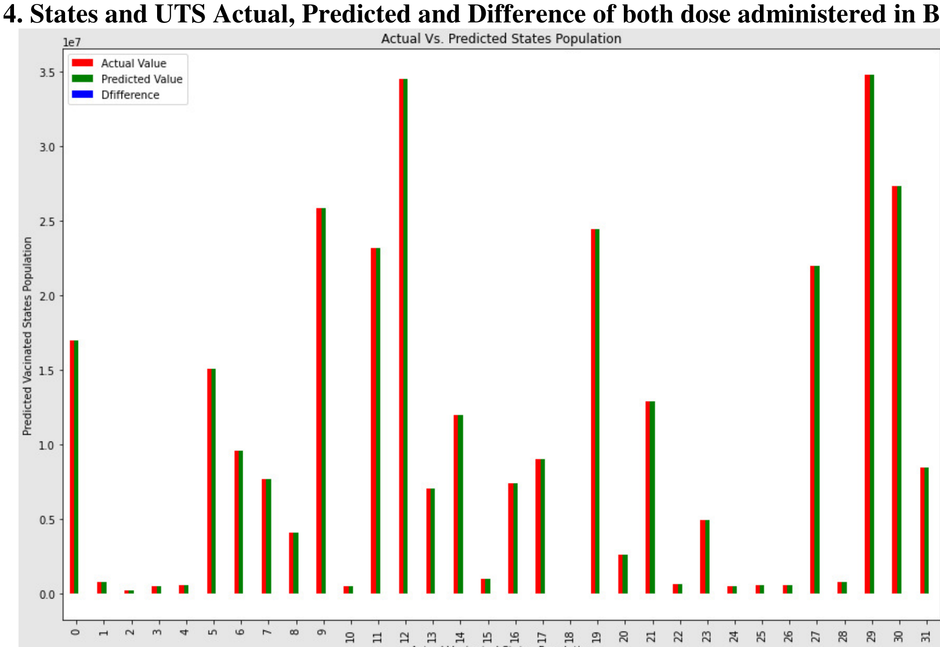


Figure 4. States and UTS Actual, Predicted and Difference of both dose administered in Bar Chart



6. CONCLUSION

This research paper predicts and identifies the differences between Total vaccinated (with different vaccination) people vs. gender using Machine Leering (ML) method of Multiple Regression Analysis (MRA) algorithm. The secondary sources of data were collected from Ministry of Health and Family Welfare Department, India. Four main parameters are used in this research paper and they are Total vaccinated population, Male, Female and Transgender population. The independent variable X(Total vaccinated population) and Y (Gender) are dependent variables. Initially the data set is split into two categories, test and training data of total vaccinated population of the entire states and UTs of the country. The python script writes and imports the entire database and split database into train and test category in the ratio of 30:70. The 30 percent of data is used for training and remaining 70 percentages for testing. The predicted results of MRA with original database and also identified the differences between actual and predicted data of entire datasets. The statistical function displays the result of descriptive statistics. In this research, the vaccination of two doses with total population and gender are fitted best model, and the R-square value indicates 0.999999. The error rate result is showed in table 2. The scatter diagram shows the linear relationship of results of Actual and predicted value of vaccinated population. This means that, the MRA algorithm is very accurate and makes good predictions.Finally the data visualized MRA in the form of graphical representation, this results shows that the actual and predicted MRA results achieved hundred percent between total population and predicted population. Visualization also clearly displayed in MRA results by line and simple bar diagrams. The more vaccinated states are Uttar Pradesh followed by Rajasthan, Karnataka, West Bengal, Tamilnadu, Kerala, Telengana and Odisha, etc. The more vaccinated Union Territories are Goa followed by Puduchery, Andaman Nicobar Lakshadweep, Ladakh and Dadra and Nagar Haveli and Daman and Diu, etc.

REFERENCES

1. *Brian D.A., Baric R.S. Coronavirus genome structure and replication. Curr Top MicrobiolImmunol. 2005;287:1-30.*
2. *9. Burrell J.C., Howard C.R., Murphy F.A. Fenner and white's medical virology. Fifth ed. © 2016 Elsevier Inc; 2017.Corona viruses.*
3. *"Navigating "Down Time" During COVID-19.Available from https://www.ncbi.nlm.nih.gov/pmc /articles/PMC7158834/.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7158834.*
4. *Rajagopalan, S., Huang, S., Brook, R.D.: Flattening the curve in COVID-19 using personalised protective equipment: lessons from air pollution. Heart (2020).URL https://dx.doi.org/10.1136/ heartjnl-2020-317104.*
5. *Rajagopalan, , Al-Kindi, S.G., Brook, R.D.: Air Pollution and Cardiovascu- lar Disease: JACC State-of-the-Art Review. Journal of the American College of Cardiology 72(17), 2054-2070 (2018).DOI 10.1016/j.jacc.2018.07.099.Available from https://www.onlinejacc.org/content/72/ 17/2054. Publisher: Journal of the American College of Cardiology Section: The Present and Future.*
6. *SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with in- nate immune genes Nature Medicine. Available from https://www.nature.com/articles/s41591- 020-0868-6.*
7. *Estimation of the final size of the COVID-19 epidemic medRxiv. Available from https://www.medrxiv.org/content/10.1101/2020.02.16.20023606v5.*
8. *Forecasting the novel coronavirus COVID-19. Aavailable from https://journals.plos.org/plosone /article?id=10.1371/journal.pone.0231236*
9. *Gaurav Pandey, Poonam Chaudhary, Rajan Gupta, Saibal Pal. SEIR and Regression Model based COVID-19 outbreak predictions in India. arXiv preprint arXiv:2004.00958, 2020*
10. *Rajan Gupta, Saibal Kumar. Trend Analysis and Forecasting of COVID-19 outbreak in India. Rajan Gupta, Saibal Kumar PalmedRxiv, 2020. Available from doi: https://doi.org/10.1101/ 2020.03.26.20044511*
11. *Manimannan G. et.al. (2021), International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-10 Issue-1, pp. 16-20.*
12. *State level data comes from Ministry of Health & Family Welfare Testing data and vaccination data comes from covid19 India.*
13. *Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Lawrence Erlbaum Associates Publishers.*

DOI No. : 10.5958/2394-9309.2022.00021.X
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

IFGP* - NEIGHBORHOOD IN INTUITIONISTIC FUZZY TOPOLOGICAL SPACES

Chandhini* Uma. N**
J Research scholar, Department of Mathematics, Sri Ramakrishna College of Arts and Science (SNR Sons), Coimbatore, Tamil Nadu. (INDIA) Prof.& Head, Department of Mathematics, Sri Ramakrishna College of Arts and Science (SNR Sons), Coimbatore, Tamil Nadu.
E-mail : chandhujk@gmail.com, uma.n@srcas.ac.in

ABSTRACT
Earlier, the intuitionistic fuzzy gp* - closed sets, Intuitionistic fuzzy gp* - open sets, characteristic of Intuitionistic fuzzy gp* - closed sets were introduced in Intuitionistic Fuzzy Topological Spaces. In this paper, the neighborhood of intuitionistic fuzzy gp* - closed set is introduced. We have also investigated their properties in Intuitionistic Fuzzy Topological Spaces.
Keywords : closed sets, open sets and neighborhood of fuzzy gp* - closed sets, intuitionistic fuzzy gp* - closed sets; intuitionistic fuzzy gp* - open sets; neighborhood of intuitionistic fuzzy gp*.

1. INTRODUCTION

Fuzzy set (FS) was proposed by **Zadeh [19]** in 1965, there have been a number of generalizations if this fundamental concept. The concept of generalized closed sets in fuzzy topological spaces was introduced by **Thakur** and**Malvika[17]. Benchalli** and **Siddapur [5]** initiated the concept of fuzzy generalized star pre closed sets. **Bhattacharyya** and **Anjana [6]**proposed the concept of fuzzy alpha generalized closed sets. **Bayaz Daraby** and **Nimse[4]**found the concept of fuzzy generalied alpha closed sets. The concept of fuzzy generalized pre closed sets was presented by **Fukutake**, **Saraf,Caldas** and **Mishra [11]**. **Firdose Habib** and **Khaja Moinuddin [10]**found the concept of fuzzy generalized pre star closed sets. The notion of intuitionistic fuzzy sets developedby **Atanassov [2]** is one among them. Using the notion of intuitionist fuzzy set, **Coker [9]**institutedthe notion of intuitionist fuzzy topological spaces. **Thakur** and **Rekha Chaturvedi [18]**introduced the concept of Intuitionistic fuzzy generalized closed sets. **Jyothi Pandey Bajpai** and **Thakur [14]** initiated the concept of Intuitionistic fuzzy generalized star pre - closed sets. **Kalamani,Sakthivel** and**Gowri [13]**instituted the concept of Intuitionistic fuzzy generalized alpha closed sets. The concept of Intuitionistic fuzzy generalized # closed sets was developed by **Abhirami[1]. Sakthivel [16]**proposed the concept of Intuitionistic fuzzy alpha generalized - closed sets. **Rajarajeswari** and **Senthil kumar [15]**launched the concept of Intuitionistic fuzzy generalized pre closed sets.In this paper we introduce IFgp*- neighborhoodin topological spaces by using the notions of IFgp* - open sets and study some of their properties.

2. PRELIMINARIES

2.1The following definitions are the fuzzy topological spaces.

Definition 2.1.1[7]: A **fuzzy pre - open set** if $A \subseteq int\big(cl(A)\big)$and a **fuzzy pre - closed set** if $cl(int(A) \subseteq A$.

Definition 2.1.2[3]: A **fuzzy semi - open set** if $A \subseteq cl(int(A))$ and a **fuzzy semi - closed set** if $int\big(cl(A)\big) \subseteq A$.

Definition 2.1.3[7]: A **fuzzy α - open set** if $A \subseteq int(cl(int(A)))$ and a **fuzzy α - closed set** if $cl(int(cl(A)) \subseteq A$.

Definition 2.1.4[17]: Let A be a subset of topological spaces(X, τ), is called a **fuzzy generalized closed set** (briefly fg-closed) if $cl(A) \subseteq U$whenever$A \subseteq U$ and$U$ is fuzzy open in$X$.

Definition 2.1.5[11]: Let A be a subset of topological spaces (X, τ), is called a **fuzzy generalized pre closed set** (briefly fgp-closed) if $pcl(A) \subseteq U$whenever $A \subseteq U$ and $U$is fuzzy open in$X$.

Definition 2.1.6[4]: Let A be a subset of topological spaces (X, τ), is called a **fuzzy generalized α closed set** (briefly fgα-closed) if $\alpha cl(A) \subseteq U$whenever $A \subseteq U$and U is fuzzy α open in$X$.

Definition 2.1.7[6]: Let A be a subset of topological spaces (X, τ), is called a **fuzzy α generalized closed set** (briefly fαg-closed) if $\alpha cl(A) \subseteq U$whenever $A \subseteq U$and $U$is fuzzy open in$X$.

Definition 2.1.8[5]: Let A be a subset of topological spaces (X, τ), is called a **fuzzy generalized pre closed set** (briefly fg*p-closed) if $pcl(A) \subseteq U$whenever $A \subseteq U$and $U$is fuzzy generalized open in$X$.

Definition 2.1.9[10]: Let A be a subset of topological spaces (X, τ), is called a **fuzzy generalized pre star closed set** (briefly fgp*-closed) if$cl(A) \subseteq U$whenever $A \subseteq U$and$U$ is fuzzy generalized pre open in$X$.

Definition 2.1.10[10]: Let x be the point in topological space X and let $x\epsilon X$. A subset N of X is said to be a gp*- nbhd of x iff there exists a gp* - open set G such that $x \in G \subset N$.

Definition 2.1.11[10]: A subset N of space X is called gp* - nbhd of $A \subset X$ iff there exists a gp*- open set G such that $A \subset G \subset N$.

2.2 The followings are the intuitionistic fuzzy topological spaces definitions.

Definition 2.2.1[2]: Let X be a nonempty fixed set. An intuitionistic fuzzy set(IFS in short) A in $X$is an object having the form $A = \{< x, \mu_A(x), \nu_A(x) > / x\epsilon X\}$ where the functions $\mu_A(x): X \to [0,1]$and$\nu_A(x): X \to [0,1]$ denote the degree of membership (namely $\mu_A(x)$) and the degree of non - membership (namely $\nu_A(x)$) of each element $x\epsilon X$ to a set A respectively and $0 \leq \mu_A(x) + \nu_A(x) \leq 1$ for each $x\epsilon X$.

Definition 2.2.2[2]: Let A and B be IFSs of the form $A = \{< x, \mu_A(x), \nu_A(x) > / x\epsilon X\}$ and $B = \{< x, \mu_B(x), \nu_B(x) > / x\epsilon X\}$. Then

a) $A \subset B$ if and only if $\mu_A(x) \leq \mu_B(x)$ and $\nu_A(x) \geq \nu_B(x)$ for all $x\epsilon X$,

b) $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$,

c) $A^C = \{< x, \mu_A(x), \nu_A(x) > / x\epsilon X\}$,

d) $A \cap B = \{< x, \mu_A(x) \cap \mu_B(x), \nu_A(x) \cup \nu_B(x) > / x\epsilon X\}$,

e) $A \cup B = \{< x, \mu_A(x) \cup \mu_B(x), \nu_A(x) \cap \nu_B(x) > / x\epsilon X\}$.

For the sake of simplicity, the notation $A = < x, \mu_A, \nu_A >$ shall be used instead of $A = \{< x, \mu_A(x), \nu_A(x) > / x\epsilon X\}$. Also for the sake of simplicity, we use the notation $A =< x, (\mu_A(x)\mu_B(x)), (\nu_A(x), \nu_B(x)) >$ instead of $A =< x, (A/\mu_A, B/\mu_B), (A/\nu_A, B/\nu_B) >$. The intuitionistic fuzzy sets $0_\sim = \{< x, 0,1 > / x\epsilon X\}$ and $1_\sim = \{< x, 1,0 > / x\epsilon X\}$ are the empty set and the whole set of X, respectively.

Definition 2.2.3[9]: An intuitionistic fuzzy topology (IFT in short) on a non empty set X is a family τ of IFSs in X satisfying the following axioms:

a) $0_\sim, 1_\sim \in \tau$,

b) $G_1 \cap G_2 \in \tau$ for any $G_1, G_2 \in \tau$,

c) $UG_i \in \tau$ for any arbitrary family $\{G_i | i \in J\} \subseteq \tau$.

In the case, the pair (X, τ) is called an intuitionistic fuzzy topological spaces(IFTS in short) and any IFS in τ is known as an intuitionistic fuzzy open set (IFOS in short) in X. The complement A^C of an IFOS A is an IFTS (X, τ) is called an intuitionistic fuzzy closed set(IFCS in short) in X.

Result 2.2.4[9]: Let A and B be any two intuitionistic fuzzy sets of an intuitionistic fuzzy topological space (X, τ). Then

a) A is an intuitionistic fuzzy closed set in $X \Leftrightarrow cl(A) = A$,

b) A is an intuitionistic fuzzy closed set in $X \Leftrightarrow int(A) = A$,

c) $cl(A^C) = (int(A))^C$,

d) $int(A^C) = (cl(A))^C$,

e) $A \subseteq B \Rightarrow int(A) \subseteq int(B)$,

f) $A \subseteq B \Rightarrow cl(A) \subseteq cl(B)$,

g) $cl(A \cup B) = cl(A) \cup cl(B)$,

h) $int(A \cap B) = int(A) \cap int(B)$.

Definition 2.2.5[9]: Let (X,τ)be an IFTS and $A=<x,\mu_A,\nu_A>$be an IFS in$X$. Then the intuitionistic fuzzy interior and an intuitionistic fuzzy closure are defined by

a) $int(A)=\cup\{G/G\ is\ an\ IFOS\ in\ X\ and\ G\subseteq A\}$,

b) $cl(A)=\cap\{K/K\ is\ an\ IFCS\ in\ X\ and\ A\subseteq K\}$.

Definition 2.2.6: An IFS$A=\{<x,\mu_A(x),\nu_A(x)>/x\epsilon X\}$is an IFTS$(X,\tau)$ is said to be

a) A **intuitionistic fuzzy pre - open set** [12] if $A\subseteq int(cl(A))$ and a **intuitionistic fuzzy pre - closed set** if $cl(int(A)\subseteq A$.

b) A **intuitionistic fuzzy semi - open set**[12] if $A\subseteq cl(int(A))$and a **intuitionistic fuzzy semi - closed set** if $int(cl(A))\subseteq A$.

c) A **intuitionistic fuzzy α- open set**[9] if $A\subseteq int(cl(int(A)))$ and a **intuitionistic fuzzy α- closed set** if$cl(int(cl(A))\subseteq A$.

d) A **intuitionistic fuzzy generalized closed set** [18] (briefly IFg-closed set) if$cl(A)\subseteq U$whenever A $\subseteq U$and$U$ is intuitionistic fuzzy open in X.

e) A **intuitionistic fuzzy generalized pre closed set**[15] (briefly IFgp-closed set) if $pcl(A)\subseteq U$whenever A $\subseteq U$and U is intuitionistic fuzzy open in X.

f) A **intuitionistic fuzzy generalized # closed set** [1] (briefly IFg#-closed set) if$cl(A)\subseteq U$whenever A $\subseteq U$and U is intuitionistic fuzzy generalized α open in X.

g) A **intuitionistic fuzzy generalized α closed set**[13] (briefly IFgα-closed set) if $\alpha cl(A)\subseteq U$whenever A $\subseteq U$and U is intuitionistic fuzzy α open in X.

h) A **intuitionistic fuzzy α generalized closed set** [16] (briefly IFαg-closed set) if $\alpha cl(A)\subseteq U$whenever A $\subseteq U$and U is intuitionistic fuzzy open in X.

i) A **intuitionistic fuzzy generalized pre closed set** [14] (briefly IFg*p-closed set) if $pcl(A)\subseteq U$ whenever A $\subseteq U$and U is intuitionistic fuzzy generalized open in X.

j) A **intuitionistic fuzzy pre generalized closed set** (briefly IFpg-closed set) if$pcl(A)\subseteq U$ whenever A $\subseteq U$and U is intuitionistic fuzzy pre open in X.

Definition 2.2.7: An intuitionistic fuzzy sets A of a intuitionistic fuzzy topological sets (IFTS) $(X,\tau)$is an called **intuitionistic fuzzy generalized pre starclosed** (briefly IFgp* - closed) if $cl(A)\subseteq U$ whenever$A\subseteq U$and $U$is intuitionistic fuzzy generalized pre - open set in X. The family of all IFgp*cs of ifts (X,τ) is denoted by IFgp*c(X).

Remark 2.2.8: For any IFTS (X,τ), we have the following:

a) Every IF closed sets are IFgp* - closed sets.

b) Every IFg*p - closed sets are IFgp* - closed set.

c) Every IFpg- closed sets are IFgp* - closed sets.

d) Every IFgp*- closed sets are IFg # - closed sets

e) Every IFgp* - closed sets are IFαg - closed sets.

Definition 2.2.9: Suppose a Intuitionistic fuzzy set A is intuitionistic fuzzy generalized pre star closed set in IFTS(X, τ), Then its complement (i,e) 1 - A is called **intuitionistic fuzzy generalized pre star open set** (briefly IFgp* - open) in(X, τ).

Remark 2.2.10: For any IFTS(X, τ), we have the following:

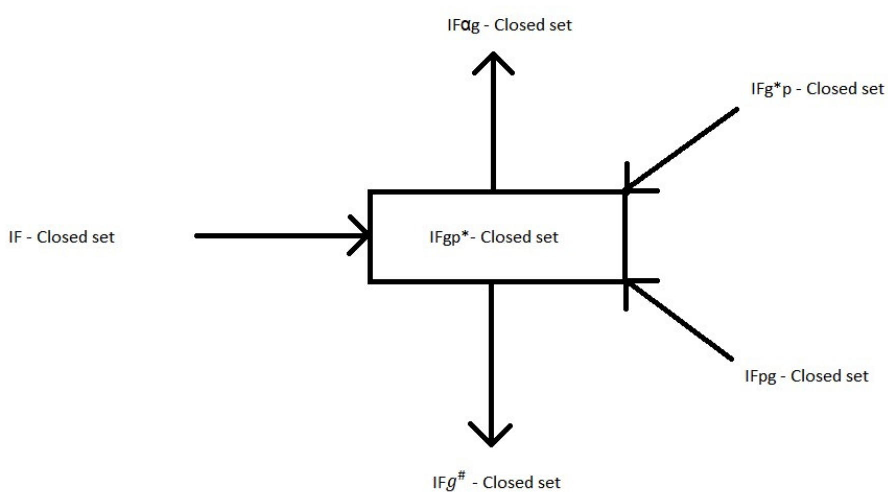
(a) Every IFg*pos is an IFgp*os.

(b) Every IFpgos is an IFgp*os.

(c) Every IFgp*os is an IFg #os.

(d) Every IFgp*os is an IFαgos.

Remark 2.2.11: The following diagram depicts the relation of intuitionistic fuzzy gp* closed set.



3. INTUITIONISTIC FUZZY gp* - NEIGHBORHOOD

In this section we introduce IFgp*- neighborhood in topological spaces by using the notions of IFgp* - open sets and study some of their properties.

Definition 3.1: Let x be the point in intuitionistic fuzzy topological space X and let $x \epsilon X$. A subset N of X is said to be a IFgp*- nbhd of x iff there exists a IFgp* - open set G such that $x \in G \subset N$.

Definition 3.2: A subset N of space X is called IFgp* - nbhd of $A \subset X$ iff there exists a IFgp*- open set G such that $A \subset G \subset N$.

Theorem 3.3: Every intuitionistic fuzzy nbhd N of $x \in X$ is a IFgp* - nbhd of X.

Proof: Let N be a intuitionistic fuzzy nbhd of a point $x \in X$. To prove that N is a IFgp* - nbhd of$X$. By definition of intuitionistic fuzzy nbhd, there exist an intuitionistic fuzzy open set Gsuch that $x \in G \subset N$.Hence N is a IFgp* - nbhd of$x$.

Theorem 3.4: If a subset N of a space $X$is IFgp* - open, then N is IFgp* - nbhd of each of its point.

Proof: suppose N is IFgp* - open. Let $x \in N$. We claim that N is IFgp* - nbhd of x. For N is IFgp* - open set such that$x \in N \subset N$. Since x is an arbitrary point of N, it follows that N is a IFgp* - nbhd of each of its point.

Theorem 3.5: Let X be a intuitionistic fuzzy topological space. If F is IFgp* - closed subset of X and $x \in F^C$. Prove that there exists a IFgp* - nbhd N of x such that $N \cap F = \varphi$.

Proof: Let F be IFgp* - closed subset of X and $x \in F^C$. ThenF^C is a IFgp - open set of X. So by definition 3.2, F^C contains a IFgp* - nbhd of each of its points. Hence there exists a IFgp* - nbhd N of x such that $N \subset F^C$ that is $N \cap F = \varphi$.

Definition 3.6: Let x be a point in a intuitionistic fuzzy topological space X. The set of all IFgp*- nbhd of x is called the IFgp* - nbhd system at x, and is denoted by IFgp* - $N(x)$.

Theorem 3.7: Let a IFgp*-nbhd N of Xbe a intuitionistic fuzzy topological space and each $x \in X$, let IFgp* -$N(X,\tau)$ be the collection of all IFgp* - nbhd of x. Then we have the following results.

(i) For all $x \in X$, IFgp∗ - $N(x) \neq \phi$

(ii) $N \in$IFgp∗ -$N(x) \Rightarrow x \in N$

(iii)$N \in$IFgp∗ - $N(x), M \supset N \Rightarrow M \in IFgp * -N(x)$.

(iv)$N \in$ IFgp∗ - $N(x), M \in IFgp * -N(x) \Rightarrow N \cap M \in IFgp * -N(x)$.

(v) $N \in$ IFgp∗-$N(x) \Rightarrow$ there exist $M \in IFgp * -N(x)$such that $M \subset N$and $M \in IFgp * -N(y)$ for every $y \in M$.

Proof:

(i) Since X is IFgp* - open set, it is a IFgp* - nbhd of every $x \in X$. Hence there exist at least one IFgp* - nbhd (Namely X) for each $x \in X$. Therefore IFgp∗ - $N(x) \neq \phi$ for each $x \in X$.

(ii) If N $\in$IFgp∗ - $N(x)$, then N is IFgp* - nbhd of x. By definition of IFgp* - nbhd, $x \in N$.

(iii)Let $N \in$ IFgp∗ - $N(x)$ and $M \supset N$. Then there is a IFgp* - open set G such that $x \in G \subset N$. Since N $\subset$ M, $x \in G \subset M$ and so M is IFgp* - nbhd of x. Hence $M \in IFgp * -N(x)$

(iv)Let $N \in$ IFgp∗ - $N(x)$, $M \in IFgp * -N(x)$. Then by definition of IFgp* - nbhd, there exist IFgp* - open sets $G_1$and G_2 such that $x \in G_1 \subset N$ and $x \in G_2 \subset M$. Hence $x \in G_1 \cap G_2 \subset N \cap M$ −−−−− (1) Since $G_1 \cap G_2$ is a IFgp* - open set. [Being the intersection of two regular open sets], it follows from (1) that $N \cap M$ is a IFgp* - nbhd of x. Hence $N \cap M \in IFgp * -N(x)$.

(v) Let $N \in$ IFgp$*$ - $N(x)$, then there is a IFgp* - open set M such that$x \in M \subset N$. Since M is IFgp* - open set, it is IFgp* - nbhd of each of its points. Therefore $M \in IFgp * - N(y)$ for every $y \in M$.

CONCLUSION

In this paper wehave investigated a new form of intuitionistic fuzzy neighborhood called intuitionistic fuzzy gp*-neighborhood which contain the classes of intuitionistic fuzzy gp*-closed sets, intuitionistic fuzzy gp*-open sets. We have also investigated their properties in Intuitionistic Fuzzy Topological Spaces.

REFERENCE

1. ***S. Abhirami, R. Dhavaseelan***, *Intuitionistic fuzzy g# closed sets, International journal of research in advent technology, vol. 2, no. 3, March 2014.*
2. ***K. T. Atanassov***, *Intuitionistic fuzzy sets, Fuzzy sets and systems, 20(1986), 87 - 96.*
3. ***K. K. Azad***, *On fuzzy semi continuity, Fuzzy almost continuity and fuzzy weakly continuity, J.Math.Anal.Appl. 82(1) (1981), 14 - 32.*
4. ***Bayaz Daraby*** *and* ***S. B.Nimse***, *On fuzzy generalized alpha closed sets and its applications, Faculty of sciences and mathematics, university of Nis, Serbia, Filmat 21:2(2007), 99 - 108.*
5. ***S. S. Benchalli*** *and* ***G. P. Siddapur***, *On fuzzy g* pre continuous maps in fuzzy topological spaces, International journal of computer applications, vol. 16, NO. 2, February 2011.*
6. ***Bhattacharyya, Anjana***, *fg*α continuous function in fuzzy topological spaces, International journal of scientific and engineering research, Vol. 4, issue 8(2013), 973 - 979.*
7. ***Bin Shanna, A. S.,****On fuzzy strongly semi continuity and fuzzy pre continuity, fuzzy sets and systems 4(11) (1991), 330 - 338.*
8. ***R. Chaturvedi***, *Some classes of generalized closed sets in intuitionistic fuzzy topological spaces, Ph. D dissertation, Rani Durgavati Vishwavidyalaya, Jabalpur, India (2008).*
9. ***D. Coker***, *An introduction to intuitionistic fuzzy topological spaces, Fuzzy sets and systems 88, 1997, 81 - 89.*
10. ***Firdose Habib and Khaja Moinuddin***, *Fuzzy gp* closed sets in fuzzy topological space, South East Asian J. of mathematics and mathematical sciences, vol. 16, no. 2(2020), pp. 151 - 160.*
11. ***T. Fukutake, R. K. Saraf, M. Caldas*** *and* ***S. Mishra***, *Mapping via fgp - closed sets, Bull. of Fukuoka Univ. of Edu. Vol. 52(2003) part III, 11 - 20.*

12. ***H. Gurcay, A. Haydar*** *and* ***D. Coker****, On fuzzy continuity in Intuitionistic fuzzy topological spaces, Jour. of fuzzy math, 5(1997), 365 - 378.*

13. ***D. Kalamani, K. Sakthivel*** *and* ***C. S. Gowri****, Generalized alpha closed sets in intuitionistic fuzzy topological spaces, Applied mathematical sciences, vol. 6, 2012, no. 94, 4691 - 4700.*

14. ***Jyothi Pandey Bajpai, S. S. Thakur,****Intuitionistic fuzzy g*p closed sets, Global Journal of pure and applied mathematics, ISSN 0973 - 1768 Volume 14, Number 7(2018), pp.955 - 975.*

15. ***P. Rajarajeswari*** *and* ***L. Senthil kumar****, Generalized pre-closed sets in Intuitionistic fuzzy topological sapces, International journal of fuzzy mathematics and systems, 3(2011), 253 - 262.*

16. ***K. Sakthivel****, Intuitionistic fuzzy alpha generalized closed sets and intuitionistic fuzzy alpha generalized open sets, The mathematical Education 4(2012).*

17. ***S. S. Thakur*** *and* ***R. Malvika****, Generalized closed sets in fuzzy topology, Math Notae, 38(1995), 137 - 140.*

18. ***S. S.Thakur*** *and* ***Reka Chaturvedi****, some classes of generalized closed sets in intuitionistic fuzzy topological spaces, Ph. D dissertation, Rani Durgavati Vishwavidyalaya, Jabalpur, India (2008).*

19. ***L. A. Zadeh****, Fuzzy sets, Information control, 8(1965), 338 - 353.*

DOI No. : 10.5958/2394-9309.2022.00022.1
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ESTIMATING SURVIVAL TIME OF CHRONIC KIDNEY DISEASE PATIENTS USING MODEL

Elizabeth Sangeetha T*, P. Pandiyan, M Meganathan*****
*Research scholar, Department of Statistics, Annamalai University, Annamalai Nagar, Tamil Nadu, India.
**Professor of Statistics, Department of Statistics, Annamalai University, Annamalai Nagar, Tamil Nadu, India
***Associate Professor, Department of Pharmacology, AVMC&H, Pondicherry
Corresponding author : est3383@gmail.com

ABSTRACT
Researchers may use mathematical models to study laws and scams. To aid in the decision-making process for claims in our community. It has the ability to predict and analyse likely events. They demonstrate many scenarios before they need to be thoroughly studied. The kidneys were studied using randomised models to see how they were impacted by the human system. Shot modelling is a term used to describe the process of creating a video. We use five variables to illustrate complicated issues. Our model gives you a behind-the-scenes look at these five aspects that have an impact on patients' human systems. with the help of a shock simulator with a survival function. The study found that it is easier to treat a sickness if it is found early on.
Keywords: Chronic Kidney Diseases, Factors, Model, Stage

1. INTRODUCTION

It should be noted that this is not the case in the general population. The preceding goals were different since they were based on co-morbidity concerns rather than age, race, or sex. Kidney disease does not develop overnight; it occurs in phases. Early on, those with symptoms may be unable to recognise anything odd. Kidney diseases may be delayed or prevented if they are recognised and treated early. Chronic kidney disease (CKD) is a serious public health issue with enormous financial implications for the human health system. CKD has been linked to a worse quality of life as well as a higher chance of mortality and degradation. Body of the coronary arteryKidney function decline is linked to an increased risk of mortality and cardiovascular issues. CKD is a condition that is often linked to high blood pressure, diabetes, anaemia, and mineral problems. element/bone The early stages of CKD include diabetes and high blood pressure.

CKD may be classified into five defined phases of infection according to the National Kidney Foundation's recommendation and based on the estimated glomerular filtration rate (eGFR) range. In Western societies, the worldwide prevalence of CKD ranges from 10.5 percent to 13.1 percent, with two-thirds of CKD resulting in diabetes and high blood pressure. In India, there are 800 people per million with CKD, which is a measure of how effectively the kidneys are performing. Creatinine levels in the blood will be checked. The estimated glomerular filtration rate is then calculated using the rates (eGFR). Natural renal function declines with ageing in healthy people. Adults aged 20–30, for example, have an eGFR of roughly 115 ml/min/1.73 sq m, while those aged 60–69 have an eGFR of approximately 85 ml/min/1.73 sq m.

The epidemic model may be used to determine the transmission of infectious illnesses as an analytical tool. The epidemic model provides a framework for infectious disease data that is easy to understand and apply. Its primary goal, however, is to generate ideas regarding biological and sociocultural transmission mechanisms. Current epidemic models include a number of assumptions about how diseases spread. These assumptions are estimates of the original process's scope. The following are the most common symptoms of these conditions:

2. MODEL DEVELOPMENT

Let Y be the random variable which has the cdf defined as

$$F(x;\theta) = \left[1 - e^{-\left(\frac{x}{\sigma}\right)^{\beta}}\right]^{\alpha} \qquad ; x > 0 \tag{1}$$

$$1 - e^{-\left(\frac{x}{\sigma}\right)^{\beta}}$$

and has the probability density function (p.d.f)

$$f(x;\theta) = \frac{\alpha\beta}{\sigma}\left(\frac{x}{\sigma}\right)^{\beta-1} e^{-\left(\frac{x}{\sigma}\right)^{\beta}}\left[1 - e^{-\left(\frac{x}{\sigma}\right)^{\beta}}\right]^{\alpha} \qquad ; x > 0$$

The corresponding survival function is $\bar{H}(x) = 1 - F(x) = e^{-\left(\frac{x}{\sigma}\right)^{\beta}}$

One is interested in an object for which shock resistance varies significantly across individuals. It's possible that inspecting a single object to establish its y threshold is impractical. The threshold in this scenario must be a random variable. The shock survival probability is calculated as follows:

$$P(X_i < Y) = \int_0^{\infty} g^*(x)\bar{H}(x)dx = \int_0^{\infty} g^*(x)\left[e^{-\left(\frac{x}{\sigma}\right)^{\beta}}\right]dx = \left[g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]^{k} \tag{2}$$

Even if the shocks are separate, it is possible that they will grow more effective in inflicting damage. This suggests that V k (t), the kth damage's distribution function, decreases with k = 1, 2,... for each t. A renewal process is a counting process in which the time until the first event happens has some distribution F, the time between the first and second events has the same distribution F regardless of the time of the first event, and so on. When something happens, we refer to it as a "renewal." The renewal procedure has also shown that

$$P(T > t) = \sum_{k=0}^{\infty} V_k(t)\, P(X_i < Y) \tag{3}$$

$$L(t) = 1 - S(t)$$

Taking Laplace Transformation of L(T), we get

$$= 1 - \left\{\sum [F_k(t) - F_{k+1}(t)]\left[g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]^{k}\right\}$$

$$l^*(S) = \frac{\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\right] f^*(S}{\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta} f^*(S)\right]}$$

$$= \frac{\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]\left(\frac{c}{c+s}\right)}{\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\left(\frac{c}{c+s}\right)\right]} \tag{4}$$

$$E(T) = \frac{d}{ds} l^*(S) \; given \; s = 0 = \frac{1}{c\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]}$$

$$E(T^2) = \frac{d^2}{ds^2} l^*(S) \; given \; s = 0 = \frac{2}{c^2\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]^2}$$

$$g^*(.) \sim \exp(\mu), g^*(\lambda) \sim exp\left(\frac{\mu}{\mu+\lambda}\right), \qquad g^*\left(\frac{x}{\sigma}\right)^{\beta} \sim exp\left(\frac{\mu}{\mu + \left(\frac{x}{\sigma}\right)^{\beta}}\right)$$

Then

$$E(T) = \frac{1}{c\left[1 - g^*\left(\frac{x}{\sigma}\right)^{\beta}\right]} = \frac{\mu\sigma^{\beta} + x^{\beta}}{c[\mu\sigma^{\beta} + x^{\beta} - \mu\sigma^{\beta}]} \tag{5}$$

$$E(T) = \frac{\mu\sigma^{\beta} + x^{\beta}}{cx^{\beta}} \tag{6}$$

$$E(T^2) = \frac{\left(\mu\sigma^{\beta} + x^{\beta}\right)^2}{c^2(x^{\beta})^2}$$

$$= \frac{\left(\mu\sigma^{\beta} + x^{\beta}\right)^2}{c^2x^{2\beta}}$$

$$V(T) = \frac{2\left(\mu\sigma^{\beta} + x^{\beta}\right)^2}{c^2x^{2\beta}} - \frac{\left(\mu\sigma^{\beta} + x^{\beta}\right)^2}{c^2x^{2\beta}} = \frac{\left(\mu\sigma^{\beta} + x^{\beta}\right)^2}{c^2x^{2\beta}} \tag{7}$$

Where,

μ – Stage I

σ – Stage II

β – Stage III

χ – Treatment after cell growth

c – Class interval

Figure: 1. Expected time for Stages and after treatment

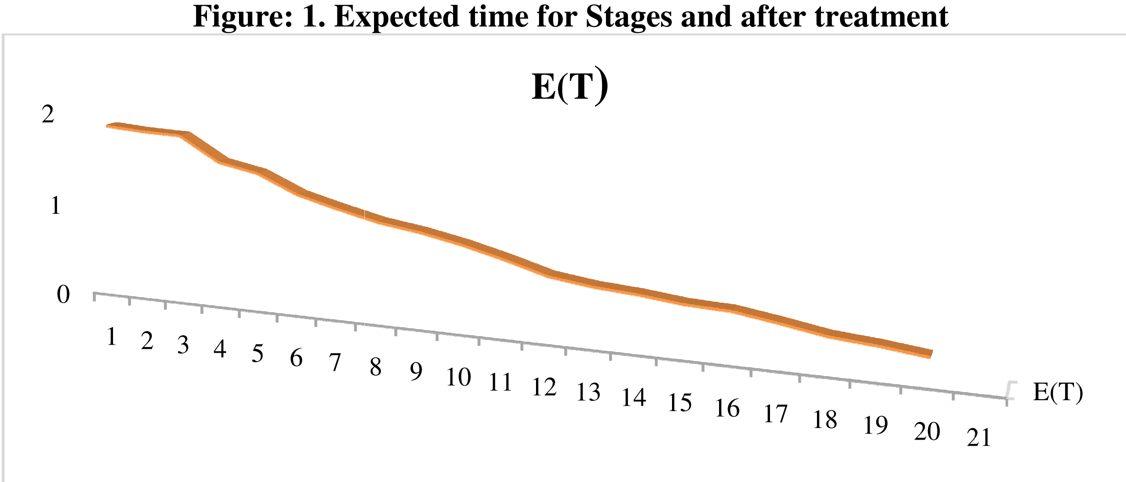
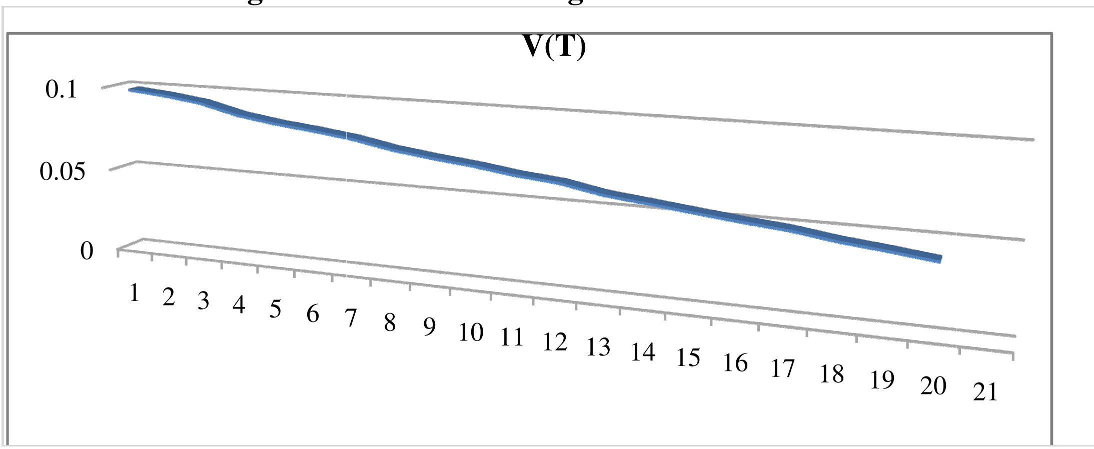


Figure 2 Variance for Stages and after treatment



3. CONCLUSION

When the inter-arrival time 'c', which follows an exponential distribution, is kept the same, it becomes a bigger and bigger number.Because of this, the value of the expected time E (T) for a cancer patient to cross the threshold is going down for all of the parameter values.When the value of the parameter increases, the expected time is decreasing. This is indicated in pictures. Same thing happens with the result of a cancer patient who had the variance V (T) that is shown in the pictur model supports us in identifying the essential relation of the developed CKD. The study works out that the misclassification of stages in process of CKD may take place in the behaviour or lack of prognostic considerations. Covariates like high blood pressure and diabetes may speed up the stages of CKD, which means that patients' lives are getting shorter.

REFERENCES

1. *Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C-y. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. New England Journal of Medicine. Vol.351, No.13, pp.1296–1305.*
2. *Grover G., Gadpayle AK, Swain PK, Deka B. (2013). A multistate Markov model based on CD4 cell count for HIV/AIDS patients on antiretroviral therapy (ART). Int J Stat Med Res. Vol.2, pp.144-151.*
3. *Hassan S. Bakouch et.al., (2012). An Exponentiated exponential binomial distribution with application, Statistics and probability letters, Vol.82, No.6, pp.1043-1218.*
4. *Turin TC, James MT, Jun M, Tonelli M, Coresh J, Manns BJ, et.al.,(2014). Short-term change in eGFR and risk of cardiovascular events. Journal of the American Heart Association. Vol.3, No.5.*
5. *Varma PP. (2015). Prevalence of chronic kidney disease in India-Where are we heading? Indian J Nephrol. Vol.25, pp.133-135.*
6. *XiaoxuWu, YongmeiLu,SenZhou, LifanChen,BingXu. (2016). Impact of climate change on human infectious diseases: Empirical evidence and human adaptation Environment International.Vol.86, pp.14-23.*

DOI No. : 10.5958/2394-9309.2022.00023.3
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

FACTORS AFFECTING THE IMPULSIVE BUYING BEHAVIOR OF CUSTOMERS - A STUDY WITH SPECIAL REFERENCE TO CUDDALORE CITY

Mr. V. Pushpanathan*, **Dr. G. Ravi****
*Ph. D Research Scholar, PG and Research Department of Commerce, Presidency College (Autonomous), Chennai-600006.
**Assistant Professor, PG & Research Dept. of Commerce, Presidency College (Autonomous), Chennai – 600006.

ABSTRACT
The markets are different and categorized by an increased competition, as well a continuous innovation in products and services available and a greater number of companies in the same market. Impulse buying is an important aspect of consumer's behaviour which comprises an interesting point for understanding the marketing efforts. Impulse buying is one of the current forms of buying. This form of buying is a divisively issue for researchers of consumers' behaviour not only due to its difficulty but also because it is observed among a wide range of product. In this paper researcher attempts to make an factors influencing the impulsive customers in the Cuddalore City.
Keywords: Chi Square value, Cuddalore, Impulsive Products, Perception level, Indian retail Market etc.

1. INTRODUCTION

Now a day's Indian retail market is constantly increasing, people are purchasing goods as there is a rise of income of common people as well as change in perceptions and preferences of consumers. The markets are different and categorized by an increased competition, as well a continuous innovation in products and services available and a greater number of companies in the same market. In this scenario it is essential to know the consumer well (Varadarajan, 2020).It is significant for the retail players to be able to understand the different factors affecting the extent in impulse buying behaviour. Impulse buying is an important aspect of consumer's behaviour which contains an interesting point for understanding the marketing efforts. Impulse buying takes place when person immediately purchases a product unplanned and without thinking. This form of buying is accidental because the individual has not been watching for a particular product and did not have the slightest plan of buying. Impulse buying is an important aspect of consumer's behaviour which comprises an interesting point for understanding the marketing efforts. Impulse buying is one of the current forms of buying. This form of buying is a divisively issue for researchers of consumers' behaviour not only due to its difficulty but also because it is observed among a wide range of product. Impulse buying is defined as a quick strong propensity to buy products which the customer had not planned to buy them and has bought them without deep observation.

An impulse buy is an unintended decision to buy a product or service, made just before a purchase. Impulse buying is unreflective in that the purchase is made without appealing in a great deal of evaluation. Individuals purchasing impulse is less likely to reflect the significances or to think carefully before making the purchase. Impulse buying has been measured a pervasive and characteristic phenomenon in the American lifestyle and has been receiving increasing attention from consumer researchers and theorists.

Impulse buying has been studied from several perspectives, namely: (i) rational processes; (ii) emotional resources; (iii) the cognitive currents arising from the theory of social judgment; (iv) persuasive communication; (v) and the effects of advertising on consumer behavior (Malter et al., 2020).

2. LITERATURE REVIEW

A literature review is a scholarly paper that presents the current knowledge including substantive findings as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources and do not report new or original experimental work. Early studies on impulse buying stemmed from managerial and retailer interests.

Verma Priyanka and Verma Rooble (2012) Titled on " An On-Field-Survey of the Impulse Buying Behaviour of Consumers in Consumer non Durable Sectors in the Retail Outlets in the City of Indore, India" is an attempt to find the impact of communication mix that effects customer impulse buying behaviour in non-durables in the city of Indore, a buzzing mercantile city of the state of Madhya Pradesh in India. The impact of various impulse buying factors like advertising, sales and promotions, personal selling, public relations, direct marketing on customer impulse buying behaviour has been analyzed.

JIYEON KIM(2000) his study entitled on "College students' apparel impulse buying behaviors in relation to visual merchandising" in this research examined the relationship between college students' apparel impulse buying behaviors and visual merchandising. This research provides information as to why visual merchandising should be considered an important component of a strategic marketing plan in support of sales increase and positive store/company image. This research also provides insights to retailers about types of visual merchandising that can influence consumers' impulse buying behaviors. The result of the present study proves that there is a pivotal relationship between college students' impulse buying behaviors and two type of visual merchandising practices: in store form/mannequin display and promotional signage.

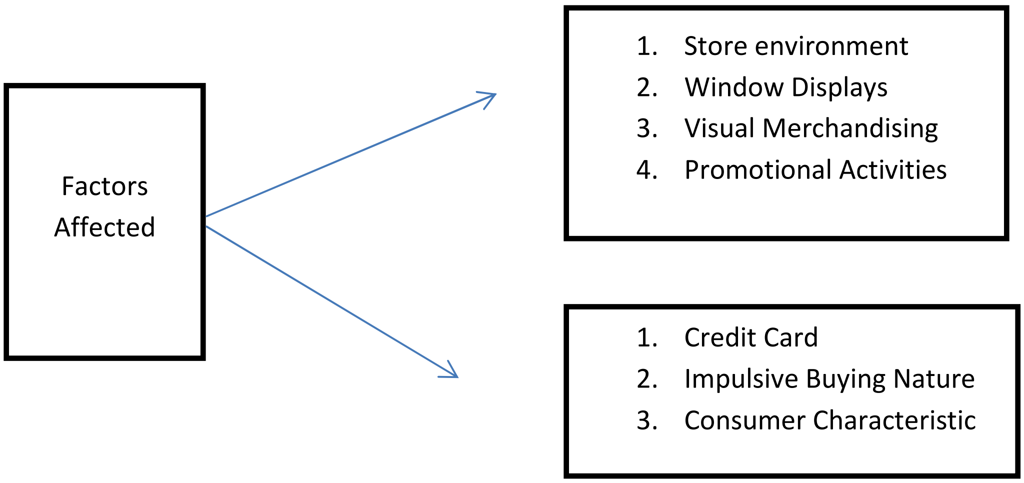
Cho, James, Ching, Gregory S. and Luong, Thai-Ha (2014)- his study entitled on **"**Impulse buying behavior of Vietnamese consumers in superstore setting" This paper was classify the factors affecting consumer impulse buying behaviors at superstores in Vietnam. This study attempts to analyze the impact of several variables extracted from internal, external, demographics, social perspectives on consumer impulse buying behavior. This consumer behavior is on a great rise due to pricing strategies, store characteristics, situational factors and promotional activities.

Maryam Sarikhani Khorrami, Mohammad Rahim Esfidani & Sajad Delavari(2015) in this paper focused for the aim of this study is to identify the effect of situational factors on impulse buying and compulsive buying. They using factors in Time pressure, available money, variety of selection, store environment, word of mouth, social norm, impulse buying and compulsive buying are variables that designed our model. Tools used in quantitative research with descriptive approach, where required data were gathered through questionnaires. The study found that available money and word of mouth have significant effect on impulse buying and impulse buying has significant effect on compulsive buying.

Factors affecting the Impulse Buying Behavior:

Most of the researchers make two categories to classify the factors which have an impact on impulse buying behavior. These categories are named 'internal factors' (shopper-related factors) and 'external factors' (environmental factors, marketer controlled or sensory stimuli emanating from the marketing systems) (Youn and Faber, 2000:179; Impulse buying behavior of the shopper influenced by number of factors which could be related to shopping environment, customer's personal traits, products' characteristics, the diverse demographic and sociocultural dimensions.

In 1987, Rook addressed that merchandising stimulus such as location of shelf and shelf space affect impulse buying. Also, consumers` demographics and lifestyle have impact on impulse buying. Impulse buying research increased and extended to investigations of how merchandising stimuli like retail shelf location and amount of shelf space influence impulse buying, and some studies determined the types of circumstances in which consumers buy things without prior planning and examined the relationships between consumers` demographic and lifestyle characteristics and their impulse buying susceptibility (Rook, 1987:190).



Internal Factors of Impulsive Buying:

Store environment:

The most of the customers are being exaggerated by internal environmental factors in Impulsive attitude of consumer. Marketers be a focus for the meditation of consumers for impulse buying through marketing stimuli external factors, when consumers "displaying stimulus as advertising inducement" (Rundh 2005). Impulse buying of consumers is generally formed by "the stimulus".

Window Displays:

A consumer's choice of a store is influenced by the physical attractiveness of a store. So window display is importance of relation to consumers' buying behavior has received minimal attention in the literature. Moreover, suggest that Widow Display associating with consumer's purchasing attitude has physical charisma and charm of a store influences customer choice of store (Alireza & Hasti 2011)."

Visual Merchandising:

Visual Merchandising means visual presentation, its aim is to communicate the retail and company fashion value, quality, building separate identity in consumers. The purpose of visual merchandising is **to attract, engage and motivate the customer towards making a purchase**, where it creates an impact on the consumer buying behavior.

Promotional Activities:

Promotional activities mean Sales promotion. It carries out a change in the requirement pattern of products and services. It helps to introduce new products or services in souk. Sales promotion plays a significant role to create and attract new customers. A sales promotional strategy is coupons, Buy-one-Get-one free price discount and free sample on customer's impulsive buying behavior.

External Factors of Impulse Buying:

External factors of impulse buying refer to marketing cues or stimuli that are placed and controlled by the marketer in attempt to lure consumers into purchase behavior (Youn and Faber, 2000).

Credit card:

The credit card is one of the importance factors of impulsive buying behavior. Easy access to credit cards eliminate the immediate need for money to buy something, cause consumer to overspending (Schor, 1998) and likely accelerate the development of impulse buying (Robert and Jones, 2001).Since impulse buying behavior may be accelerated by the credit card use (Roberts and Jones, 2001; Kim, 2001a; Kim, 2001b), a need exist to investigate the relationship between impulse buying behavior and credit card use.

Impulsive Buying Nature:

The nature of impulsive buying means an attitude of people is totally unreflective so the purchase cannot be made without judging the product. Everyone Impulsive purchasing that are little probable for deeming a fine or to believe carefully right earlier than doing purchase (Rook 1987). Impulse buying an unweights in buying that is composed of not appealing into contract for estimation. People just focus on instant satisfaction by reacting to insist on to purchase fairly than on resolving pre-existing troubles in looking up thing for filling set desire.

Consumer Characteristic:

Consumers characteristics are often used to predict how likely a group of people are to purchase a specific product. Consumer characters of people engaged in acquiring, using, and throwing away economic services and goods, inclusive of the decision-making procedures that come before and after such actions. Impulsive buyers have low levels of self-esteem, high levels of anxiety, depression and negative mood and a strong tendency to develop obsessive-compulsive disorders.

Research Objectives:

The Research Objectives are as following:

1. To investigate whether the demographic profile of gender, age and education the difference exists in impulsive buying behavior of consumers.
2. To study the factors affecting the impulsive buying behavior of consumers.

Hypothesis of the study

- There is no association between income level and perception of impulsive buying motives of consumers.
- Marital status does not influence the factors of impulsive buying behavior of consumers.

Limitations of the study

- Few customers are not given the accurate response to the interviewer.
- Some respondents were hesitant to give true response.
- The data was collected for two months only.
- The inference only given for the Cuddalore city particular with the selected shop only

3. RESEARCH METHODOLOGY

The study is based on the primary data collected from various super markets from the area of the city of Cuddalore with the help of structured questionnaire about consumer intentions and behaviour related to the analysis of the respondents' impulsive buying behaviours. 180 samples are collected from various super markets in cuddalore city using by convenience sampling method**.** Collecting data regarding the potential customers from the existing outlets of various super market. Data analysis has been done using MS-Excel software. The main limitation was the sample size, and the fact that the research was based on the consumers' sincerity in answers, without verifying their actual behaviour and actual purposes.

Data analysis and interpretation:

In this research data was analyzed in the form of descriptive statistics showed how samples were distributed in terms of demographic variables (gender, age, marital status, education, & income level) as well as factors influencing buying behavior of consumers and perception of impulsive buying behaviour by chi squared, variance and frequency in the form of tables.

Table 1 – Demographic Profile of The Impulsive Customers

S. No	Demographic Profile	Factors	No. of Respondence	Percentage
1	Gender	Male	74	41
		Female	106	59
2	Age(in years)	Below 25	17	9
		25 – 35	36	20
		35 – 45	64	36
		45 – 55	32	18
		Above 55	31	17
3	Marital status	Married	123	68
		Unmarried	57	32
4	Occupation	Self-employment	22	12
		Govt. Employee	46	26
		Non – Govt. Employee	79	44
		Farmer	33	18
5	Income level	Below 20000	20	11
		20000 – 30000	48	27
		30000 – 40000	69	38
		40000 – 50000	21	12
		Above 50000	22	12
6	Qualification	Below higher secondary	45	25
		UG	78	43
		PG	40	22
		Professional	17	10

Source: Primary Data.

Above table 1 shows that most of the impulsive customer 59 per cent are female , they are income earned people in the family in the research area as compared to women in the Cuddalore city. In this research 68 per cent of the retail customers are married. On a percentage basis, 35- 45 years of age groups fall on 36 per centage. 43 per cent of the customers qualifications only Under Graduate level in the study area. Out of 180 interviewees , 44 per cent of the customers are non-government employee and remaining fall under the other category. Out of 180 retail customers 38 per cent of them earn the monthly income of Rs.30,000 – Rs. 40,000 (i.e. Non-Government employee and Government employee).

Table 2- Factors Influencing Impulsive Buying Behavior

S. No	Factors	No. of Consumers	Percentage
1.	Elevation Scheme	45	25
2.	Cash Payment	22	12
3.	Convenience of time	30	17
4.	Debit/Credit cards/ATM	15	8
5.	Availability of many product	18	10
6.	Consumer Attitude	28	16
7.	Store Outline	22	12

Source: Primary data

Table 4.2 clearly mentioned factors influencing impulsive buying behaviour out of 180 respondents 25 per cent of the customers were influenced for the elevation scheme, 17 per cent of the retail customers are influencing the factors for Convenience of time and consumer attitude respectively followed by cash payment and the rest 12 per cent of the customers are influencing the cash payment and store outline.

Table 3 - Perception of Impulsive Buying Behavior

S. No	Perceptions	No. of Respondents	Percentage
1.	Rational	45	25
2.	Emotional	14	8
3.	Social judgment	29	16
4.	Persuasive communication	30	17
5.	Advertising	62	34
Total		**180**	**100**

Source: Primary data

Table 3 clearly enumerates that the perception of impulsive buying behaviour of the Cuddalore City consumers. Out of 180 respondents 34 per cent of the customers percepts for advertising; 25 per cent of the respondents are rational perception followed by 16 per cent of the customers aware about Social Judgement and Persuasive Communication respectively and the rest 8 per cent of the customers were awareness about emotional buying.

Verification of hypothesis

In order to prove this fact, the interviewer has framed the following null hypothesis that H1- “There is no association between income level and perception of impulsive buying motives of consumers”. To validate this promulgation, the researcher has used Chi-Square Technique.

Table 4 – Relationship between income level and perception of the impulsive buying motives of the consumers

Income level	Rational	Emotional	Social Judgment	Persuasive Communication	Advertising	Total
Below 20k	4	1	5	3	7	**20**
20k – 30k	13	4	9	8	14	**48**
30k – 40k	19	7	8	10	25	**69**
40k – 50k	3	2	2	4	10	**21**
Above 50k	6	0	5	5	6	**22**
Total	**45**	**14**	**29**	**30**	**62**	**180**

Source: Primary Data (Computed)

Calculated Chi square $\chi 2$= 9.805; p value = 26.296; the hypothesis is accepted.

As p value 26.296 is higher than α error (0.05%), hypothesis H0 is accepted, so that we can conclude that there is no association between income level and perception of impulsive buying motives of consumers. i.e. spontaneous purchases guided by the current emotions. As 69 respondents 30 to 40 thousands of income level having is more than the impulsive buying behavior. Second more than impulsive buying behavior of the above total respondents is 20 to 30 thousands of income level of consumers.

H2- Martial status do not influencing the factors of impulsive buying behavior of consumers.

Table 5 – Relationship between Marital status and Factors influencing impulsive consumer behavior

Factors	**Married**	**Unmarried**	**Total**
Elevation Scheme	38	7	45
Cash Payment	17	5	22
Convenience of time	25	5	30
Debit/Credit cards/ATM	10	5	15
Availability of many product	13	5	18
Consumer Attitude	10	18	28
Store Outline	10	12	22
Total	**123**	**57**	**180**

Source: Primary Data (Computed)

Calculated Chi square $\chi 2$= 29.411; p value = 12.592; the hypothesis is rejected.

In the chi squared analysis the p value is 12.592 which is lower than the alpha 0.05 (5% level of significance) level. So it was found in the study that Martials status do influencing the factors of buying behavior of consumers. The responses of these customers reveal that several factors, such as: Elevation Scheme, Convenience of time and Cash Payment are more influencing the marital status of impulsive buying behavior of consumers. And remaining factors do not much more of influencing the marital status of impulsive buying behavior.

Findings of the study

- It is lucid that 59 per cent of the respondents are Female in this study area.
- It is clear that 36 per cent of the respondents are fall under the age group of 35 to 45 years.
- It is found that 68 per cent of the sample force are married remaining unmarried.
- It is vivid that very few customers occupations are government and business.
- Most (25%) of the customers are influencing the buying the products only for elevation scheme.
- Majority (34%) of the consumers are awareness about the advertising for retail shops.

4. RECOMMENDATIONS AND CONCLUSION

This study found the new understanding of the situational factors and impulsive buying behavior in the witnessed in the Cuddalore city. Results of this study expressed by responses reveal that most situational factors to some extent prompt shoppers to buy on impulse. This study is to found that to investigate whether the demographic profile of gender, age and education the difference exists in impulsive buying behavior of consumers. Hypothesis compared to there is no association between income level and perception of impulsive buying motives of consumers. This study found the five percent level of significant hypothesis is accepted. So there is no association between income level and perception of impulsive buying motives of consumers. This study exposed that impulsive buying behavior does significantly depend on the respondents' marital status of number of household members The query which is stood here is how much the respondents, as consumers, are actually conscious of the impact of situational factors to be able to observe them and distinct those that do affect their buying behavior. For a more thorough description of separated segments, demographic data, i.e., socio-economic variables, were also used. Connecting these variables to the impulsive buying produced the results that might be expected to some extent. When the customers get products on a discounted price, they purchase more than they intended to buy. So, price discount enhances unplanned purchasing and it also helps the retailers to clear the stock. When the customers are offered price discount, they also incline to purchase other brand products also. So, it also helps in brand switching.

REFERENCE

1. *Alireza, K., and Hasti, Y., 2011. "Evaluating Effective Factors on Consumer Impulse Buying Behavior", Asian Journal of Business Management Studies 2 (4): 174-181.*
2. *Verma Priyanka1 and Verma Rooble- An On-Field-Survey of the Impulse Buying Behaviour of Consumers in Consumer non Durable Sectors in the Retail Outlets in the City of Indore, India" Research Journal of Management Sciences ISSN 2319–1171 Vol. 1(4), 1-5, November (2012) Res. J. Management Sci.*
3. *Rook, D. W. (1987), "The Buying Impulse", Journal of Consumer Research, Vol. 14 No.2, 189-197*
4. *Jiyeon kim "college students' apparel impulse buying behaviors in relation to visual merchandising" bfa, American Intercontinental University (2000)*
5. *Cho, James, Ching, Gregory S. and Luong, Thai-Ha-" Impulse buying behavior of Vietnamese consumers in supermarket setting", International Journal of Research Studies in Management, 2014 October, Volume 3 Number 2, 33-50*
6. *Maryam Sarikhani Khorrami, Mohammad Rahim Esfidani & Sajad Delavari The Effect of Situational Factors on Impulse Buying and Compulsive Buying: Clothing nternational Journal of Management, Accounting and Economics Vol. 2, No. 8, August, 2015*
7. *Seounmi Youn and Ronald J. Faber, 2000. Impulse Buying: Its Relation To Personality Traits and Cues", Advances in Consumer Research, 27: 179-185.*
8. *Malter M., Holbrook M., Kahn B., Parker J., Lehmann D. (2020). The past, present, and future of consumer research.* Market. Lett. *31, 137–149. 10.1007/s11002-020-09526-8 [PMC free article] [PubMed] [CrossRef] [Google Scholar]*
9. *Varadarajan R. (2020). Customer information resources advantage, marketing strategy and business performance: a market resources based view.* Indus. Market. Manag. *89, 89–97. 10.1016/j.indmarman.2020.03.003 [CrossRef] [Google Scholar]*
10. *Sanjay Pramanik , Rupali Pramanik(2018) Sales Promotion Strategies and Customers Impulsive Buying Behavior: A Study of Indore City, International Journal of Research, e-ISSN: 2348-6848 p-ISSN: 2348-795X.*

DOI No. : 10.5958/2394-9309.2022.00024.5
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A STUDY ON LOAD BALANCING TECHNIQUES AND OPTIMIZED LAYOUTS IN CLOUD ANALYTICS

R. Nathiya*, Dr. SK. Piramu Preethika**
*Research Scholar, Department of IT, School of Computing Sciences, VISTAS, Chennai, Tamilnadu. India
**Assistant Professor, Department of IT, School of Computing Sciences, VISTAS, Chennai, Tamilnadu. India

ABSTRACT
In cloud system, allocating or scheduling of user tasks is considered to be an NP-hard optimization problem. As per the cloud environment, tasks of every user are assigned to denote balanced overload or under load of work. The process of task scheduling with the help of load balance might be done on independent or dependent work with virtual machines (VMs) is considered to be more important aspect as well. Load balancing is the technical process to detect and balance the loads of underloaded or overloaded nodes in the cloud task. In addition to this, optimized data layout focusing on maximize block skipping process in line with block assignment strategy to work with defined workload and dataset in the cloud environment. There are some approaches on load balancing mechanism in cloud computing in order to improve efficiency over various performance parameters. This study presents a detailed overview of load balancing and optimized layouts in cloud analytics.

Key Words: Cloud analytics, load balance, task scheduling, virtual machine, optimized data layout (MTO-Multiple table optimizer)

1. INTRODUCTION

Nowadays, all businesses are reportedly working on analytics that help efficient performance on multiple historical and operational data in order to attain profit in competitive advantages and their business insights. Therefore, such analytical reports are running under cloud-based work analytics services like Microsoft Azure Synapse, Google BigQuery, Snowflake and Amazon Redshift. Such services are processed data in columnar formats with compressed and retained in large blocks [1]. Each part consists of thousands or millions of data records in it. From the cloud storage, disk I/O usage is considered to be one of the dominant costs in query processing problem for these availed services. In order to reduce the query execution time and lower the costs of I/O, these data services often need to balance per-block metadata like zone maps, which helps to skip blocks by assessing not necessary points to the query. Moreover, the potency of block skipping process through zone maps is particularly dependent on how the data records are scheduled to the blocks as well. Recent research proposed the instance-optimized data layouts which focusing on maximize block skipping process in line with block assignment strategy in a specialized way to work with defined workload and dataset in the cloud environment. Hence, already existing models owing to optimize the single table layout.

Modern cloud-based data analytics services perform various techniques to lower the data access and data movement while maintaining larger volumes of data process in efficient manner [2]. Therefore, a standard unique technique is required for columnar storage in order to avoid accessing on columns which does not have any right to make hard copies of any part or digitalization process. Furthermore, this work used for personal or classroom purpose could be granted and no fee provided for those obtaining copies or not to make for profit distribution or any commercial advantage as well as that copies get warning notices as the full citation on the first page. In zone maps, it can skip blocks 1 and 2 in the first query but it cannot able to skip any of the blocks in the second query session which are relevant to the query. Amazon S3 consists of data which are often retained in remote cloud storage and its query is processed by accessing “compute nodes” at execution. These storage systems records data into large blocks consisting of hundreds or thousands or millions of records to maximize their compression ratios. During query processing, a block is considered to be smallest unit of I/O operations from cloud computing storage and is used for minimize I/O and to maximize throughput operations at per second. Per-block metadata are used to skip blocks and also often cache in memory for resist blocks accessing particular not relevant to the query processing. Zone maps are most known form of per-block metadata that reserve maximum and minimum value for each column in a data block. It is effective and cheap to maintain. Zone maps effectiveness is accessed at block skipping process highly depends on how much data layout is allocated to each block. Each table is sorted by the systems in default access by certain sort column option and will find and place some actual chunks of records in a same block [3]. With the help of basic blocking scheme, filtered queries are sorted in one column undergo skip block based on zone maps whereas other columns in filters do not have skipping opportunity. Moreover, Zorder is a multi-dimensional sorting technique not used often in practice due to its simplicity. In columns with Z-order defined very effective in block skipping process and select manual relative orders in careful manner which permits poor tuning resulted degraded performance. The existing shortcomings of data layout techniques overcome by the newly proposed idea of instance-optimized data layouts are available now. These approaches signed as “learn” which is a specialized in a blocking scheme especially to attain high performance in block skipping. This study highlights the load balancing and optimized layouts in cloud analytics.

Literature review:

2. LOAD BALANCING

In cloud computing multi-objective system consist of load balancing algorithm which leads create a NP-complete problem. It comprises some main objectives of 15 energy saving efficiency, minimizing make span scheduling problems, maximized network throughput and so on. Hence, researchers are launched many techniques in heuristic approaches such as sub-optimal search algorithmsto find out a sub-optimal decoding solution for performing load balancing algorithms in the cloud environment. The loan balancer are required to maintain stability for the load of the system. As in cloud computing, theresearches on the heuristic technique in two types of load balancing system such as static and also dynamic strategy are identified [4].

Static strategies:

Usually, the cloud computing strategies of the static load are categorized under two types of assumptions. Initially, the first is that arrival time of initial task and also the second one is that checking of physical machines availability [5]. After evaluation of each and every task, the information resource is to be updated on regular basis. Some heuristics static strategies are MET, OLB, GA, MCT, TABU, A∗40 algorithm, Min-Max, Switching Algorithm and Min-Min.

Dynamic strategies

In cloud computing environment, it is considered to be important strategy due to distribution of load balance among physical machines at the execution-time. Hence, task's arrival time may be unusual, as well as depending on the type of input task, the creation of virtual machines is determined. Such dynamic algorithms based on heuristic model are used for load balancing and of two categories: Off-line mode also called Batch mode and immediate mode or called as On-line mode. At some predefined moments only, task is distributed in Off-line mode heuristics. It is used to fix actual time for executing multiple tasks at a time. In batch mode, these heuristics technique are presented as Min-min, Max-min, and Sufferage algorithm. Thus, in immediate mode or on-line mode, the user tasks are overlapped with computing node where it entered as soon as the scheduler. Each task are carried out for scheduling work at once in a time which resulted that remains unchanged one. Heuristics design for immediate (On-line) modes is presented as MET, OLB, MCT, and SA. OLB. In a cloud environment, OLB(Opportunistic load balancing)heuristic model utilized both of static strategy as well as dynamic (immediate mode) strategy [5].

Hardware vs. Software LB

In traditional method, LB solutions worked on proprietary hardware encountered in a data center. It requires team work with highly trained IT personnel in order to maintain, tune, and install the system. Hence, large organization along with big IT budgetscan only utilizes the benefits of high performance and reliability. Hardware-based solutions have undergone serious drawback in the age of cloud computing. They did not support cloud load balancing system due to the vendors of cloud infrastructure did not allow any proprietary or customer hardware in their work station.

In a good way, software based load balancers could able to pay out reliability benefits and good performance at a lower cost rate of hardware based solutions. It is purely based on commodity hardware which they run and they are considered to be affordable for smaller companies. They are also ideal for LB as even they can carry out the cloud performance as like with other software application [6].

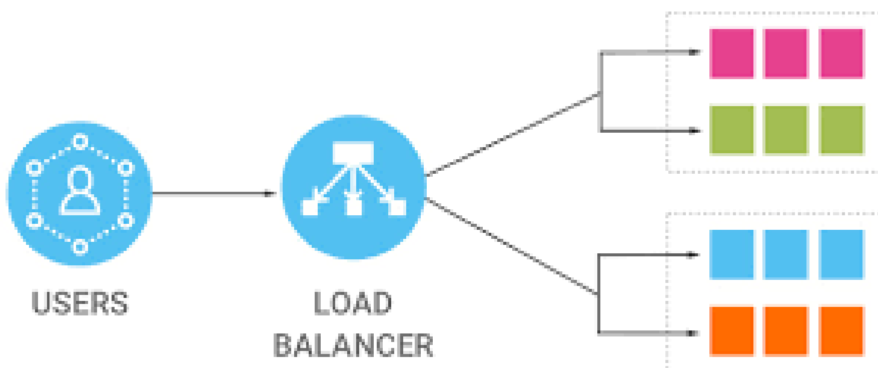
Benefits of Cloud Load Balancing

The global character of the cloud itself proves their benefits in load balancing particularly from scalable point.

- The speed and ease of scaling system in the cloud demonstrates that can handle all type of traffic spikes without any degraded image in order to place cloud load balancer at a group of application. This can attain the level of demand in quick auto scale reaction.
- The derived ability to place an application at multiple cloud hubs all over the world which enhance the reliability power as well [7].

3. OPTIMIZED DATA LAYOUT

In order to rule out computation and load balance in computational resources, Cloud base analytics usually distribute data throughout the partitions or multiple nodes.Data is considered to be distributed purely based on using range, ingestion time, round-robin distribution or hash schemes. Automatic design advisors typically use data mining and what-if analysesfor auto-tune the partitioning scheme and physical design.Some approaches in automated tasks are expert in terms of analytic and transactional workloads. MTO (Multi Table Optimizer) is distributed on these schemes within each partition or node [8].



Instance-Optimized Databases

Recent research trend has been found in regards to the instance-optimized database components and systems. In traditional systems, the design decisions often done by heuristics or manual tuning and sometimes machine learning are also used for the instance-optimized goal system also accomplish a particular case with algorithms and database components automatically.Qd-tree and MTO used as framework for instance-optimized data layouts [9].

Sideways Information Passing

To get predicate information from joins, column equivalence, data-induced predicates,and magic-set rewritingare used just similar to MTO database. The performance of these methods is based on data layout and it does not help skip blocks. During execution, MTO based join-induced predicate system explicitly used to frame a data layout which maximize block skipping opportunities [10]. Hence, sideways information passed between two connected sub-expressions form or tables are in the form of semi-join reduction in order to speed up joins and skip blocks during the execution. MTO also do sideways information in multiple tables for better joint layout during offline optimization. Other auxiliary data structures also cache needful data about joining tables which included join indexes, materialized views and join zone maps as well. MTO is not replicate any of database where these data structures utilize extra storage space as well as require maintenance overhead.

4. DISCUSSION

Load balance is an important in the cloud computing process which helps for business partners in process of easy handling to customers as well as end-users of cloudbased applications. It is considered to be advantageous for cloud computation where it could be working under a single server to carry out tremendous workloads [11]. It enhances raised service availability and crucially responded in times for some business operations which SLAs could permit. In absence of LB, new spinning virtual servers are never to be access or to adapt the incoming traffic in an aligned fashion. Apart from few virtual servers, it also left to access for zero traffic even though others could not do without LB. LB is also used to find some unobtainable servers as well as redirect traffic with those which are in operating as well.

Many research works portrays the load balancing techniques are operated under four categories such as geographical distributions, natural phenomenon based LB, general LB, and scheduling Network-aware task LB [12]. The geographical distribution channel of nodes are considered to be very important in line with commercial largescale applications like Facebook and Twitter. Geographic distribution is relatively considered as a resolution based on relocation of VMs, the digital deployment and/or executing activities or the digital deployment in geographically dispersed data centers in which deadlines assigned in systems for virtual machines or SLAs activities reached in order to reduce the operating expense costs for cloud environment.

Some kind of LB techniques are carried out by the general LB such asrandomized algorithm, round robin, threshold algorithm, OLB, min-min, max-min,OLB + LBMM, central LB Strategy for virtual machines, spread equal current execution algorithm, throttled LB, join idle queue and stochastic hill climbing. It is quick and efficient technique. It typically could not find connected servers which lead to incompatible resource distribution. The actual system state is quite taken into consideration for decision making process and this kind of measure make a biggest problem with the techniques. The natural phenomena-based LB has followed some techniques like genetic algorithm, ant colony algorithm, honey bee foraging, hybrid, artificial bee colony algorithm, osmosis LB algorithm, ant colony & complex network LB, bee colony optimization algorithm, and LB honey bee foraging. This process is influenced by biological behavior or natural phenomena [13].

The network-aware task scheduling LB consists of some type of LB techniques such as task scheduling strategy based on LB,shortest job scheduling LB algorithm,biased random sampling and active clustering [14].By improving the overall performance of the system and better resource utilization leads to achieve good load balancing process. In next generation, cloud dynamics launch multiservice concepts by which each server clusters could solely execute a unique multimedia task as well as each client could request a special multimedia carrier in a particular time. Load balancer is used to achieve the diligent loan distribution where different locations of task are obtained and then it will be distributed to the concern data center. Alongside with prolific development and innovation in the cloud arena devices and sensors are bringing people to relate closer computing. In upcoming generation, technology with Internet of Things (IoT) will play vital role that billions of smart objects could able to talk others diligently inn order to make human lives as more comfortable. Thus, IoT is wireless sensor community (WSN) device, and also Zigbee is one of the most leading WSN protocols. However, Zigbee AODV does not involve in stable load distribution mechanism to handle bursty traffic.

Load balancing technique collaborates with data processing system which consist of embedded sensors into the infrastructure take place of communication, transportation, buildings, various utilities, healthcare and sensors on user devices, home equipment and wearableswould be a result of the upcoming Internet of Things (IoT). Load balancing cloud technique along with IoTprovide better improvement over the accuracy and efficiency rate as well as reduces human intervention and accessing data effectively. The resource involved in IoT “things” in the IoT context where the data processing is obtained from biometric sensors, microchips, sensors on mobile phones and electrical. Thus, key venture is considered to be end-to-end safety in network system which has wifi networks, sensor networks, RFID devices, edge nodes, cloud information centers, public and personal clouds. These are required to be combined one in order to accomplish IoT systems [15]. This data processing technique involved to get impenetrable reprogrammable protocols that allow the authentication response by triggering a network characteristic and also trying to stop malicious installations. Here, we could try to use load balancing algorithm in the form of BYOD movement, which acts as end-users to derive their own information into syncing of non-public cloud services of streaming, and storage. Therefore, the traditional physical sensors would be big issues always, in future IoT devices enable the strong integration of physical sensors, human sensors and people-centric sensors will have been promoted. BDaaS is considered as a new financial model not only a new technology but it is a new data aid usage pattern. LB is played a specific role of hardware gateway which is using distinct protocol and software. Generally, LB technologies surrounded in two categories of level such as world and local. Global dealt to clear up the troubles of LB in one area between exclusive architectures in multiple clusters.

5. CONCLUSIONS

The study highlights load balancing and optimized data layout needed for cloud analytics. In the cloud based data analytics services, the query processing is applicable for managing large data blocks from cloud storage. The effectiveness is highly based on how the records are assigned to blocks. An effective approaches for optimizing data layouts targets a single table and performance affects using join based queries. In order to resolve the data layout issue, MTO technique is used for optimizing the blocking strategy for all tables for query workload. This technique generates 93% and 75% reduction in block accessed and end-end query times on commercial cloud based data analytics service.

REFERENCES

1. *H. Geng, "INTERNET OF THINGS AND DATA ANALYTICS IN THE CLOUD WITH INNOVATION AND SUSTAINABILITY," in Internet of Things and Data Analytics Handbook, John Wiley & Sons, Inc., 2016, pp. 1–28.*
2. *"Load Balancing in Cloud Computing," International Journal of Recent Trends in Engineering and Research, no. 3, pp. 118–125, Mar. 2018, doi: 10.23883/ijrter.2018.4105.vqpyq.*
3. *D. Anoshin, D. Shirokov, and D. Strok, "Getting Started with Cloud Analytics," in Jumpstart Snowflake, Apress, 2019, pp. 1–17.*
4. *N. K. Sehgal and P. C. P. Bhatt, "Analytics in the Cloud," in Cloud Computing, Springer International Publishing, 2018, pp. 159–169.*
5. *W. Smith and H. Sun, "Analytics Cloud Builder," in Pro Salesforce Analytics Cloud, Apress, 2016, pp. 39–75.*
6. *A. Harraz and M. Zbakh, "Cloud Load Balancing and Reinforcement Learning," in Advances in Business Information Systems and Analytics, IGI Global, 2018, pp. 266–291.*
7. *P. Membrey, D. Hows, and E. Plugge, "Load Balancing in the Cloud," in Practical Load Balancing, Apress, 2012, pp. 211–224.*
8. *R. Kaur and N. S. Ghumman, "Task-Based Load Balancing Algorithm by Efficient Utilization of VMs in Cloud Computing," in Advances in Intelligent Systems and Computing, Springer Singapore, 2017, pp. 55–61.*
9. *V. Kunwar, N. Agarwal, A. Rana, and J. P. Pandey, "Load Balancing in Cloud—A Systematic Review," in Advances in Intelligent Systems and Computing, Springer Singapore, 2017, pp. 583–593.*
10. *S. Afzal and G. Kavitha, "Load balancing in cloud computing – A hierarchical taxonomical classification," Journal of Cloud Computing, no. 1, Dec. 2019, doi: 10.1186/s13677-019-0146-7.AlRawas and Valeo, "ISSUES WITH FLASH FLOOD MODELING IN THE CAPITAL REGION OF SULTANATE OF OMAN", JOUR, vol. 1, no. 1, 2008. [Accessed 26 March 2021].*
11. *Manos Athanassoulis, Kenneth S. Bøgh, and Stratos Idreos. 2019. Optimal Column Layout for Hybrid Workloads. Proc. VLDB Endow. 12, 13 (Sept. 2019), 2393–2407.*
12. *Benjamin Hilprecht, Carsten Binnig, and Uwe Röhm. 2020. Learning a Partitioning Advisor for Cloud Databases. In Proceedings of the 2020 ACM SIGMOD International Conference on*

Management of Data (Portland, OR, USA) (SIGMOD '20). Association for Computing Machinery, New York, NY, USA, 143–157.

13. *Hongzi Mao, Malte Schwarzkopf, Shaileshh Bojja Venkatakrishnan, Zili Meng, and Mohammad Alizadeh. 2019. Learning Scheduling Algorithms for Data Processing Clusters. In Proceedings of the ACM Special Interest Group on Data Communication (Beijing, China) (SIGCOMM '19). Association for Computing Machinery, New York, NY, USA, 270–288.*
14. *Anil Shanbhag, Alekh Jindal, Samuel Madden, Jorge Quiane, and Aaron J. Elmore. 2017. A Robust Partitioning Scheme for Ad-Hoc Query Workloads. In Proceedings of the 2017 Symposium on Cloud Computing (Santa Clara, California) (SoCC '17). Association for Computing Machinery, New York, NY, USA, 229–241.*
15. *Zongheng Yang, Badrish Chandramouli, Chi Wang, Johannes Gehrke, Yinan Li, Umar Farooq Minhas, Per-Åke Larson, Donald Kossmann, and Rajeev Acharya. 2020. Qd-Tree: Learning Data Layouts for Big Data Analytics. In Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data (Portland, OR, USA) (SIGMOD '20). Association for Computing Machinery, New York, NY, USA, 193–208. https://doi.org/10.1145/3318464.3389770.*

DOI No. : 10.5958/2394-9309.2022.00025.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

DESIGNING OF QUICK SWITCHING SAMPLING SYSTEM FOR PERCENTILE LIFE UNDER LOG-LOGISTIC DISTRIBUTION

V Kaviyarasu*, Anila P**
*Assistant Professor, Department of Statistics, Bharathiar University, Coimbatore, TN
**Department of Statistics, Cochin University of Science and Technology, Kerala
E-mail :

ABSTRACT
Quality and reliability are two important criteria for product acceptance, there are various methods of inspection in quality control for improving the quality of products, reliability acceptance sampling plans plays a vital role in manufacturing industries to ensure defect less products are marketed which saves the inspection time and cost. In the article, a new reliability sampling system is proposedwhich is referred as Quick Switching Sampling (QSS) of type-1whichensure to study the percentile lifeof the products when the lifetime of items follows the Log-logistic distribution. The principal objective of this paper is to design an optimal sampling system using the planparameters of the proposed distribution under single sampling plan as reference plan. The operating characteristic curves and its associated probability values are developed for sampling system. The minimal sample size are determinedfor the optimal parameters of the QSS under two points on the OC curve approach. Tables are constructed to select the optimal parameters of the proposed system given with numerical illustrations. A real time example are presented to see the effectiveness and its implementation of proposed Quick switching system.

Keywords : Reliability, Quick switching System, Acceptance sampling plans,OC curve.

1. INTRODUCTION

Lifetime is a significant quality characteristic of any product to achieve its desirable purpose, here the sampling plans are used to determine the acceptability of a product, respecting its life time are known as reliability sampling plan. A common constraint in life testing is the duration of total time consumed on the life test. It will be really time consuming to wait all items fail when the lifetime of a product is high. Therefore a need arises to study about the truncated life test is developed and implemented. A life test sampling plan is a procedure which postulates the number of units of product from a lot which is to be tested and the criterion for determining acceptability of the lot is determined.

Acceptance sampling plans are developed based on the mean or medianlife are broadly studied under truncated life test which may not serve detailed analysis towards the strength or braking stress of the products required for manufacturing or designing engineer consideration. Therefore the manufacturer may pay more attention to study the percentile life of an item than the mean life of an item. The percentile life provides more detailed information than the mean life of a life testing experiments. When the life distribution is symmetric the 50th percentile or median is equal to the mean life, 25th percentile is considered as the first quartile deviation which will give the first quartile of failure of item. Then the 10th percentile is considered as the deciles which willlead an item initial failure.

Even though a number of sampling plans are available in the literature, special purpose sampling plans such as chain sampling plan, skip-lot sampling plan, etc., have taken an important place in the literature of acceptance sampling. Quick switching system is one of the important special purpose plans. Generally, special purpose plans can be applied when the lots are coming serially in the order of production for inspection like the production of cloth, float glass etc.In this article, it is propose to study QSS-1, for assuring product lifetime under Log-logistic distribution. In QSS-1 the procedure start with normal inspection and continue until a rejection occurs.If rejection occurs switch to tightened inspection and continue the tightened inspection until the lot is accepted. If a lot is accepted in tightened inspection, switch to normal inspection due to instantaneously switching between normal and tightened this system is referred as Quick Switching System.

Many authors studied on log-logistic distribution as life time distribution for various plan parameters including Shah and Dave(1963), Tadikamalla and Johnson (1982) and O'Quigley and Struthers (1982)whogave the survival models for the log-logistic and logistic distributions. The moments of order statistics of the distribution were studied, and linear unbiased estimator of the log-logistic distribution was given by Balakrishnan and Malik (1987). Kantam et al. (2001) introduced the log-logistic distribution in acceptance sampling plans based on truncated life tests. They considered the single sampling plan as a single-point approach using the median life r as the quality parameter of the submitted product.

2. CONDITIONS FOR APPLICATION OF QSS-1

Following are the conditions of Quick Switching System-1

- The production is steady, so that results on current and preceding lots are largely suggestive of a continuous process.
- Lots are submitted significantly in the order of production.

This QSS is described by five parameters namely sample size n, normal acceptance numbers $c_N$and tightened acceptance number $\boldsymbol{c_T}$,and normal probability of acceptance $P_N$and tightened probability of acceptance and P_T. Romboski (1969) has derived the OC function of QSS-1 (n; c_N, c_T) and the OC function of QSS-1 (n; c_N, c_T) is given by,

$$P_a(p) = \frac{P_T}{1-P_N+P_T} (1)$$

Where P_N and P_T.can be obtained from below equations

$$\mathrm{P_N(p)}=\sum_{\mathrm{d=0}}^{\mathrm{c_N}} \binom{n}{d} p^d (1-p)^{n-d} (2)$$

$$\mathrm{P_T(p)}=\sum_{\mathrm{d=0}}^{\mathrm{c_T}} \binom{n}{d} p^d (1-p)^{n-d} (3)$$

3. OPERATING PROCEDURE OF QSS UNDER TRUNCATED TEST

The following operating procedure can be executed in order to implement the proposed QSS under truncated life test studies.

Step 1: Select a random sample of size n units from the submitted lot during normal inspection and put them on life test. Count the number of defects in the sample as $d_N$before the specified timet_0

Step 2: Accept the lot if $\boldsymbol{d_N} \leq \boldsymbol{c_N}$or if $\boldsymbol{d_N} > c_N$,reject the lot and switch to tightened inspection.
Step 3: Under tightened inspection,select a random sample of size *n* units. Count the number of defects in the sample before the specified timet_0, denote it as d_T
Step 4: Accept the lot if $\boldsymbol{d_T} \leq \boldsymbol{c_T}$or if $\boldsymbol{d_T} > c_T$reject the lot and go to step 1and the process will terminate when the continuous production is stopped.
where n = sample size,
c_N = acceptance number of the normal inspection
c_T = acceptance number of the tightened inspection($c_N > c_T$)

4. CONSTRUCTION OF THE PLAN

Suppose that the life-time‘t’ of a product follows a Log-Logistic distribution. The probability that the product fails before the experiment time t_0 is given by

$$p = F(t_0) \qquad (4)$$

Log-Logistic Distribution has a great application in survival analysis and reliability. The Log-Logistic distribution is the probability distribution of a random variable whose logarithm has a Logistic distribution. The Probability Density Function (PDF) is given by,

$$f(x)=\frac{\left(\frac{\gamma}{\sigma}\right)\left(\frac{t}{\sigma}\right)^{\gamma-1}}{\left[1+\left(\frac{t}{\sigma}\right)^{\gamma}\right]^2} \qquad t\geq 0,\ \gamma > 0,\ \sigma > 0 \qquad (5)$$

and the Cumulative Distribution Function (CDF) of the Log-Logistic distribution is given by

$$F(x)=\frac{\left(\frac{t}{\sigma}\right)^{\gamma}}{1+\left(\frac{t}{\sigma}\right)^{\gamma}} \qquad t\geq 0,\ \gamma > 0,\ \sigma > 0 \qquad (6)$$

where γ is known shape parameter and σ is unknown scale parameter.
It is important to note that the CDF depends only on *t/σ*, since the shape parameter is known. The q^{th} percentile of the Log-logistic distribution is given as

$$\theta_q = \sigma\left(\frac{q}{1-q}\right)^{\frac{1}{\gamma}} (7)$$

The probability that the product fails before the experiment time $t_0$under the Log-logistic Distribution is given as

$$p=\frac{\left(\frac{t_0}{\sigma}\right)^{\gamma}}{1+\left(\frac{t_0}{\sigma}\right)^{\gamma}} \qquad (8)$$

According to Aslam and Jun (2009), one can write the experiment time as a multiple of the specified q^{th} percentile lifeθ_0. That is, $t_0 = a\theta_0$ for a constant experiment termination ratio ‘a’. Therefore equation (8) can be rewritten as

$$p=\frac{a^{\gamma}\left(\frac{q}{1-q}\right)\left(\frac{\theta_0}{\theta_q}\right)^{\gamma}}{1+a^{\gamma}\left(\frac{q}{1-q}\right)\left(\frac{\theta_0}{\theta_q}\right)^{\gamma}} \qquad (9)$$

5. DESIGNING THE SAMPLING PLAN

In this paper the designing method QSS-1 is proposed with the objective of minimizing the ASN, the significance of such optimization eventually eases the inspection cost and time. The ASN of proposed QSS is its sample size n. Therefore this paperproceeds with two points on the OC curve approach are used to design the QSS-1with the intention of minimizing the ASN at both AQL and LQL. The optimal parameters are determined such that both the producer and consumer risks are satisfied with minimum sample size using the optimization problem. With the intention of determine the optimal parameters, the following optimization problem can be used:

$$\text{Minimize ASN (p)=n}$$
$$\text{Subject to} P_a(p_1) \geq 1-\alpha$$
$$P_a(p_2) \leq \beta$$
$$\text{n} > 1, c_N > c_T \geq 0$$

where $p_1$is the quality level corresponding to the producer's risk and $p_2$is quality level corresponding to the consumer's risk also $P_a(p_1)$ is the acceptance probability of the lot under at $p_1$and $P_a(p_2)$ is the acceptance probability of the lot under at $p_{2.}$, which is given by,

$$P_a(p_1) \ = \frac{\sum_{\text{d=0}}^{\text{cT}}\binom{n}{d}{p_1}^d(1-p_1)^{n-d}}{1-\sum_{\text{d=0}}^{\text{cN}}\binom{n}{d}{p_1}^d(1-p_1)^{n-d}+\sum_{\text{d=0}}^{\text{cT}}\binom{n}{d}{p_1}^d(1-p_1)^{n-d}} \qquad (10)$$

$$P_a(p_2) \ = \frac{\sum_{\text{d=0}}^{\text{cT}}\binom{n}{d}{p_2}^d(1-p_2)^{n-d}}{1-\sum_{\text{d=0}}^{\text{cN}}\binom{n}{d}{p_2}^d(1-p_2)^{n-d}+\sum_{\text{d=0}}^{\text{cT}}\binom{n}{d}{p_2}^d(1-p_2)^{n-d}} \qquad (11)$$

the percentile ratio is described by $\frac{\theta_q}{\theta_0}$=2, 2.5, 3, 3.5, 4 at producer's risk and that ratio is taken as 1 at consumer's risk.

It is assumed that the lot size is sufficiently large so that the binomial model may be applied to determine the probability of acceptance. Stephens (2001) referred for more information on the use of binomial distribution. Accordingly the probabilities of acceptance for single acceptance sampling truncated life test plan under binomial model are given by

$$L(\text{p})=\sum_{\text{d=0}}^{c}\binom{n}{d}\, p^d(1-p)^{n-d}(2.12)$$

In this study, we determine the optimal parameters of the proposed Quick Switching System-1 for assuring 20$^{\text{th}}$ and 50$^{\text{th}}$ product lifetime percentile under Log-logistic distribution .For determining the optimal parameters, various values of shape parameters are considered as 1,2,3 ($\gamma = 1,2,3$). The producer's risk is fixed to be $\alpha = 0.05$, the consumer risks are taken as β=0.25, 0.10, 0.05, 0.01 and the experiment test termination ratio is considered as following two cases a=0.5 and a=1.0.

Table:1 Optimal parameters of the QSS-1 for 20th percentile under Log-logistic distribution with γ =1

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	30	5	0	0.9573	0.2115	19	6	0	0.9715	0.1757
	2.5	26	4	0	0.9758	0.2313	15	4	0	0.9655	0.1764
	3	22	3	0	0.9748	0.2513	12	3	0	0.9733	0.2507
	3.5	19	2	0	0.9513	0.2311	10	2	0	0.9527	0.25
	4	19	2	0	0.9674	0.2311	10	2	0	0.9683	0.25
0.10	2	38	6	0	0.9423	0.0848	20	6	0	0.9561	0.1174
	2.5	34	5	0	0.9746	0.0968	19	5	0	0.968	0.0812
	3	31	4	0	0.9748	0.0916	17	4	0	0.9712	0.0852
	3.5	28	3	0	0.9616	0.0889	15	3	0	0.9598	0.0909
	4	28	3	0	0.9777	0.0889	15	3	0	0.9765	0.0909
0.05	2	44	7	0	0.9484	0.0488	24	7	0	0.9472	0.0503
	2.5	37	5	0	0.9559	0.0542	23	6	0	0.9705	0.0356
	3	34	4	0	0.9585	0.0531	21	5	0	0.9785	0.0384
	3.5	34	4	0	0.981	0.0531	19	4	0	0.9749	0.0422
	4	32	3	0	0.9604	0.0457	17	3	0	0.9587	0.0475
0.01	2	61	10	0	0.9692	0.0105	34	10	0	0.9584	0.0081
	2.5	52	7	0	0.9629	0.0100	28	7	0	0.962	0.0105
	3	49	6	0	0.9768	0.0103	27	6	0	0.9699	0.0084
	3.5	47	5	0	0.9728	0.0092	25	5	0	0.9719	0.0098
	4	44	4	0	0.9613	0.0101	24	4	0	0.9523	0.0087

Table 2 Optimal parameters of the QSS-1 for 20th percentile under Log-logistic distribution with γ =2

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	36	2	0	0.9698	0.2405	10	2	0	0.9683	0.2499
	2.5	29	1	0	0.9574	0.2507	10	2	0	0.9918	0.2499
	3	29	1	0	0.9796	0.2507	9	1	0	0.9712	0.1923
	3.5	29	1	0	0.9891	0.2507	9	1	0	0.9845	0.1923
	4	29	1	0	0.9936	0.2507	9	1	0	0.991	0.1923
0.10	2	53	3	0	0.9799	0.0958	15	3	0	0.9765	0.0909
	2.5	47	2	0	0.9822	0.0988	13	2	0	0.9801	0.0994
	3	42	1	0	0.9564	0.0987	12	1	0	0.947	0.0866
	3.5	42	1	0	0.9765	0.0987	12	1	0	0.9714	0.0866
	4	42	1	0	0.9863	0.0987	12	1	0	0.9833	0.0866
0.05	2	60	3	0	0.9662	0.0527	17	3	0	0.9587	0.0475
	2.5	56	2	0	0.9692	0.0493	15	2	0	0.9681	0.0552
	3	56	2	0	0.9899	0.0493	15	2	0	0.9895	0.0552
	3.5	52	1	0	0.9637	0.0496	14	1	0	0.9605	0.052
	4	52	1	0	0.9788	0.0496	14	1	0	0.9769	0.052
0.01	2	84	4	0	0.9654	0.0110	23	4	0	0.9623	0.0117
	2.5	81	3	0	0.9808	0.0103	22	3	0	0.9787	0.0109
	3	78	2	0	0.9717	0.0104	21	2	0	0.9686	0.0111
	3.5	76	1	0	0.9221	0.0105	21	2	0	0.9877	0.0111
	4	76	1	0	0.9542	0.0105	21	1	0	0.9463	0.0097

Table 3 Optimal parameters of the QSS-1 for 20th percentile under Log-logistic distribution with γ =3

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	70	2	0	0.9965	0.2455	10	2	0	0.9961	0.2499
	2.5	57	1	0	0.9934	0.2502	8	1	0	0.9927	0.2525
	3	57	1	0	0.9978	0.2502	8	1	0	0.9976	0.2525
	3.5	57	1	0	0.9991	0.2502	8	1	0	0.999	0.2525
	4	57	1	0	0.9996	0.2502	8	1	0	0.9996	0.2525
0.10	2	103	3	0	0.9989	0.0995	15	3	0	0.9986	0.0909
	2.5	153	2	0	0.995	0.0106	13	2	0	0.9988	0.0994
	3	149	1	0	0.9846	0.0107	12	1	0	0.9942	0.0866
	3.5	149	1	0	0.994	0.0107	12	1	0	0.9977	0.0866
	4	149	1	0	0.9973	0.0107	12	1	0	0.999	0.0866
0.05	2	119	3	0	0.998	0.0500	17	3	0	0.9975	0.0475
	2.5	110	2	0	0.9982	0.0495	16	2	0	0.9976	0.0416
	3	110	2	0	0.9997	0.0495	16	2	0	0.9995	0.0416
	3.5	102	1	0	0.9972	0.0503	15	1	0	0.9964	0.0405
	4	102	1	0	0.9988	0.0503	15	1	0	0.9984	0.0405
0.01	2	166	4	0	0.999	0.0105	23	4	0	0.9989	0.0117
	2.5	160	3	0	0.9996	0.0100	22	3	0	0.9995	0.0109
	3	153	2	0	0.9991	0.0106	21	2	0	0.9989	0.0111
	3.5	150	1	0	0.9939	0.0104	21	1	0	0.9926	0.0097
	4	150	1	0	0.9973	0.0104	21	1	0	0.9967	0.0097

Table 4 Optimal parameters of the QSS-1 for 50th percentile under Log-logistic distribution with γ =1

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	13	7	0	0.9778	0.1291	11	9	0	0.9889	0.0769
	2.5	10	5	0	0.9851	0.1847	10	8	0	0.9973	0.0833
	3	9	4	0	0.9822	0.1522	7	5	0	0.9900	0.1111
	3.5	9	4	0	0.9918	0.1522	5	3	0	0.9660	0.1429
	4	6	2	0	0.9589	0.2155	5	3	0	0.9799	0.1429
0.10	2	14	7	0	0.9483	0.0561	10	8	0	0.9799	0.0833
	2.5	11	5	0	0.9669	0.0865	8	6	0	0.9864	0.100
	3	10	4	0	0.964	0.0752	7	5	0	0.9900	0.1111
	3.5	8	3	0	0.9683	0.1311	7	4	0	0.9579	0.0333
	4	8	3	0	0.9814	0.1311	7	4	0	0.9782	0.0333
0.05	2	14	7	0	0.9483	0.0561	13	10	0	0.9603	0.0108
	2.5	12	5	0	0.934	0.0416	12	9	0	0.9926	0.0125
	3	13	5	0	0.9581	0.0208	9	6	0	0.9824	0.0213
	3.5	11	4	0	0.9691	0.0385	8	5	0	0.9837	0.0263
	4	11	4	0	0.9842	0.0385	8	5	0	0.9927	0.0263
0.01	2	20	10	0	0.9534	0.0079	14	11	0	0.9766	0.0093
	2.5	20	10	0	0.996	0.0079	13	9	0	0.9676	0.0026
	3	16	6	0	0.9526	0.0058	12	8	0	0.9878	0.0033
	3.5	14	5	0	0.9706	0.0109	11	7	0	0.9920	0.0043
	4	13	4	0	0.9554	0.0113	11	7	0	0.9973	0.0043

Table 5 Optimal parameters of the QSS-1 for 50th percentile under Log-logistic distribution with γ =2

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	14	4	0	0.9979	0.2530	6	4	0	0.9939	0.125
	2.5	10	2	0	0.9918	0.2500	4	2	0	0.9832	0.1667
	3	9	1	0	0.9712	0.1923	4	2	0	0.9944	0.1667
	3.5	9	1	0	0.9845	0.1923	3	1	0	0.9799	0.200
	4	9	1	0	0.991	0.1923	3	1	0	0.9882	0.200
0.10	2	15	3	0	0.9765	0.0909	7	4	0	0.9782	0.0333
	2.5	13	2	0	0.9801	0.0994	6	3	0	0.9896	0.0435
	3	12	1	0	0.947	0.0866	5	2	0	0.9857	0.0588
	3.5	12	1	0	0.9714	0.0866	4	1	0	0.9595	0.0833
	4	12	1	0	0.9833	0.0866	4	1	0	0.9762	0.0833
0.05	2	17	3	0	0.9587	0.0475	8	5	0	0.9927	0.0263
	2.5	16	2	0	0.9606	0.0416	6	3	0	0.9896	0.0435
	3	15	2	0	0.9895	0.0552	6	2	0	0.971	0.0233
	3.5	14	1	0	0.9605	0.052	6	2	0	0.9886	0.0233
	4	14	1	0	0.9769	0.052	6	2	0	0.9949	0.0233
0.01	2	24	4	0	0.9523	0.0087	9	5	0	0.9777	0.0076
	2.5	22	3	0	0.9787	0.0109	8	4	0	0.9937	0.0106
	3	21	2	0	0.9686	0.0111	7	2	0	0.949	0.0100
	3.5	20	2	0	0.9895	0.0143	7	2	0	0.9797	0.0100
	4	20	1	0	0.9514	0.0122	7	2	0	0.991	0.0100

Table 6 Optimal parameters of the QSS-1 for 50th percentile under Log-logistic distribution with γ =3

β	θq/θo	a=0.5					a=1				
		n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$	n	c_N	c_T	$P_a(p_1)$	$P_a(p_2)$
0.25	2	19	2	0	0.9961	0.2311	5	3	0	0.9988	0.1429
	2.5	15	1	0	0.9931	0.2515	4	2	0	0.9989	0.1667
	3	16	1	0	0.9974	0.2182	3	1	0	0.9959	0.2
	3.5	16	1	0	0.999	0.2182	3	1	0	0.9984	0.2
	4	16	1	0	0.9995	0.2182	3	1	0	0.9993	0.2
0.10	2	25	2	0	0.9905	0.0894	5	2	0	0.9796	0.0588
	2.5	21	1	0	0.9861	0.1082	4	1	0	0.975	0.0833
	3	21	1	0	0.9954	0.1082	4	1	0	0.9916	0.0833
	3.5	21	1	0	0.9982	0.1082	4	1	0	0.9967	0.0833
	4	21	1	0	0.9992	0.1082	4	1	0	0.9985	0.0833
0.05	2	34	3	0	0.997	0.0332	6	3	0	0.9962	0.0435
	2.5	29	2	0	0.998	0.0489	6	2	0	0.9945	0.0233
	3	29	2	0	0.9996	0.0489	6	2	0	0.999	0.0233
	3.5	27	1	0	0.997	0.0484	5	1	0	0.9945	0.037
	4	27	1	0	0.9986	0.0484	5	1	0	0.9975	0.037
0.01	2	47	5	0	0.9998	0.0092	9	5	0	0.9997	0.0076
	2.5	41	2	0	0.9941	0.0093	8	2	0	0.9843	0.0045
	3	41	2	0	0.9989	0.0093	8	2	0	0.997	0.0045
	3.5	41	2	0	0.9997	0.0093	8	2	0	0.9993	0.0045
	4	41	2	0	0.9999	0.0093	8	2	0	0.9998	0.0045

From above tables, it's understood that when the percentile ratio is increases from 2 to 4 for fixed values of α, β and a then the sample size decreases.The optimal parameters of the QSS-1 are same when γ=3 andβ =25%,for the percentile ratios from 3 to 4 at both cases of
a = 1.0 and a = 0.5.

6. PLOTTING OC CURVE

For each sampling plan there is an OC curve which portrays the performance of the sampling performance against good and bad quality. The OC curve for QSS-1 is constructed using table 1ie., the table of optimal parameter of QSS-1 for 20th percentile under log-logistic distribution with γ =1and experiment termination ratio a=0.5. Figure 1.1 shows the OC curve of QSS-2 with randomly chosen parameters n=22,$c_N = 5$,$c_T = 0$

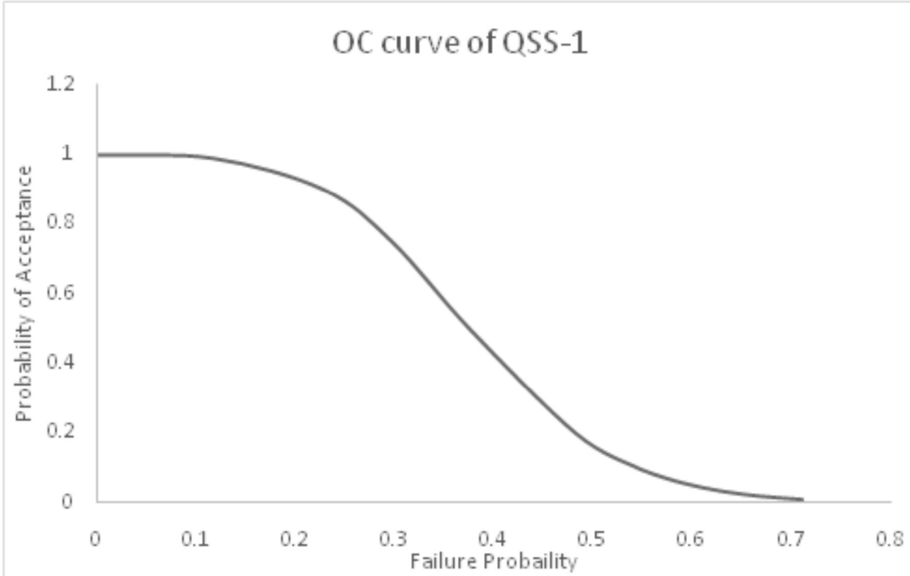


Figure 1 OC curve of QSS-2(22,5,0)

Table-7 OC values for 'p' and ' $P_a(p)$'

p	0.2816	0.222	0.2086	0.1971	0.1756	0.1503	0.133	0.1239	0.1057	0.071
$P_a(p)$	0.0102	0.0409	0.0609	0.0895	0.2009	0.4997	0.7509	0.8501	0.9574	0.9982

7. ALGORITHM

The following algorithm were applied to study the performance measure of the system,

Step 1: Specify β, α, a, σ and percentile.

Step 2: To find the failure probability corresponding to AQL, substitute the values a,γ with the specified percentile ratio in equation (9). Also, find the failure probability corresponding to LQL using the percentile ratio 1.

Step 3: Find the lot acceptance probability using equation (1) under QSS-1 for the required AQL and LQL and denote it as $P_a(p_1)$and $P_a(p_2)$respectively.

Step 4: Find the minimum value of *n* such that both two conditions such as

$P_a(p_1) \geq 1-\alpha$ and $P_a(p_2) \leq \beta$ are satisfied.

Step 5: If such an *n* exists, then the required optimum parameters are (n, c_N, c_T) .Else repeat Step 3 and Step 4 for several combinations of n, c_N, c_Ttill an optimum set of parameters is attained.

8. REAL LIFE EXAMPLE

The real data is used to implement the proposed QSS-1 for assuring 20th percentile life of the product where the life time follows Log-Logistic distribution with shape parameter γ =1.The data is discussed by

Schmee and Nelson (1977) and show the number of thousand miles at which different locomotive controls failed.The producer and consumer risks are considered as α=5% and β=25% respectively and the percentile ratio $\frac{\theta_q}{\theta_0}$=3.Let the specified percentile of the product θ_0= 49 and the experiment time be t_0=49 so that the experiment termination ratio a=1.0.For the specified requirements,Table 1.1 yields the optimal parameters as n=12, $c_N = 3$, $c_T = 0$**)**.Then using the data the proposed plan can be implemented with the above mentioned optimal parameters as follows

Start with Normal inspection and select a random sample of 12 units from the lot and put them on life test for specified time 49. Suppose that 12 units are having failure time as follows,

22.5	37.5	46.0	48.5	51.5	53.0
54.5	57.5	66.5	68.0	69.5	76.5

4 observations are failed that are 22.5,37.5,46.0,48.5,which are less than the experiment time 49 (i.e.d_N= 4). $\boldsymbol{d_N} > c_N$-Reject the lot and switch to tightened inspection for next lot.

Under the tightened inspection, the life test is conducted for specified time 49 with sample size 12 .The failure time of 12 units are as follows

77.0	78.5	80.0	81.5	82.0	83.0
84.0	91.5	93.5	102.5	107.0	108.5

d_T=0, $d_T = c_T -$ accept the lot under tightened inspection.Switched to normal inspection for next lot.

9. CONCLUSION

The work presented in this research paper is mainly focused on QSS of type 1 of (n, c_N, c_T) indexed with incoming and outgoing quality levels. In thisarticle an attribute Reliability Sampling system named as Quick Switching Single Sampling System of type1 is proposedfor assuring product lifetime to study the percentile life under Log-logistic distribution with single sampling plan as reference plan. Tables have been provided for selecting the optimal parameters of the proposed system given specification. It is concluded that the proposed QSS-1 is useful in quality control inspection and effective in reducing the inspectiontime and cost. OC curve were developed and designing method of minimum sample size, protecting the consumer and producer are given in the tables. The sample size of the Quick switching system is decreases when the percentile ratio increases. The proposed sampling system under Log-logistic distribution will be a fruitful addition in the literature of acceptance sampling. For future study this work also extended to designing a reliability sampling plan for various percentile levels for the other continuous probability distributions. The efficiency of the proposed sampling plans using a cost model can be considered as future study. This can also applied in various reference plan like Skip lot Sampling Plan, TNT plans.

REFERENCES

1. *Aslam, M., and C.-H. Jun (2010). A double acceptance sampling plan for generalizedlog-logistic distributions with known shape parameters, J. Appl. Stat. 37 (2010), no.3, 405–414.*
2. *Balamurali S, P. Jeyadurga& M. Usha (2018) Determination of optimal quick switching system with varying sample size for assuring mean life under Weibull distribution, Sequential Analysis, 37:2, 222-234*
3. *Balamurali S, P. Jeyadurga& M. Usha (2017): Optimal designing of quick switching sampling system assuring percentile life under Weibull distribution, Communications in Statistics - Simulation and Computation.*
4. *Balamurali S, P. Jeyadurga& M. Usha (2017): Economic design of quick switching sampling system for assuring weibull distributed mean life, Communications in Statistics - Theory and Methods,*
5. *Cameron, J M. (1952). Tables for Constructing and for Computing the Operating Characteristics of Single Sampling Plans. Industrial Quality Control, 9(1): 37 – 39*
6. *Govindaraju K, Subramani K. (1992). Selection of Tightened - Normal - Tightened System for Given Values of Acceptable Quality Level and Limiting Quality Level, Journal of Applied Statistics, 19(2): 241-250.*
7. *Guenther W C. (1969). Use of the Binomial, Hyper-geometric and Poisson Tables to Obtain Sampling Plans. Journal of Quality Technology, 1(2): 105 – 109.*
8. *Kantam, R.R.L. and Rosaiah, K. and Rao, G.S. (2001), Acceptance sampling based on life tests: Log-Logistic model. Journal of Applied Statistics. 28: 121-128.*
9. *Kantam, R.R.L., G. Srinivasa Rao, and G. Sriram (2006), An economic reliability test plan: log-logistic distribution, J. Appl.Statist. 33(3): 291–296.*
10. *Kaviyarasu V. (2012). Contributions to the Study on Quick Switching System through Incoming and Outgoing Quality Levels. Ph.D thesis, Bharathair University, Coimbatore, India.*
11. *Montgomery D C. (2009). Introduction to Statistical Quality Control, 6th edition. John Wiley & Sons, New Jersey.*
12. *Rao, G.S. and Kantam, R.R.L. (2010), Acceptance sampling plans from truncated life tests based on the log-logistic distribution for percentiles. Economic Quality Control.25(2), 153- 167.*
13. *Romboski L D. (1969). An Investigation of Quick Switching Acceptance Sampling Systems. Ph.D Thesis, Rutgers- The State University, New Brunswick, New Jersey.*

14. *Schilling E G. (1982). Acceptance Sampling in Quality Control. Marcel Dekker, Inc.,*
15. *Shah, B. K., & Dave, P. H. (1963). A note on log-logistic distribution. Journal of the MS University of Baroda (Science Number), 12, 15-20*
16. *Soundararajan V, Arumainayagam S D. (1988). Quick Switching System, Technical Report No.24 Department of Statistics, Bharathiar University, Tamil Nadu, India.*
17. *Soundararajan V, Arumainayagam S D. (1992). Quick Switching System for Costly and Destructive Testing. Sankaya, Series B–Part I, 54: 1-12*
18. *Tadikamalla, P. R., & Johnson, N. L. (1982). Systems of frequency curves generated by transformations of logistic variables. Biometrika, 69(2), 461-465.*

DOI No. : 10.5958/2394-9309.2022.00026.9
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A STATISTICAL STUDY ON BREAST MALIGNANCY PATIENTS SURGERY IN INDIA

Salomi M[*]. R. Rajesh[**] and Manimannan G.[***]
[*]Assistant Professor, Department of Statistics, Madras Christian College, Chennai
[**,***]Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai, Chennai

ABSTRACT
Objective: To determine the effects of breast malignancy surgery of Tertiary care centre of India. The data was collected from case sheet from a private clinic.
Study Design: A Descriptive Statistical Study.
Methodology: Seventy five patients were operated with nineteen parameters and three groups of patients. Each groups have twenty five samples in the name of Group R=Ropivacaine, Group RD=Dexmedetomidine and Group RK=Ketamine. The parameters are patients details, Anthropometric, Height, Weight, BMI, Site of Surgery, Duration (hours), Anesthesia(ASA) Grade, First request analgesia (hours), Post-operative analgesic consumption (mg),Heart Rate (beats/minutes), Systolic Blood Pressure (SBP - mmof Hg), oxygen saturation (SPo2%), Numerical Rating Scale at rest, Numerical Rating Scale at movement, Sedation Score and Patient Satisfaction Score.
Results: Ithas been foundthat Chi-square test results some insignificant relationship between the levels of site of surgery and test results of Anesthesia (ASA) status among studied groups. In ANOVA and Post-hoc LSD t-test, test results are significant in Age, duration of surgery ,Anthropometric measurements in study groups, time requirement of first rescue analgesia, post-operative analgesic consumption, patient satisfaction score (PSS) in studied groups at 24 hours. In addition, mean±Standard Deviation and p-values ofIntraoperative and postoperative hemodynamic parameters, Numerical Rating Scale (NRS) at rest in studied groups, Sedation score (Mean ± S.D) in studied groups were identified. Finally, some important clinical parameters are visualized.
Conclusion: Breast Malignancy patient Level of Site of Surgery and anesthesiahad no effective and rest of the clinical treatments are effective intraoperative and postoperative hemodynamic parameters patient satisfaction.
Keywords: Breast Malignancy, Intraoperative, Postoperative,Anesthesia, Descriptive Statistics, and Visualization.

1. BACKGROUND

Breast-Conserving Therapy (BCT) is the preferred treatment approach for most women with early stage breast cancer [1]. Complete resection of the primary lesion with tumor-free margins and adjuvant radiotherapy provide for optimal local tumor control. Margin status is an important prognostic factor for local failure after removalof invasive breast carcinoma. Pathological examination of margin status plays a key role in BCT [2]. Intraoperative evaluation of margin status allows immediate reexcision of suspicious margins, minimizing the need for secondary operative procedures.16, 17 techniques were used to evaluate margins in the operating room and are simple, rapid, and inexpensive even if they are to be incorporated into routine clinical practice. Unfortunately, all methods used to determine margin status have some technical or practical limitations [3].

Breast cancer surgery is associated with mild to moderate pain but some procedures including axillary nodes dissection are more painful [4]. In these patients, pain may impair postoperative comfort and may prevent from mobilization of the corresponding upper limb. Moreover, studies have pointed out that chronic pain syndromes may develop after breast surgery that could be, at least partly, related to the intensity of acute postoperative pain and axillary nodes dissection [5]. Local infiltration with a local anesthetic solution has been studied on several instances with disappointing results, most of the studies documenting indeed the lack of significant effect of wound infiltration on postoperative pain after breast surgery. In addition, in several trials breast infiltration is not combined with axillary infiltration [6]. In this paper, an attempt is made to identify the significance and summary statistics of randomized controlled trial on breast malignancy patients undergone breast surgery in eastern India.

2. METHODOLOGY

The secondary sources of breast malignancy patient's data were collected from tertiary care centre of eastern India. A Randomized controlled trial on breast malignancy patients undergone breast surgeries in a tertiary care center of eastern India on three groups of each 25 patients. Group R = (Ropivacaine 0.2% only), Group RD = (dexmedetomidine 1 mcg /kg with ropivacaine 0.2%) and Group RK = (ketamine 1 mg/kg with ropivacaine 0.2%), to assess the effectiveness of ropivacaine along with adjuvants in view of bringing down the post-operative pain and improving the numerical rating scale. Total sample of 75 is considered for this study with parameters age, Site of surgery, Anthropometric measurements, Site of surgery, Duration of surgery, Anesthesia (ASA) Grade, Time requirement of first rescue analgesia and post-operative analgesic consumption (mg).

3. RESULT AND DISCUSSION

The data were evaluated by an independent analyst blinded to the allocations of patient groups. Continuous variables are expressed as the mean ± SD (Standard Deviation). Differences were assessed using one-way Analysis of Variance (ANOVA). Pairwise comparison of one-way ANOVA was made using post-hoc analysis and Student–Newman–Keuls Q-test. The corrected p-value was obtained directly, and cutoff value is 0.05. The summary statistics of age distribution is given in Table 1 and Figure 1.

Table 1.Age Distribution among Studied Groups

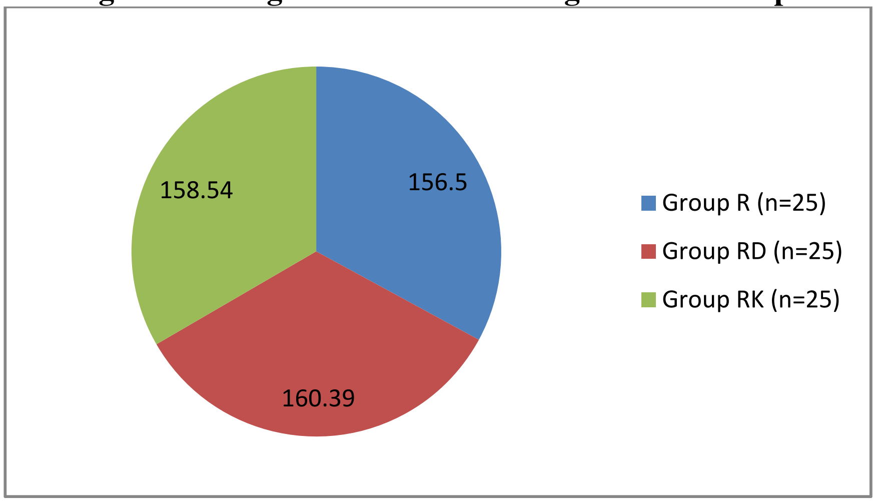
Variables	Group R (n=25)	Group RD (n=25)	Group RK (n=25)
Age (mean± S.D) years	48.25 ± 05.68	47.81 ± 06.21	46.32 ± 05.89

In ANOVA table 2, the F-ratio = 0.727, Degrees of freedom = 2 and 72, Two-tailed probability = 0.487, Written as: $F(2, 72) = 0.727$, $p=0.487$, at the .05 critical alpha level: The age distribution among studied groups arenot significant.

Table 2. Analysis of Variance (ANOVA) Table

Source	DF	SS	MS	F	p
Factor	2	51.15	25.58	0.727	.487
Error	72	2532.45	35.17		
Total	74	2583.60			

Fig 1: Mean Age Distribution Among Studied Groups



The summary statistics and chi-square test of anthropometric measurements with Group R, Group RD and Group RK results is given in Table 2 and Figure 2. This result showsthat anthropometric measurements in study groups are statistically highly significant.

Table 2: Anthropometric measurements in study groups

Variable	Group R (n=25)	Group RD (n=25)	Group RK (n=25)	p - value
Height (cm)	156.50 ± 4.12	160.39 ± 5.90	158.54 ± 6.12	$< 0.0001^{HS}$
Weight (Kg)	74.53 ± 7.60	77.24 ± 9.04	75.65 ± 5.98	0.0036^{S}
BMI	30.89 ± 3.93	25.32 ± 2.92	24.65 ± 4.65	$< 0.0001^{HS}$

NS=Not significant (p>0.05); S=Significant (p<0.05, p<0.01); HS= Highly Significant (p<0.001).

Fig 2: Mean Anthropometric Measurements in Study Groups

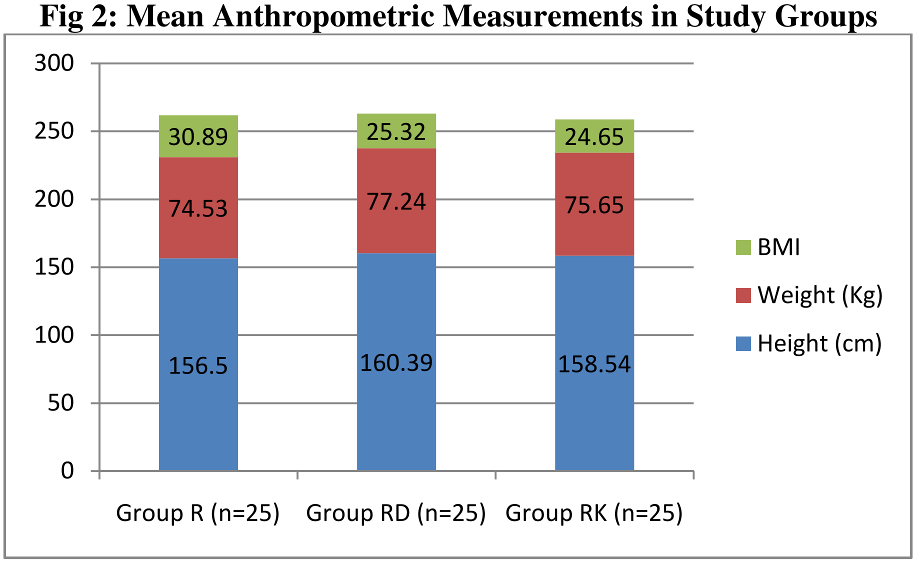
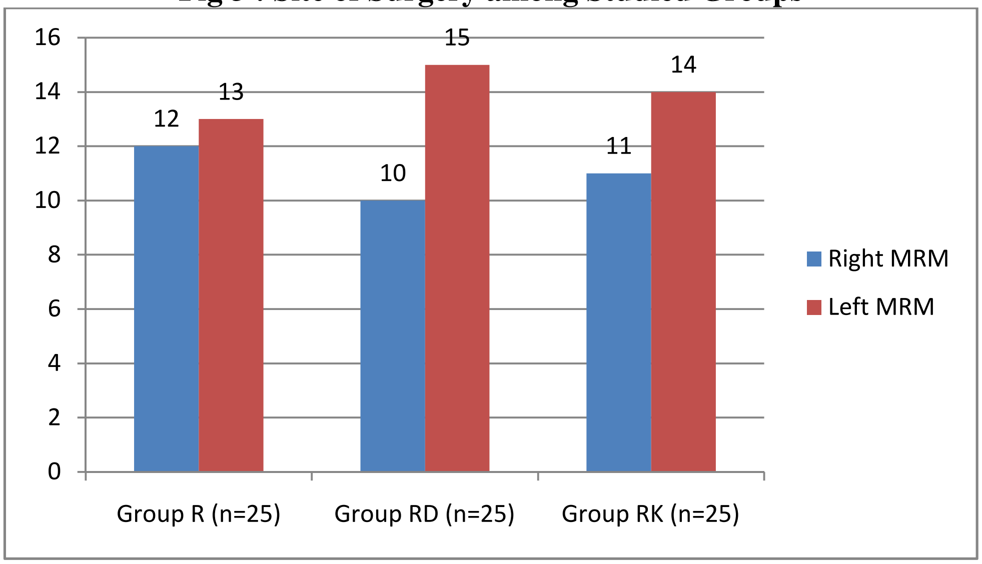


Table 3: Site of Surgery among Studied Groups

Site	Group R (n=25)	Group RD (n=25)	Group RK (n=25)	p - value
Right MRM	12 (48.00%)	10 (40.00%)	11 (44.00%)	< 0.5267NS
Left MRM	13 (52.00%)	15 (60.00%)	14 (56.00%)	<0..5132NS

The summary statistics and chi-square test of site of Surgery among studied groups with Group R, Group RD and Group RK results is showed in Table 3 and Figure 3. These results shows the Site of Surgery among study groups and are not statistically significant at 5 percent level of significance.

Fig 3 : Site of Surgery among Studied Groups



The duration of surgery among studied group's statistics isgiven in Table 4, ANOVA results in table 5 and Post-hoc Least Significant Difference (LSD) t-test in Table 6. In both methods the difference of duration surgery among study groups are significant at 5 percent level of significance.

Table 4: Duration of Surgery among Studied Groups

Duration (mean± S.D) (hrs)	Group R (n=25)	Group RD (n=25)	Group RK (n=25)
	02.24 ± 0.72	02.89 ± 0.76	02.12 ± 0.67

F-ratio = 8.3323, Degrees of freedom = 2 and 72, Two-tailed probability < 0.001 Written as: F (2, 72) = 7.799, p<0.001, Conclusion at the 0.05 critical alpha level: The difference is significant

Table 5.Analysis of Variance table

Source	DF	SS	MS	F	p
Factor	2	8.58	4.29	8.3323	<.001
Error	72	37.08	0.52		
Total	74	45.66			

Table 6.Post-hoc lsd t-tests

Groups	lsd t-test
1 & 2	t(48)=3.098, p=.003
2 & 3	t(48)=3.67, p<.001

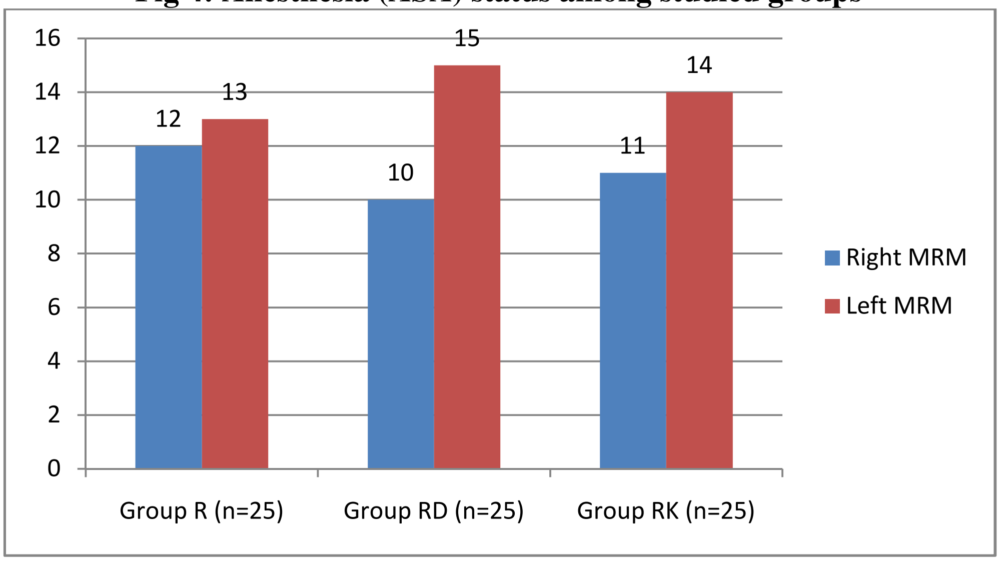
In Anesthesia (ASA) status among studied groups results are in Table 7 and figure 4. Chi-square calculated value is 1.2821 and probability value is 0.526752 which states that ASA grade and study group is not significant at 5 percent level of significance.

Table 7: Anesthesia (ASA) Status among Studied Groups

ASA Grade	Group R (n=25)	Group RD (n=25)	Group RK (n=25)	p - value
I	12 (48.00%)	14 (56.00%)	10 (40.00%)	< 0.52675 NS
II	13 (52.00%)	11 (44.00%)	15 (60.00%)	<0..521421 NS

The chi-square statistic is 1.2821. The p-value is 0.526752. The result is **not significant** at $p < 0.05$.

Fig 4: Anesthesia (ASA) status among studied groups



The time requirement of First rescue Analgesia among studied group's statistics is in Table 8 and figure 5, ANOVA results are shown in table 9 and Post-hoc Least Significant Difference (LSD) t-test in Table 10. In both methods the difference of duration time requirement of first rescue Analgesia among study group is significant at 5 percent level of significance.

Table 8: Time Requirement of First Rescue Analgesia among Studied Groups

First request analgesia (hrs.) (mean ± S.D)	Group R (n=25)	Group RD (n=25)	Group RK (n=25)
	14.26 ± 01.97	18.79 ± 02.39	12.26 ± 02.84

F-ratio = 57.108, Degrees of freedom = 2 and 72, Two-tailed probability < .001 Written as: $F(2, 72) = 57.108$, $p<0.001$, Conclusion at 0.05 critical alpha level: The difference is significant.

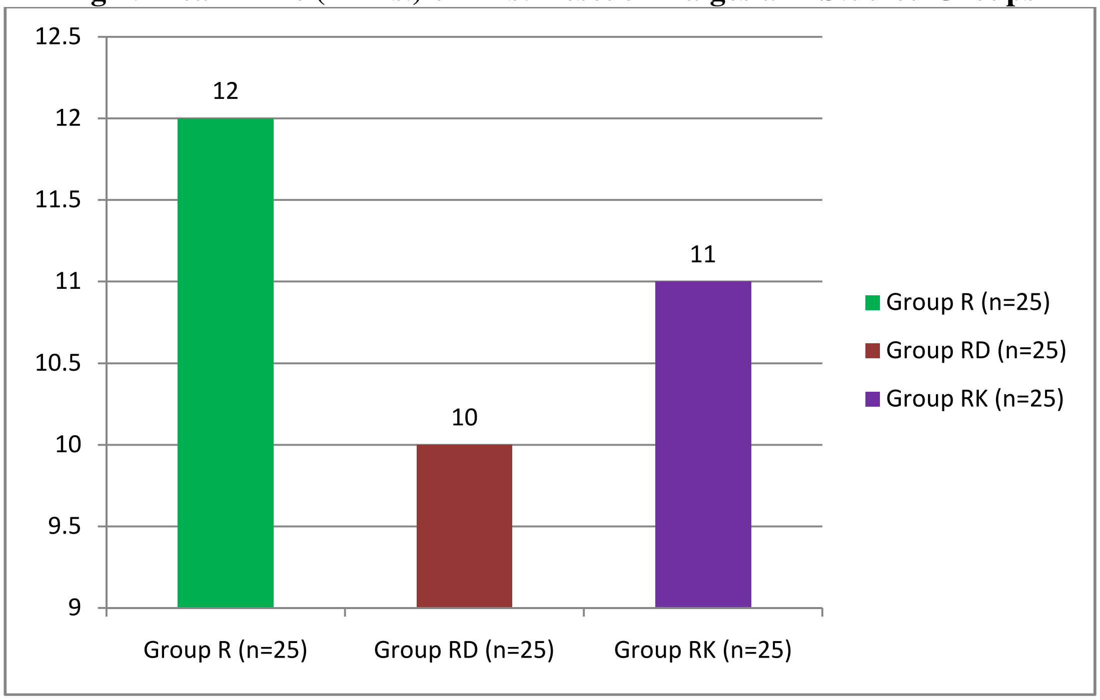
Table 9. Analysis of Variance Table

Source	DF	SS	MS	F	p
Factor	2	559.68	279.84	57.108	<.001
Error	72	352.81	4.90		
Total	74	912.50			

Table 10. Post-hoc lsd t-tests

Groups	lsd t-test
1 & 2	t(48)=7.235, p<.001
1 & 3	t(48)=3.194, p=.002
2 & 3	t(48)=10.429, p<.001

Fig 5: Mean Time (in hrs.) of First Rescue Analgesia in Studied Groups



Post-operative analgesic consumption in studied group's statistics is given in Table 11 and figure 6, ANOVA results in table 12 and Post-hoc Least Significant Difference (LSD) t-test in Table 13. The difference of total post-operative analgesic consumption in studied group is significant at 5 percent level of significance in both methods.

Table 11: Total Post-Operative Analgesic Consumption in Studied Groups

Post-operative analgesic consumption (mg) (mean ± S.D)	Group R (n=25)	Group RD (n=25)	Group RK (n=25)
	21.26 ± 03.14	14.89 ± 02.43	12.51 ± 04.62

F-ratio = 41.366, Degrees of freedom = 2 and 72, Two-tailed probability < 0.001 Written as: F (2, 72) = 41.366, p<0.001, Conclusion at the .05 critical alpha level: The difference is significant

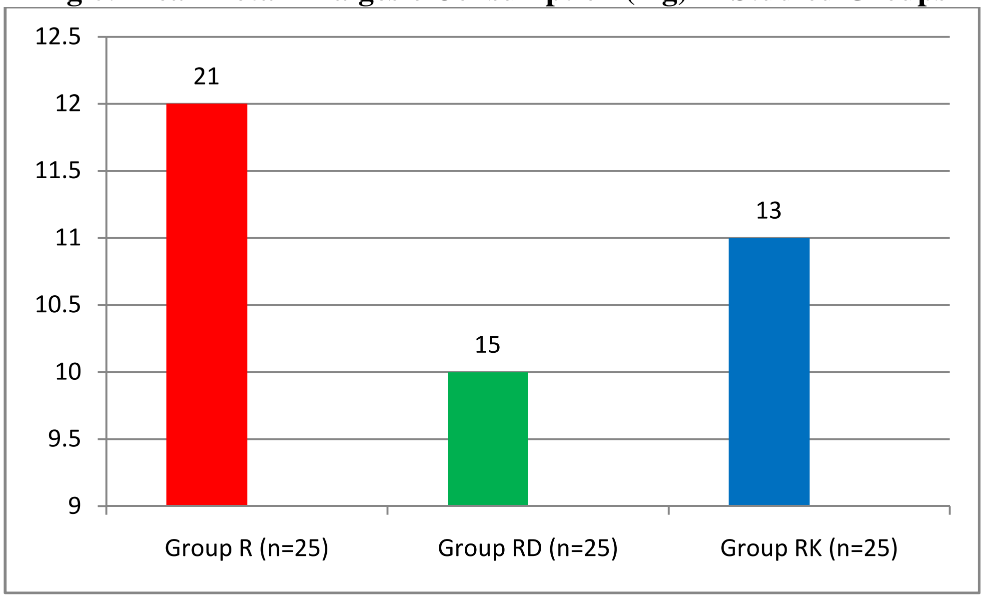
Table 12. Analysis of Variance Table

Source	DF	SS	MS	F	p
Factor	2	1023.37	511.68	41.366	<.001
Error	72	890.61	12.37		
Total	74	1913.98			

Table 13. Post-hoc lsd t-tests

Groups	lsd t-test
1 & 2	t(48)=6.403, p<.001
1 & 3	t(48)=8.796, p<.001
2 & 3	t(48)=2.393, p=.021

Fig 6: Mean Total Analgesic Consumption (mg) in Studied Groups



4. CONCLUSION

Chi-square test results clearly show some insignificant relationship between the levels of site of surgery and test results of Anesthesia (ASA) status among studied groups. ANOVA and Post-hoc LSD t-test test results exhibitsthe existence of significance in Age, duration of surgery , Anthropometric measurements,Time requirement of first rescue analgesia, post-operative analgesic consumption,Patient satisfaction score (PSS) in studied groups at 24 hours. In addition,mean±Standard Deviation and p-values ofIntraoperative and postoperative hemodynamic parameters, Numerical Rating Scale (NRS) at rest,Sedation score (Mean ± S.D) in studied groups were identified. Finally, some important clinical parameters are visualized.Breast Malignancy patient Level of Site of Surgery and anesthesia had no effect and rest of the clinical treatments are effective with intraoperative and postoperative hemodynamic parameters patient satisfaction.

REFERENCES

1. *National Institutes of Health Consensus Conference. Treatment of early-stage breast cancer. JAMA 1995; 265:391–5.*
2. *Van Dongen JA, Bartelink H, Fentiman IS, et al. Factors influencing local relapse and survival and results of salvage treatment after breast-conserving therapy in operable breast cancer: EORTC trial*

10801, breast conservation compared with mastectomy in TNM stage I and II breast cancer. Eur J Cancer 1992; 28A:801–5

3. *Klimberg VS, Harms S, Korourian S. Assessing margin status. SurgOncol 1999;8:77–84.*
4. *Tasmuth T, von Smitten K, Hietanen P, Kataja M, Kalso E: Pain and other symptoms after different treatment modalities of breast cancer. Ann Oncol 1995, 6:453-9. 4.*
5. *Gärtner R, Jensen MB, Nielsen J, Ewertz M, Kroman N, Kehlet H: Preevalence of and factors associated with persistent pain following breast cancer surgery. JAMA 2009, 302:1985-92. 5.*
6. *Marret E, Vigneau A, Salengro A, Noirot A, Bonnet F: Effectiveness of analgesic techniques after breast surgery: a meta-analysis. Ann FrAnesthRéanim 2006, 25:947-54.*

DOI No. : 10.5958/2394-9309.2022.00027.0
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

ESTIMATION OF RAILWAY QUEUE SYSTEM USING MARKOVIAN QUEUING MODEL

V. P. Murugan* and S. Syed Ali Fathima**

*Research Scholar, Department of Mathematics, Manonmaniam Sundaranar University ,Tirunelveli.
**Assistant Professor, Department of Mathematics, Sadakathullah Appa College (Autonomous), Tirunelveli.

ABSTRACT

Queuing system is a most important application of probability theory. Queuing models mainly used in the demand for service is more than the capacity to give service. In this module, we describe the M/M/C model for analyzing elements of a Queuing system. This paper elucidates the analysis of Queuing system of Railway queue. Queuing models is used to give an optimum balance between waiting time of customers and idle time of service facilities. This paper also describes a Queuing simulation for multiple server process as well as for single Queue models. This study requires an empirical data analysis which may contains the variables like arrival time in the Queue, waiting time and service time. This model is developed for minimize customer's waiting time at railway ticketing system. The interpretation of the railway system using Monte Carlo simulation method enables to reduce minimum waiting time of the passenger's in a Markovian Queues models of M/M/I and M/M/C. So as to enhance a better service to the customers in railway ticketing system.

Keywords: Queue, service facility, waiting time, Monte Carlo simulation.

1. INTRODUCTION

This is the question we ask ourselves. Waiting lines occur in our daily life. We wait in a line to eat in restaurants, we wait to have a service in banks, we wait for service in post offices, we wait to pay bills at the check-out counters in a grocery store or in a supermarket (mall), we wait to have a fuel at a petrol pump, we wait in a queue while boarding into a plane at an airport and even the cricketers wait in a dressing room for their turn during a cricket match. And waiting is experienced not only by human beings but jobs also wait to be processed on machine, airplanes wait to take off or to land at an airport, cars stop at traffic lights and at a parking lot and even the thoughts wait in a line to be thought in our brain. Due to waiting, we have to tolerate a delay to get the service[1]. Our goal is to reduce this delay to the tolerable level. For this, we study the measures of performance related to waiting lines, such as average queue length, average waiting time in a queue, average waiting time in a system and average service facility utilization.

2. QUEUING PROCESS

Customers requiring service are generated over time by an input source. The required service is then carried out for the customers using the provider mechanism, after which the purchaser leaves the queuing system [6].

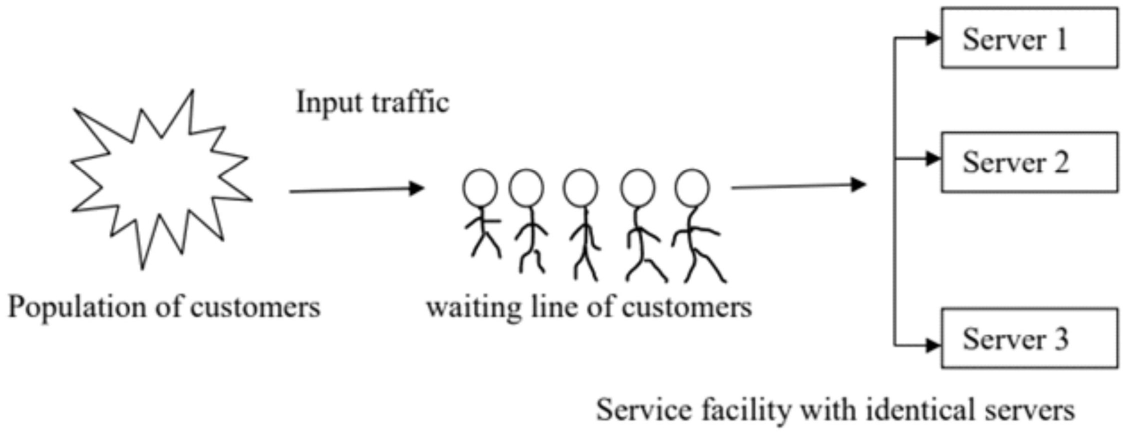


Figure: 1 Customer based queuing models

We can have the following two types of models: One model will be as Single-queue Multiple Servers model (fig.1) and the second one is Multiple-Queues Multiple-Servers model and single queue and a single server. The queuing process comprises three components such as Arrival Process, Waiting Process, and the server process [7].They are following,

a. Arrival Process:

Includes a variety of customers arriving, several types of customers, and a kind of customers' demand, deterministic or stochastic arrival distance, and arrival intensity. The method goes from event to event, i.e., the tournament "customer arrives" places the consumer in a queue, and at the equal time schedules, the event "next purchaser arrives" at some time in the future.

b. Waiting Process:

Includes the size of queues, servers' discipline (First in First out). This consists of the tournament "start serving subsequent customer from queue" which takes this consumer from the queue into the server and at the same time schedules the event "customer served" at some time in the future.

c. Server Process

Consists of a type of a server, serving price and serving time. This includes the event "customer served" which prompts the subsequent event "start serving next customer from queue."

d. Service discipline

Service discipline or order of service is a rule by which customers or selected for service from the queue. The most common discipline is „first come, first served' (FCFS) or 'first in, first out' (FIFO) according to which the customers are served in the order of their arrival. Example: cinema ticket counters, railway booking counters[9]

e. Maximum queuing system capacity:

Maximum number of customers in the system can be either finite or infinite. In some facilities only limited number of customers are allowed in the system, the new arrivals are not allowed to join the queue. In some system the queue capacity is assumed to be infinite, if every arriving customers is allowed to wait until service is provided. The author proposed, the path of the network is based on the routing table it is not fixed in Ad-hoc technique network path is also not fixed and create dynamically. It design is practical oriented But it is more cost[10]

The following are usual notations in the discussions

n = number of customers in the system (i.e. waiting for service in the queue + being served)

λ = mean arrival rate (i.e. average number of customers arriving per unit time)

μ = mean service rate per busy server (i.e. average number of customers served per unit time)

ρ = is traffic intensity or utilization factor

c = number of parallel service channels

Lq = mean length of the queue (i.e. average or expected number of customers waiting in the queue)

Ls = mean length of the system (i.e. average or expected number of customers both waiting and in service)

Wq = mean waiting time in the queue (i.e. the expected waiting time before being served)

Ws = mean waiting time in the system (i.e. the expected waiting time in the system)

P_n = steady state probability of n customers in the system [10]

f. Multiple queue, multiple server Model Formula (M/M/C)

The following parameters are calculated, and their formulas have been presented

n – Number of total customers in the system (in queue plus in service)

c – Number of parallel servers

μ– Serving rate

cμ– Serving rate when c > 1 in a system

System intensity is

$$\rho = \frac{\lambda}{c\mu} \qquad \text{--------} \ [1]$$

P_0: Steady-state Probability of all idle servers in the system / Probability that there is no customer in the system

$$p_0 = \left[\sum_{n=o}^{c-1} \frac{r^n}{n!} + \frac{r^c}{c!(1-\rho)} \right]^{-1}$$

Lq: Average number of customer in the waiting line(queue)

$$L_q = \frac{r^c \rho}{c!(1-\rho)^2} \bullet P_0$$

Ls : Average number of customers in the system

$$L_s = L_q + \frac{\lambda}{\mu}$$

Wq: Average amount of time a customer spends in queue

$$W_q = \frac{L_q}{\lambda}$$

Ws: Average amount of time a customer spends in the system [8]

$$W_s = W_q + \frac{1}{\mu}$$

3. ANALYSIS

Data analysis is considered to be a significant step in research work. Collection of data is described with the help of relevant tools and techniques, after that analyses and interprets with a view to arrive at a practical solution to the problem. The data analysis for the research study was done quantitatively with the help of inferential statistics. Standard Statistical Package for Social Sciences (SPSS) and the queueing simulation software was used for analyzing the data. Sampling Design: As the study includes Railway que system in Chennai. The sample of the survey considered as mentioned below: Sample size calculation at the Confernce level at 95%

$$n_o = \frac{z^2 pq}{e^2}$$

n_0: Average sample size.

Z: value for confidence level (1.96 for 95% confidence interval)

p: is percentage for picking a choice (p= (1-α)/2, where α = 1-CL = 1-.95)

q: is 1-P

e: is the desired level of precision (The desired level of precision is 5%).

$$n_0 = \frac{(1.96)^2(0.475|(0.525)}{(0.05)^2} = 380$$

Therefore the sample size is 380.

TABLE 1: Frequency Table for Arriving Customers

		Frequency (Arrival)				Total
		One queues	Two queues	Three queues	N - queues	
First issue	Yes	51	55	42	54	202

	No	43	44	40	51	178
Total		94	99	82	105	380

TABLE 2: Frequency Table for Servicing Customers

		Frequency (Service)				Total
		One queues	Two queues	Three queues	N - queues	
First issue	Yes	54	49	47	52	202
	No	47	43	49	39	178
Total		101	92	96	91	380

TABLE 3 : M/M/1 MODEL

System Description	SERVER - 1
Arrival rate	24.74
Service rate	26.58
System intensity	0.93
W_s : average amount of time a customer spends in the system	0.54
W_q : average amount of time a customer spends in the queue	0.51
L_s : average number of customers in the system	13.4
L_q : average number of customers in the queue	12.51
P_0 :Idle System	0.07

TABLE 4: M/M/2 MODEL

System Description	SERVER - 1
Arrival rate	24.74
Service rate	26.58
System intensity	0.46
W_s : average amount of time a customer spends in the system	0.067
W_q : average amount of time a customer spends in the queue	0.029
L_s : average number of customers in the system	1.64
L_q : average number of customers in the queue	0.71
P_0 : Idle System	0.36

4. CONCLUSION

This paper presents a queuing model for multiple servers. The average queue length can be estimated simply from raw data from questionnaires by using the collected number of customers waiting in a queue each minute. We can compare this average with that of queuing model. Two different models are used to estimate a queue length: a single-queue multi-server model, single-queue single-server. In case of more than one queue (multiple queue), customers in any queue switch to shorter queue (jockey behavior of queue). Therefore, there are no analytical solutions available for multiple queues and hence queuing simulation is run to find the estimates for queue length and waiting time .Queues are also common in computers waiting systems, queues of enquiries waiting to be processed by an interactive computer system. Queues of data base request, queues of input/output requests. From the table efficiency parameters under two different queuing models by using system utilization parameters, it has changed significantly from 93% to 46% for the models M/M/1 and M/M/2 consecutively. When comparing with single and multi channels, waiting time of the customers are slightly reduced in the multi channels rather than the single channels.

REFERENCE

1. *Balaji, N., Siva, E. P., Chandrasekaran, A. D., &Tamilazhagan, V. (2018, April). Optimal service using Matlab-simulink controlled Queuing system at call centers. In Journal of Physics: Conference Series (Vol. 1000, No. 1, p. 012167). IOP Publishing.*
2. *Amit, N., & Ghazali, N. A. (2018). Using Simulation to Model Queuing Problem at a Fast-Food Restaurant. In Regional Conference on Science, Technology and Social Sciences (RCSTSS 2016) (pp. 1055-1062). Springer, Singapore.*
3. *Kumar, J., Shinde, V., & Bharadwaj, S. (2017). Sale Checkout Operation In Big Bazaar Using Queuing Model. International Journal of Mathematical Archive EISSN 2229-5046, 8(7).*
4. *Ejaz, I., Alvarado, M., Gautam, N., Gebraeel, N., & Lawley, M. (2019). Condition-Based Maintenance for Queues With Degrading Servers. IEEE Transactions on Automation Science and Engineering.*
5. *Hanukov, G., Avinadav, T., Chernonog, T., &Yechiali, U. (2019). Performance improvement of a Service System via Stocking Perishable Preliminary Services. European Journal of Operational Research, 274(3), 1000-1011.*
6. *Murdapa, P. S., Pujawan, I. N., Karningsih, P. D., &Nasution, A. H. (2018, April). Single Stage Queuing/Manufacturing System Model that Involves Emission Variable. In IOP Conference Series: Materials Science and Engineering (Vol. 337, No. 1, p. 012008). IOP Publishing.*
7. *Chen, C., & Tiong, L. K. (2018). Using Queuing Theory and Simulated Annealing to Design The Facility Layout in An AGV-Based Modular Manufacturing System. International Journal of Production Research, 1-18.*
8. *Zhang, R., Rossi, F., &Pavone, M. (2019). Analysis, Control, and Evaluation of Mobility-on Demand Systems: a Queuing-Theoretical Approach. IEEE Transactions on Control of Network Systems, 6(1), 115-126.*
9. *Malik, S., Akhtar, N., Shoukat, S., & Javed, S. A Study of Different Factors in Single and Multiple Server Queuing Models in Banking System.*
10. *Nguyen, B. M., Tran, D., & Nguyen, G. (2016). Enhancing Service Capability with Multiple Finite Capacity Server Queues in Cloud Data Centers. Cluster Computing, 19(4), 1747-1767.*

DOI No. : 10.5958/2394-9309.2022.00028.2
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

MODELLING LONGITUDINAL DATA - A COMPARATIVE STUDY

S. Dharani*, S. Abeena Shantini**
*PG Department of Biostatistics, S.D.N.B.V.College For Women, Chrompet, Chennai.
**Assistant Professor, PG Department of Statistics, S.D.N.B.V.College For Women, Chrompet, Chennai.
E-mail : dharanishankar1411@gmail.com, abeeshan777@gmail.com

ABSTRACT

A longitudinal study (or longitudinal survey, or panel study) is a research design that involves repeated observations of the same variables over short or long periods of time. It is often a type of observational study, although they can also be structured as longitudinal randomized experiments.

METHOD:

The research aims to compare the performance of the various models namely Generalized Linear Model (GLM), Linear Mixed Effect Model (LMEM) and Generalized Estimating Equation (GEE) on a longitudinal data. To evaluate the fitting of the model the data were taken from a clinical trial of secondary data of 16 rats carried out by Hand and Crowder (1996) describes data on the body weights of rats measured over 64 days. The body weights of the rats (in grams) are measured on day 1 and every seven days thereafter until day 64, with an extra measurement on day 44. The experiment started several weeks before "day 1." There are three groups of rats, each on a different diet. Thus the data can be viewed as a balanced longitudinal data with repeated measurements. . Data was analyzed in RStudio using Descriptive measures, Correlation, Box Plot, Plot, GLM, LME, GEE.

RESULT:

A Generalized Linear Model was fit assuming independence of response variables with exponentially family error distribution for the data. The parameter estimates for models with interaction and without interaction are obtained along with AIC and BIC values. For the models without interaction performed better than the models with interaction. For linear mixed effect model the AIC criterion suggests that Model 2 (i.e.) Model with Interaction fits the data better, while using BIC criterion it may inferred that the Model 1 (i.e.) Model without Interaction performs better.

Keywords: Longitudinal study, GLM, LMEM, GEE

1. INTRODUCTION

LONGITUDINAL STUDIES

A **longitudinal study** (or **longitudinal survey**, or **panel study**) is a research design that involves repeated observations of the same variables over short or long periods of time. It is often a type of **observational study**, although they can also be structured as longitudinal randomized experiments. Longitudinal studies are often used in social-personality and clinical psychology, to study rapid fluctuations in behaviours, thoughts, and emotions from moment to moment or day to day;

in developmental psychology, to study developmental trends across the life span; and in sociology, to study life events throughout lifetimes or generations

Longitudinal studies are designed to measure intra-individual change over time. Repeated observations are made on individual subjects, usually at a set of common time points specified by the study protocol. A main objective of longitudinal studies is to relate change over time in individuals to their characteristics (exposure, sex, etc.), or to an experimental condition (drug treatment arm, time since baseline, etc.). In some studies, exposures or experimental conditions may change during the course of the study (as in crossover designs or repeated measures experiments). Outcomes may be measurements, counts, or dichotomous indicators, and may have multivariate outcomes measured at each of several occasions as well. In the ideal setting, one may have all subjects measured at the same set of occasions; this greatly facilitates the analysis and interpretation.. To some extent, standard univariate regression models and methods can be used to analyze longitudinal data, provided one uses the proper design matrices and takes into account the fact that the observations on individuals are correlated. This approach is the basis of the General Estimating Equations (GEE) approach (Diggle et al., 1994). The objective of a statistical analysis of longitudinal data is usually to model the expected value of the response variable as either a linear or nonlinear function of a set of explanatory variables. Statistical analysis of longitudinal data requires an accounting for possible between-subject heterogeneity and within-subject correlation. The model can also be used for the analysis of clustered data; here the sampling unit is the cluster and the repeated observations are individuals in the cluster.

. Longitudinal studies can be **retrospective** (looking back in time, thus using existing data such as medical records or claims database) or **prospective** (requiring the collection of new data)

2. OBJECTIVE OF THE STUDY

To compare the performance of the various models namely Generalized Linear Model (GLM), Linear Mixed Effect Model (LMEM) and Generalized Estimating Equation (GEE) for Longitudinal data.

3. REVIEW OF LITERATURE

Wei Pan Biometrics, Vol. 57, No. 1. (Mar., 2001), pp. 120-125 works on “Akaike's Information Criterion in Generalized Estimating Equations”. The study goal was to determine the risk factors for diabetic retinopathy. The binary response is the presence of diabetic retinopathy in each of two eyes from each of 720 individuals in the study. That was found that eight of them are marginally associated with the response variable. Barnhart and Williamson included only four risk factors, i.e., duration of diabetes (years), glycosylated haemoglobin level, diastolic blood pressure, and body mass index, plus the two quadratic terms such as duration of diabetes and body mass index in their final model . Due to the nature of the possible correlation between the two observations on the two eyes from the same participant, GEE is used to fit the marginal logistic regression model and QIC is applied to do model selection, all under the working independence model. According to the QIC values, the top four models are very close but different from Barnhart and Williamson's model in that proteinuria is included in the former four models.

Barnett et al (2010) works on "Using information criteria to select the correct variance-covariance structure for longitudinal data in ecology Article in Methods in Ecology and Evolution". This paper contains the Ecological data sets often use clustered measurements or use repeated sampling in a longitudinal design. Choosing the correct covariance structure is an important step in the analysis of such data, as the covariance describes the degree of similarity among the repeated observations. Three methods for choosing the covariance are: the Akaike information criterion (AIC), the quasi-information criterion (QIC), and the deviance information criterion (DIC). We compared the methods using a simulation study and using a data set that explored effects of forest fragmentation on avian species richness over 15 years. The overall success was 80.6% for the AIC, 29.4% for the QIC and 81.6% for the DIC. For the forest fragmentations study the AIC and DIC selected the unstructured covariance, whereas the QIC selected the simpler autoregressive covariance. Graphical diagnostics suggested that the unstructured covariance was probably correct. Author recommends using DIC for selecting the correct covariance structure.

Thall P.F., and Vail S.C (1990) ,Vol. 46,No. 3,657-671 works on "Some covariance Models for longitudinal count data with over dispersion Biometrics". The work presents covariance models for longitudinal count data with predictive covariates. The models accounts for over dispersion, heteroscedasticity and dependence among repeated observations. Quasi-likelihood regression approach similar to the equation formulation by Liang and Ziger was adopted. Generalized estimating equations for both the covariates parameter and variance covariance parameter are analyzed. A longitudinal count data was taken to test the proposed model and the data accounts for over dispersion hence the estimated value deviates from the predicted value due the nature of the data and main focus is on the study of decline of counts over a period of time. This phenomenon will arise almost invariably in any setting where a predictive model is fit to the data. Thus the proposed model accounts for over dispersed data.

4. DATA SOURCE AND METHODOLOGY

DATA SOURCE :

The data is retrived from the RStudio using nlme package.

TYPES OF STUDY: Longitudinal Study

SAMPLE SIZE: 16

VARIABLES USED:

- **Weight** : a numeric vector giving the body weight of the rat (grams).
- **Time** : a numeric vector giving the time at which the measurement is made (days)
- **Rat :** an ordered factor with levels

2 < 3 < 4 < 1 < 8 < 5 < 6 < 7 < 11 < 9 < 10 < 12 < 13 < 15 < 14 < 16

Identifying the rat whose weight is measured.

- **Diet** : a factor with levels 1 to 3 indicates the diet that the rat receives.

BIOSTATISTICAL TOOLS:

- DESCRIPTIVE MEASURES
- CORRELATION
- BOX PLOT
- GLM
- LME
- GEE

PACKAGE USED:

- RStudio
- SPSS

5. DATA SOURCE:

The data is come from a clinical trial of 16 rats carried out by Hand and Crowder (1996) describes data on the body weights of rats measured over 64 days. The body weights of the rats (in grams) are measured on day 1 and every seven days thereafter until day 64, with an extra measurement on day 44. The experiment started several weeks before "day 1." There are three groups of rats, each on a different diet. Thus the data can be viewed as a balanced longitudinal data with repeated measurements.

6. ANALYSIS AND INTERPRETATION

GENERAL DESCRIPTIVE

The Exploratory analysis of the Body weight data is given the following Table

Table Data Descriptive for Diet 1group (n=8)

VARIABLE	MEAN	S.D	VARIANCE
DAY_1	250.625	15.22158	231.6964
DAY_8	258.75	15.52648	241.0714
DAY_15	259.75	15.52648	241.0714
DAY_22	261.875	13.60081	184.9821
DAY_29	264.625	11.05748	122.2679
DAY_36	265	11.78377	138.8571
DAY_43	267.375	10.95364	119.9821
DAY_44	267.25	10.19454	103.9286
DAY_50	269.5	14.56022	212
DAY_57	271.5	10.75706	115.7143
DAY_64	273.75	12.44129	154.7857

Table Data Descriptive for Diet 2 group (n=4)

VARIABLE	MEAN	S.D	VARIANCE
DAY_1	407.5	137.2042	18825
DAY_8	413.75	136.8317	18722.92
DAY_15	414.75	136.8317	18722.92
DAY_22	426	144.0185	20741.33
DAY_29	433.25	144.4608	20868.92
DAY_36	438	147.4879	21752.67
DAY_43	436.75	145.1858	21078.92
DAY_44	438.25	145.0342	21034.92
DAY_50	446.75	154.8276	23971.58
DAY_57	453.75	153.6107	23596.25
DAY_64	460.25	159.0479	25296.25

Table Data Descriptive for Diet 3 group (n=4)

VARIABLE	MEAN	S.D	VARIANCE
DAY_1	508.75	27.80138	772.9167
DAY_8	506.25	28.39454	806.25
DAY_15	507.25	28.39454	806.25
DAY_22	518.25	24.54078	602.25
DAY_29	523.75	25.07821	628.9167
DAY_36	529.25	24.40458	595.5833
DAY_43	522.75	20.59733	424.25
DAY_44	530	18.4029	338.6667
DAY_50	538.25	21.25049	451.5833
DAY_57	542.5	17.0196	289.6667
DAY_64	550.25	18.89224	356.9167

INTERPRETATION:

The average body weight is greater in Diet 3 when compared to the Diet 1 and Diet 2.The increase in average body weight from the Day 1 to the Day 64 is 41.5grms in the Diet 3 group where as it is 52.75grms in Diet 2 and 23.125grms in Diet 1. Therefore it can be concluded that the Diet 2 has an impact on the increase of body weight.

CORRELATION ANALYSIS

The following Table shows the correlation between the variables of interest, which are the repeated measures from the same rat over a period of 64 days.

Table Correlation among the weight taken at different time points (Diet 1)

	Day_1	Day_8	Day_15	Day_22	Day_29	Day_36	Day_43	Day_44	Day_50	Day_57	Day_64
Day_1	1	.608	.739*	.811*	.825*	.884**	.872**	.804*	.888**	.848**	.793*
Day_8	.608	1	.737*	.723*	.717*	.617	.747*	.688	.714*	.804*	.856**
Day_15	.739*	.737*	1	.960**	.960**	.898**	.911**	.930**	.951**	.929**	.924**
Day_22	.811*	.723*	.960**	1	.988**	.949**	.950**	.958**	.986**	.980**	.965**
Day_29	.825*	.717*	.960**	.988**	1	.963**	.957**	.951**	.987**	.982**	.963**
Day_36	.884**	.617	.898**	.949**	.963**	1	.979**	.967**	.975**	.948**	.898**
Day_43	.872**	.747*	.911**	.950**	.957**	.979**	1	.980**	.971**	.971**	.940**
Day_44	.804*	.688	.930**	.958**	.951**	.967**	.980**	1	.962**	.951**	.925**
Day_50	.888**	.714*	.951**	.986**	.987**	.975**	.971**	.962**	1	.980**	.952**
Day_57	.848**	.804*	.929**	.980**	.982**	.948**	.971**	.951**	.980**	1	.991**
Day_64	.793*	.856**	.924**	.965**	.963**	.898**	.940**	.925**	.952**	.991**	1

Table Correlation among the weight taken at different time points (Diet 2)

	Day_1	Day_8	Day_15	Day_22	Day_29	Day_36	Day_43	Day_44	Day_50	Day_57	Day_64
Day_1	1	.999**	.998**	.996**	.995**	.991**	.990**	.987*	.988*	.986*	.983*
Day_8	.999**	1	1.000**	.999**	.998**	.996**	.995**	.993**	.994**	.993**	.991**
Day_15	.998**	1.000**	1	1.000**	.999**	.997**	.996**	.995**	.996**	.994**	.993**
Day_22	.996**	.999**	1.000**	1	1.000**	.999**	.998**	.997**	.998**	.997**	.995**
Day_29	.995**	.998**	.999**	1.000**	1	.999**	.999**	.998**	.998**	.998**	.996**
Day_36	.991**	.996**	.997**	.999**	.999**	1	1.000**	1.000**	1.000**	.999**	.999**
Day_43	.990**	.995**	.996**	.998**	.999**	1.000**	1	1.000**	1.000**	.999**	.999**
Day_44	.987*	.993**	.995**	.997**	.998**	1.000**	1.000**	1	1.000**	1.000**	.999**
Day_50	.988*	.994**	.996**	.998**	.998**	1.000**	1.000**	1.000**	1	1.000**	1.000**
Day_57	.986*	.993**	.994**	.997**	.998**	.999**	.999**	1.000**	1.000**	1	1.000**
Day_64	.983*	.991**	.993**	.995**	.996**	.999**	.999**	.999**	1.000**	1.000**	1

Table Correlation among the weight taken at different time points (Diet 3)

	Day_1	Day_8	Day_15	Day_22	Day_29	Day_36	Day_43	Day_44	Day_50	Day_57	Day_64
Day_1	1	.974*	.967*	.936	.929	.927	.814	.769	.855	.805	.766
Day_8	.974*	1	.997**	.978*	.988*	.983*	.913	.880	.938	.905	.767
Day_15	.967*	.997**	1	.963*	.991**	.992**	.933	.904	.956*	.927	.807
Day_22	.936	.978*	.963*	1	.967*	.946	.867	.834	.890	.858	.617
Day_29	.929	.988*	.991**	.967*	1	.997**	.961*	.939	.975*	.956*	.774
Day_36	.927	.983*	.992**	.946	.997**	1	.972*	.952*	.986*	.967*	.821
Day_43	.814	.913	.933	.867	.961*	.972*	1	.997**	.997**	1.000**	.828
Day_44	.769	.880	.904	.834	.939	.952*	.997**	1	.989*	.998**	.815
Day_50	.855	.938	.956*	.890	.975*	.986*	.997**	.989*	1	.996**	.845
Day_57	.805	.905	.927	.858	.956*	.967*	1.000**	.998**	.996**	1	.830
Day_64	.766	.767	.807	.617	.774	.821	.828	.815	.845	.830	1

Interpretation

It is evident that correlation between the body weights measured on different days for all the three Diets are highly statistically significant.

DISPERSION IN THE STUDY VARIABLE

Box plot representing the dispersion on the study variable. The following plot shows the weight of Rats day wise for the three Diets.

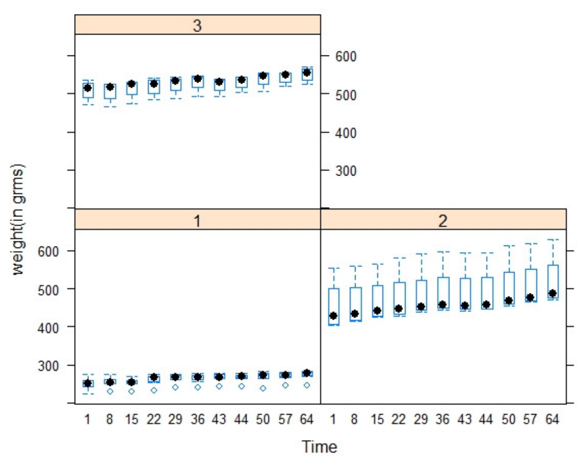


Figure Box plot of weight of Rats daywise for the three Diets.

Interpretation:

The above Box plot reveals that there is over dispersion in average body weight for Diet 2 on all days as compared with those of Diet 1 and Diet 2.

PLOT

The figure 3.2 shows the changes in the Body weight over a period of 64 days for all sixteen rats. The Rat 1 to 8 received Diet 1, Rat 9 to 12 received Diet 2 and Rat 13 to 16 received Diet 3.

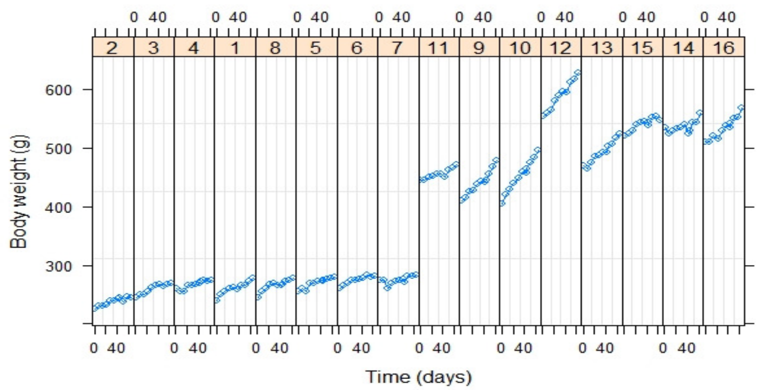


Figure 3.2

Interpretation:

From the above plot, it is observed that the increase in average body weight seems to be steep for Rats on Diet 2 followed by Rats on Diet 3 than those on Diet 2.

PARAMETER ESTIMATION USING GLM:

The generalized linear model has been constructed for the body weight data both with interaction and without interaction using R codes given in Appendix 1, the results of which are presented below.

Table Model 1: Model without interaction

Variable	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	244.0689	5.7725	42.281	< 2e-16 ***
Time	0.5857	0.1331	4.402	1.88e-05 ***
Diet 2	220.9886	6.3402	34.855	< 2e-16 ***
Diet 3	262.0795	6.3402	41.336	< 2e-16 ***

Table Model 2: Model with interaction

Variable	Estimate	Std. Error	t- value	Pr(>|t|)
Intercept	251.6517	7.2635	34.646	<2e-16 ***
Time	0.3596	0.1873	1.920	0.0565
Diet 2	200.6655	12.5807	15.950	<2e-16 ***
Diet 3	252.0717	12.5807	20.036	<2e-16 ***
Time:Diet2	0.6058	0.3244	1.867	0.0636 .
Time:Diet3	0.2983	0.3244	0.920	0.3591

The following Table shows the value of information criteria AIC and BIC for models with and without Interactions.

Table Information Criteria for various Model

	DF	AIC	BIC
Model 1	4	1750.19	1766.043
Model 2	6	1750.516	1772.709

Inference:

The point estimates and the standard error for the models with interaction and for the model without interpretation vary slightly. The factors Time, Diet 2 and Diet 3 are statistically significant for the Model without Interaction. In the Model with Interaction, the factors Diet 2 and Diet 3 turns out to be statistically significant where as Time with Diet 2 and Time with Diet 3 are not statistically significant.

On comparison of the models on the basis of Information Criteria AIC and BIC shows that the model without interaction is the better fit to the Body Weight.

PARAMETER ESTIMATION USING LMM:

The generalized linear model has been constructed for the body weight data both with interaction and without interaction using R codes given in Appendix 2, the results of which are presented below.

Table Model1: Model Without Interaction

Variable	Value	Std. Error	DF	t-value	p-value
Intercept	246.45734	13.083678	159	18.837007	0***
Time	0.58568	0.088283	159	6.634179	0***
Diet2	214.58718	22.387529	13	9.585121	0***
Diet3	258.92723	22.387529	13	11.565691	0***

Table Model2: Model With Interaction

Variable	Value	Std. Error	DF	t-value	p-value
Intercept	251.65165	13.094025	157	19.218816	0.0000***
Time	0.35964	0.091140	157	3.946019	0.0001***
Diet2	200.66549	22.679516	13	8.847873	0.0000***
Diet3	252.07168	22.679516	13	11.114509	0.0000***
Time:Diet2	0.60584	0.157859	157	3.837858	0.0002***
Time:Diet3	0.29834	0.157859	157	1.889903	0.0606

The following Table shows the value of information criteria AIC and BIC for models with and without Interactions.

Table Information Criteria for various Model

	DF	AIC	BIC
Model 1	4	1175.335	1200.515
Model 2	6	1171.720	1203.078

Inference:

The point estimates and the standard error for the models with interaction and for the model without interaction vary slightly. For the Model without Interaction the factors Time, Diet 2 and Diet 3 are statistically significant, whereas in the Model with Interaction, the factors Time, Diet 2 and Diet 3 and Time with Diet 2 are statistically significant where as Time with Diet 3 are not statistically significant in the with interaction model.

The AIC criterion suggest that Model 2 (i.e.) Model with Interactions fits the data better, while using BIC criterion it may inferred that the Model 1 (i.e.) Model without Interaction performs better.

PARAMETER ESTIMATION USING GEE

Parameter Estimation using Generalized Estimating Equation with quasi likelihood function has been done for Body weight data, the results of which are presented below.

Table Model 1: GEE Model using Independent Covariance Structure

Variable	Estimate	Std. Error	Wald	Pr(>|W|)
Intercept	244.06890	4.94385	2437.22	< 2e-16 ***
Time	0.58568	0.08548	46.95	7.30e-12 ***
Diet2	220.98864	30.89378	51.17	8.48e-13 ***
Diet3	262.07955	10.33385	643.19	< 2e-16 ***

Table Model 2: GEE Model using Autoregressive Order 1 Covariance Structure

Variable	Estimate	Std. Error	Wald	Pr(>|W|)
Intercept	243.5067	5.3184	2096.3	< 2e-16 ***
Time	0.5770	0.0864	44.6	2.4e-11 ***
Diet2	223.7082	30.8354	52.6	4.0e-13 ***
Diet3	266.9057	10.4153	656.7	< 2e-16 ***

Table Model 3: GEE Model using Exchangeable Covariance Structure

Variable	Estimate	Std. Error	Wald	Pr(>|W|)
Intercept	244.0689	4.9438	2437.2	< 2e-16 ***
Time	0.5857	0.0855	47.0	7.3e-12 ***
Diet 2	220.9886	30.8938	51.2	8.5e-13 ***
Diet3	262.0795	10.3339	643.2	< 2e-16 ***

The following Table shows the values of information criteria QIC, QICu and CIC for the different correlation structure.

Table Information Criteria for the Three Models

Correlation Structure	QIC	QICu	CIC
Independent	202897.4	202823.5	40.9
Auto Regressive (1)	203743.3	203740.3	5.5
Exchangeable	548.57	547.87	3.35

Inference:

Modelling body weight data for Rats using different covariance structures are considered, it is inferred that the factors, Time, Diet 2 and Diet 3 are statistically significant at α (0.05) level of significance in all the three models. Comparison of models based on the information criteria QIC and QICu indicates that, the model with Exchangeable covariance structure ranks first, followed by independent and AR (1). Whereas the CI Criterion seems to suggest that Independent Covariance structure is best suited for modelling this data.

FINDINGS AND CONCLUSION

Basic Descriptive measures are performed for the Rat's body weight dataset on different Diets. It shows that the average body weight is greater in Diet 3 when compared to the Diet 1 and Diet 2. The increase in body weight from the Day 1 to the Day 64 is 41.5grms in the Diet 3 group where as it is 52.75grms in Diet 2 and 23.125grms in Diet 1. Therefore it can be concluded that the Diet 3 has more impact on the increase of body weight.

Correlation was performed to find the correlation between the variables of interest, which are the repeated measures from the same rat over a period of 64 days for every seven days thereafter until day 64, with an extra measurement on day 44. From this, we conclude that correlation between the body weight measured on different days for all the three diets are highly statistically significant. The Box plot reveals that there is an over dispersion in average body weight for Diet 3 on all days as compared with those of Diet 1 and Diet 2.

From plot, it is observed that the increase in average body weight seems to be steep for Rats on Diet 2 followed by Rats on Diet 3than those on Diet 2.

From the Generalized Linear Model**,** the point estimates and the standard error for the models with interaction and for the model without interaction vary slightly. The factors Time, Diet 2 and Diet 3 are statistically significant for the Model without Interaction. In the Model with Interaction, the factors Diet 2 and Diet 3 turns out to be statistically significant where as Time, Time with Diet 2 and Time with Diet 3 are not statistically significant.

On comparison of the models on the basis of Information Criteria AIC and BIC shows that the model without interaction is the better fit to the Body Weight.

From the Linear Mixed Model, the point estimates and the standard error for the models with interaction and for the model without interaction vary slightly. For the Model without Interaction the factors Time, Diet 2 and Diet 3 are statistically significant, whereas in the Model with Interaction, the factors Time, Diet 2 and Diet 3 and Time with Diet 2 are statistically significant where as Time with Diet 3 are not statistically significant in with interaction model.

The AIC criterion suggest that Model 2 (i.e.) Model with Interactions fits the data better, while using BIC criterion it may inferred that the Model 1 (i.e.) Model without Interaction performs better.

Modelling body weight data for Rats using different covariance structures are considered, it is inferred that the factors, Time, Diet 2 and Diet 3 are statistically significant at α (0.05) level of significance in all the three models. Comparison of models based on the information criteria QIC and QICu indicates that, the model with Exchangeable covariance structure ranks first, followed by independent and AR (1). Whereas the CI Criterion seems to suggests that Independent Covariance structure is best suited for modelling this data.

BIBLIOGRAPHY

1. *Zeger S.L & Liang K.Y.& Albers P.S (1988) : Models for longitudinal data: A generalized estimating equation approach,Biometrics,Vol. 44,No.4,pp. 1049-1060*
2. *Wei Pan(2001): Akaike's Information Criterion in Generalized Estimating Equations,Biometrics,Vol. 57,No. 1.PP. 120-125*
3. *Verbeke G, and Molenberghs G, (2000), Linear Mixed Models for Longitudinal Data, Springer.*
4. *Ballinger G.A (2004): Using generalized Estimating Equations for Longitudinal Data analysis., Organizational Research Methods, 7:127-150*
5. *Carey V.J and Wang Y.G (2011): Working covariance model selection for generalized equations,(2011),Statistics in Medicine,DOI: 10.1002/sim.4300*

DOI No. : 10.5958/2394-9309.2022.00029.4
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Online) : 2394-9309

MODEL EVALUATION AND CLASSIFICATION OF TAMIL ARTICLES USING ORANGE DATA MINING TECHNIQUES

A. M. Priyanga[*] and Manimannan G.[**]
[*]Assistant Professor, Department of Mathematics, Apollo Arts and Science College Chennai, Chennai
[**]Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai, Chennai

ABSTRACT

The present paper deals with classification and cross validation of articles of three known authors written by contemporary Tamil scholars of the same period, namely Mahakavi Bharathiar (MB), Subramniya Iyer (SI), and T. V. Kalyanasundaranar (TVK). These three popular scholars had written number of articles on India's Freedom Movement during the pre-independence period and published in a magazine called, India. This research paper discusses the assignment of articles of ambiguous to Mahakavi Bharathiar (MB), Subramaniya Iyer (SI), and T. V. Kalyanasundaranam (TVK). The application of machine learning methods of Neural Network, AdaBoost, Support Vector Machine and Random Forest as data mining tools to explore the classification model and cross validate in the present database structure. Classification models were applied and extracted unique results for all authors. Area under the Curve, Classification Accuracy, F1 Score, Precision and Recall are all closer to unity. The results of the present study indicates that machine learning algorithms of Data Mining tools can be used as a feasible tool for analysis of large set of author's data. Finally, the three author's model writing structure and their stylistic features are classified, cross validated and visualized using scatter and bar plots.

Keywords: Authorship, Stylistic Features, Machine Learning Algorithms, Test and Score, Confusion Matrix, Distribution and Scatter Plot.

1. INTRODUCTION

Stylometry is a study of quantifiable of human language or the statistical analysis of literary style (Holmes, 1995; Holmes and Forsyth, 1995). This involves attempting to formally capture the creative, unconscious elements of language particular to individual writers and speakers. Although researchers have studied writing for centuries, the discipline of stylometry is fairly recent, while its origins date back to the late 19th century, the field as it is now began with work on the Federalist Papers in 1968 (Mosteller and Wallace, 1968).

The problem of authorship attribution began with discipline known as Stylometry. Stylometry mainly concerns itself with authorship attribution studies, although chronological studies on the dating of work within the corpus of an author have also investigated. Writing in a forensic background, Bailey (1979) proposed three rules to define the situation necessary for authorship attribution:

1. The number of putative authors should constitute a well- defined set.
2. The lengths of the writings should be sufficient to reflect the linguistic behavior of the author of the disputed text and also those of the candidates.
3. The texts used for comparison should be commensurate with the disputed writing.

A computational stylistic study of doubtful authorship should involve comparisons of the disputed text with works by each of the possible candidate authors using suitable statistical tools on quantifiable features of the texts-features which reflect the style of the writing as defined above. One modern addition to the tools available for computational stylometry is that of the Artificial Neural Network (ANN). These are the computational methods closely based on the concept of biological neuron, idea being that simple, trained processing elements will result in much more difficult behavior when used in combination. In recent years, many scholars have successfully demonstrated that this technique of machine learning field can be applied to authorship attribution. Merriam and Mathews (1993, 1994) have trained a multi-layer perceptron network to distinguish the works of Shakespeare and Marlowe.

Tweedie et al. (1996) have provided a useful review of the applications of ANNs in the area of computational stylometry and have used this machine-learning package for reanalysis of Federalist Papers. Kjell (1994) have taken up authorship study using letter-pair frequency features with neural network classification. Recently, authorship identification problem is also attempted by the authors using Radial Basis Function Network (Chandrasekaran and Manimannan, 2008). The present study attempts to use Machine Learning Algorithm models for prediction, cross validation and Classification for authorship database.

2. REVIEW OF LITERATURE

In recent days, data science plays a vital role of almost all fields of research, Statistics, Engineering, Computer Science, Text Mining, Marketing and Medical Sciences, etc. Machine Learning Models are used for prediction, classification and cross validation of given database in any field of data science. The following section describes the literature review of various machine learning models. The discipline associated with *Stylometry* paved path to the problem of authorship attribution

Tweedie *et. al*. (1996) discussed the application of neural networks in stylometry and their usefulness for a number of reasons:

1. Neural Networks can learn from the data themselves. Implementing a rule-based system in linguistic computing may become complex as the number of distinguishing variables increases and even the most complex rules may still not be good enough to completely characterize the training data. In essence, neural networks are more adaptive.
2. Neural networks can generalize. This ability is particularly required in literary field, as only limited data may be available.

3. Neural networks can capture non-linear interactions between input variables.
4. Neural networks are capable of fault tolerance. Hence a particular work, which is not in line with the usual writing style of an author, will not affect the network to a considerable extent. Thus neural networks appear to promise much for the field of stylometry. Their application would appear to be worthy of investigation.

The pioneering work in the application of neural nets in Stylometry was undertaken by Merriam and Mathews (1993). In their paper, a very small set of function word frequencies is used as input to a multiplayer perceptron (a neural net having a hidden layer) to examine four plays that have been attributed both to Shakespeare and John Fletcher.

In the initial stage the linguistic studies have often been made based on either complete or on non-probabilistic sampling methods (Butler, 1985).Later, block sampling or spread sampling on the type of problems under study. The major disadvantage of this sampling method is that the block chosen may be a representative in a number of ways (Holmes. 1985). To avoid this kind of bias a sample has to be drawn as uniformly as possible over the whole texts in a methodological way (Herdan, 1985). Nowadays, the availability of computer readable text corpora makes all the literary studies completes enumerative studies. The entire texts have been divided into different blocks of certain number of words and each block is considered as a sample. Manimannan G. and Bagavandas M (2004) have considered each sentence of a text as a sample in their complete enumerative literary studies.

A computational stylistic study of doubtful authorship should involve comparisons of the disputed text with works by each of the possible candidate authors using suitable statistical tools on quantifiable features of texts-features which reflect the style of writing as defined above. One modern addition to the tools available for computational stylometry is Artificial Neural Network (ANN). It is a computational method closely based on the concept of biological neuron, idea being that simple, trained processing elements will result in much more difficult behavior when used in combination.

The present study attempts to use the Orange Data Mining tools like, Neural Network, AdaBoost, Random Forest, and Support Vector Machine as suitable machine learning algorithms of data mining classification tool.

3. DATABASE

The present study deals with literary works of existing Tamil scholars, namely MB, SI and TVK. During pre-independence period, these three scholars have written a number of articles on India's freedom movement in the magazine called India. Initially, the three authors' wrote their articles by attributing their names in this magazine. Because of opposition of the British regime, the articles written on the same topic appeared in the same magazine without author's name. For the quantitative study all the attributed articles on India's freedom movement published in 1906 in the magazine India are considered. Accordingly Mahakavi Bharathiar (MB), consist of 34 blocks of sentences, the Subramaniya Iyer (SI) consists of 38 blocks of sentences and Thiru.V. Kalyanasundaranar (TVK) consists of 31 of block sentences. The data matrix consists of block of sentences of MB size of (34*24), (38*24) of SI and (31*24) of TVK respectively in case function words. The exact lists of variable of this research paper with their abbreviations are given below.

Table 1. List of Function Words with Abbreviations

Function Words	Meaning	Variable Name
மேலும்	Also	Frequency of *Also*
போன்ற	As	Frequency of *As*
ஆக	For	Frequency of *For*
என்றால்	If	Frequency of *If*
ஆக	As	Frequency of *As*
மிகவும்	Very much	Frequency of *Very much*
போல	Like	Frequency of *Like*
என்றும்	And	Frequency of *And*
மீது	On	Frequency of *on*
இருந்து	From	Frequency of *From*
எனவும்	Also	Frequency of *Also*
உள்ளே	Inside	Frequency of *Inside*
குறிக்கப்படாமல்	Unmarked	Frequency of *Unmarked*
உடன்	With	Frequency of *With*
காரணத்தினால்	Because	Frequency of *Because*
ஆம்	If	Frequency of *If*
க்கு	To	Frequency of *To*
என்னுடைய	My	Frequency of *My*
இல்	In	Frequency of *In*
உடன்	with	Frequency of *with*
ஒவ்வொரு	Every	Frequency of *Every*
குறைந்தது	Atleast	Frequency of *At least*
இன்	of	Frequency of *of*
பற்றி	About	Frequency of *About*

The main objective of this research paper is (1) to identify the authorship attribution of three scholars of the study, Namely MB, SI and TVK using twenty four function words. (ii) to identify four data mining techniques of MLP, SVM, Naïve Baye;s and RFM method of classification and (3) to visualize three authorship attribution with the help of twenty four function words.

4. METHODOLOGY

4. 1 Neural Network

A multi-layer perceptron (MLP) algorithm used with back propagation Inputs and Outputs process.

Inputs

1. Data: input dataset
2. Preprocessor: preprocessing methods

Outputs

1. Learner: multi-layer perceptron learning algorithm
2. Model: trained model

The Neural Network widget uses sklearn's Multi-Layer Perceptron Algorithm that can learn non-linear models as well as linear.

4.2 Adaptive Boosting (AdaBoost)

An ensemble meta-algorithm that combined is weak learners and adapts to the 'hardness' of each training sample.

Inputs

1. Data: input dataset
2. Preprocessor: preprocessing method(s)
3. Learner: learning algorithm

Outputs

1. Learner: AdaBoost learning algorithm
2. Model: trained model

The AdaBoost (Adaptive Boosting) widget is a machine-learning algorithm, formulated by Yoav F. and Robert S. It can be used with other learning algorithms to boost their performance. It does so by tweaking the weak learners.

4.3 Random Forest Classifier (RFC)

Predict using an ensemble of decision trees.

Inputs

1. Data: input dataset
2. Preprocessor: preprocessing method(s)

Outputs

1. Learner: random forest learning algorithm
2. Model: trained model

Random Forest is an ensemble learning method used for classification, regression and other tasks. It was first proposed by Tin Kam Ho and further developed by Leo Breiman (Breiman, 2001) and Adele Cutler. Random Forest builds a set of decision trees. Each tree is developed from a bootstrap sample from the training data. When developing individual trees, an arbitrary subset of attributes is drawn (at Random), from which the best attribute for the split is selected. The final model is based on the majority vote from individually developed trees in the forest.

Support Vector Machine

Support Vector Machines map inputs to higher-dimensional feature spaces.

Inputs

1. Data: input dataset
2. Preprocessor: preprocessing method(s)

Outputs

1. Learner: linear regression learning algorithm
2. Model: trained model
3. Support Vectors: instances used as support vectors

Support Vector Machine (SVM) is a machine learning technique that separates the attribute space with a hyperplane, thus maximizing the margin between the instances of different classes or class values. The technique often yields supreme predictive performance results. Orange embeds a popular implementation of SVM from the LIBSVM package (Orange 3). This widget is its graphical user interface. The workflow of orange machine learning algorithms is visualized in figure 1.

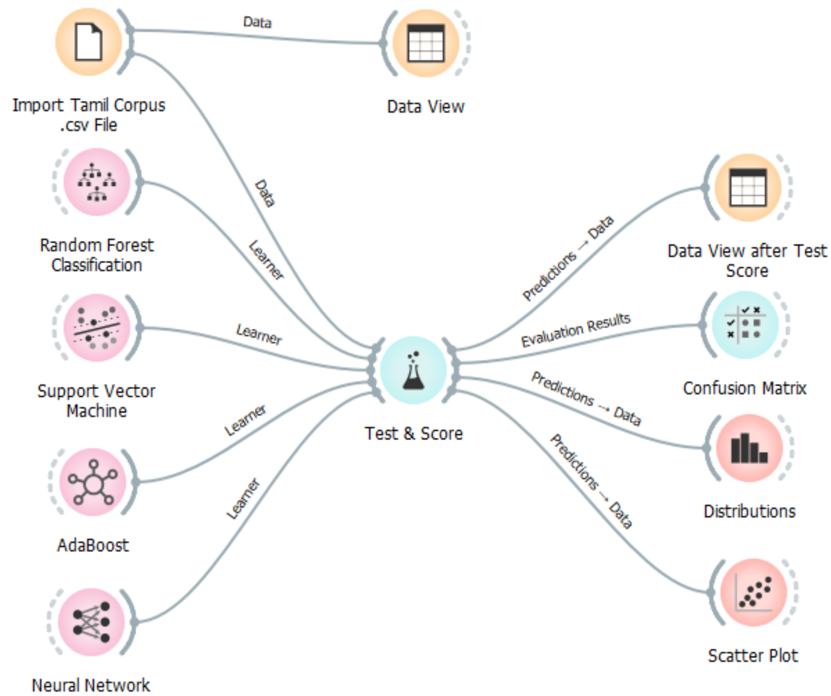
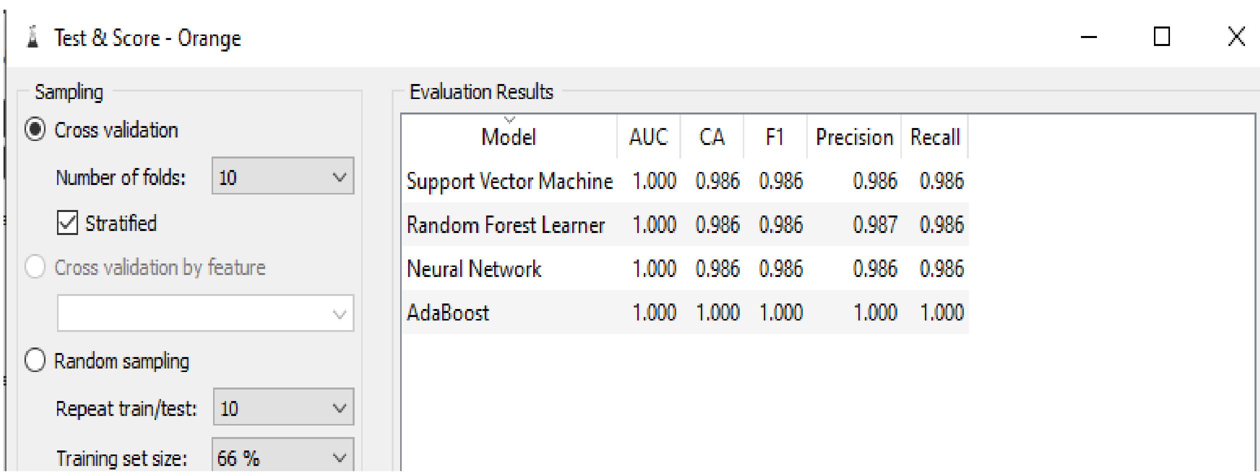


Figure 1. Work flow and widget of Machine Learning Algorithms



Model	AUC	CA	F1	Precision	Recall
Support Vector Machine	1.000	0.986	0.986	0.986	0.986
Random Forest Learner	1.000	0.986	0.986	0.987	0.986
Neural Network	1.000	0.986	0.986	0.986	0.986
AdaBoost	1.000	1.000	1.000	1.000	1.000

Table 2: Test and Score of Various Different Machine Learning Algorithms

The above table is a classification task on authorship dataset. The researcher compares the results of Neural Network with other machine learning algorithms. The AUC, CA, F1, Precession and Recall are closer to 1 and shows best classification and accuracy of this dataset. All the authors have unique writing style in the same topic of Indian independence. The Test and Score results are closer to 99 percent correct classification and remaining one percent is misclassified due to database normality or outlier. Distribution and Scatter plot are visualized in Figure 2 to 7. Classification matrix results are shows that table 3 to 6.

Table 3: Random Forest Classifier

Actual \ Predicted	MB	SI	TVK	Σ
MB	74	0	0	74
SI	0	38	0	38
TVK	0	2	29	31
Σ	74	40	29	143

Table 4: AdaBoost

Actual \ Predicted	MB	SI	TVK	Σ
MB	74	0	0	74
SI	0	38	0	38
TVK	0	0	31	31
Σ	74	38	31	143

Table 5: Neural Network

Actual \ Predicted	MB	SI	TVK	Σ
MB	74	0	0	74
SI	0	37	1	38
TVK	0	1	30	31
Σ	74	38	31	143

Table 6: Support Vector Machine

Actual \ Predicted	MB	SI	TVK	Σ
MB	74	0	0	74
SI	0	37	1	38
TVK	0	1	30	31
Σ	74	38	31	143

Original Database is visualized in figure 3 and machine learning algorithms of AdaBoost methods achieved hundred percent classification. Neural Network, Support Vector Machine and Random Forest achieved 99 percent of classification accuracy when compared to original database. In figure 2 visualization shows that the usage of function words in each corpus is unique and independent writing style of three authors.

Figure 2, Author Vs. Function Words

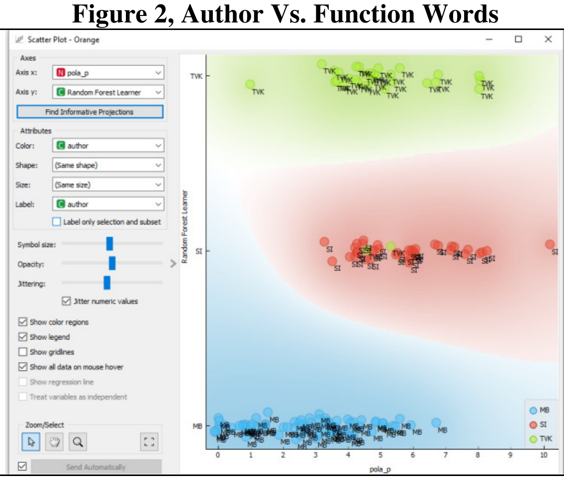


Figure 3.Author Corpus

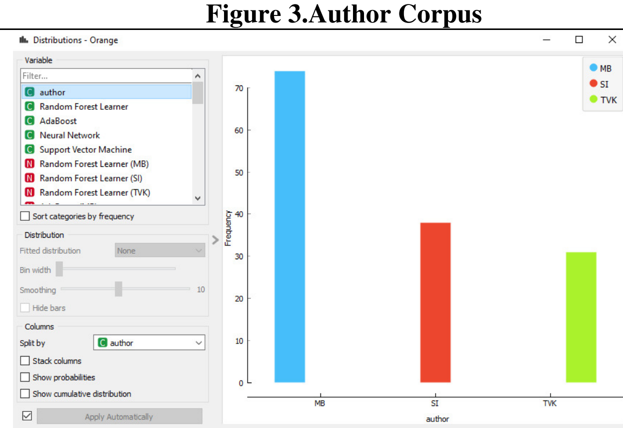


Figure 4. Random Forest Learner

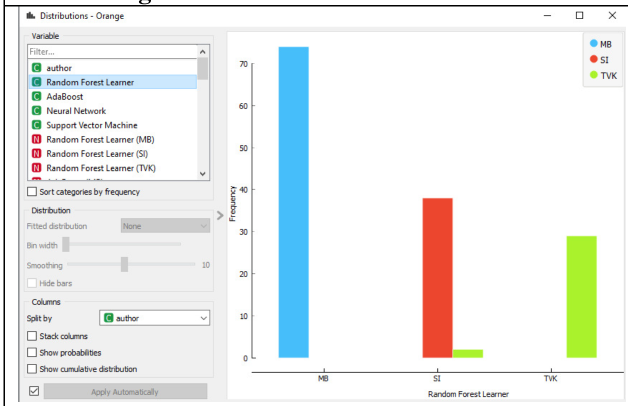
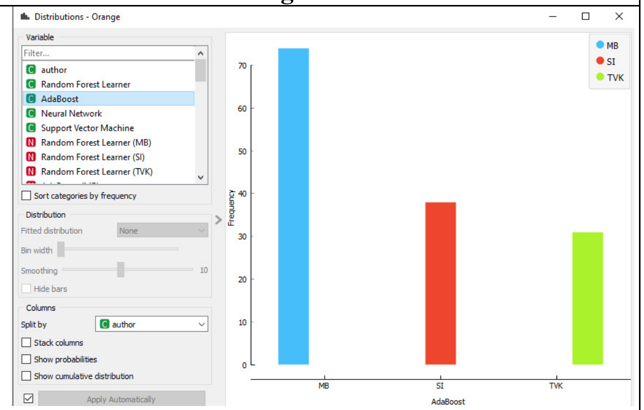


Figure 5. AdaBoost



Figur 6. Neural Network

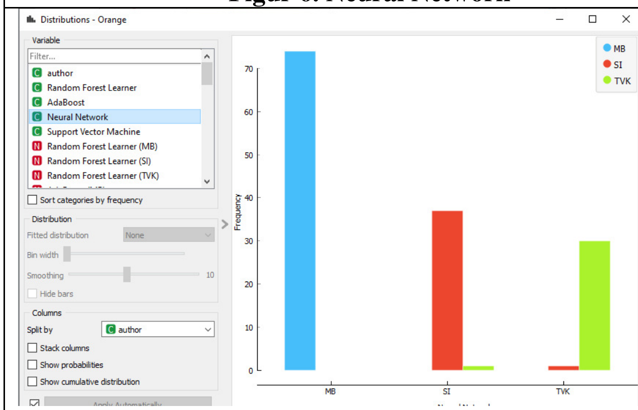
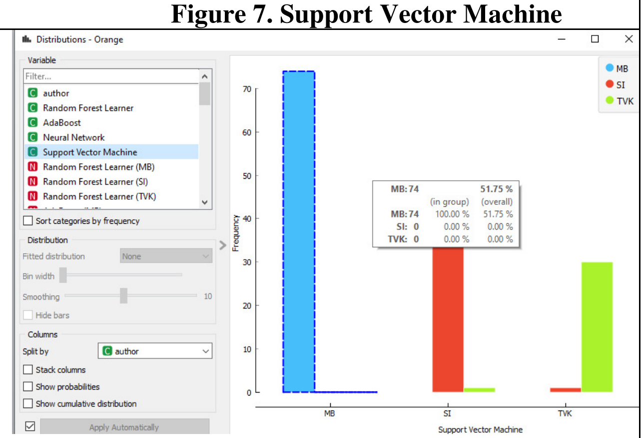


Figure 7. Support Vector Machine



5. CONCLUSION

Application of Machine Learning Algorithms models has increased considerably in areas of data mining and classification problems in the field of *Stylometry* over the last decade. The authorship attribution problem is attempted using a Machine Learning Algorithms for attributing Mahakavi Bharathiar (MB) consists of 34 blocks of sentences; the Subramaniya Iyer (SI) consists of 38 blocks of sentences and Thiru.V. Kalyanasundaranar (TVK) consists of 31 block sentences. The data matrix consists of block of sentences of MB size of (34*24), (38*24) of SI and (31*24) of TVK respectively in case function words. These results supported by the claims made by three scholars have been unique and writing style is independent on the same topic of Indian independence.

REFERENCES

1. *Bailey, R. W. (1979), The Future of Computational Stylistics, Association for Literary and Linguistic Computing Bulletin, Vol. 7, 4-11. England.*
2. *Chandrasekaran, R. and Manimannan, G. (2008), Neural Network Classification and Authorship Attribution of Articles of Unknown Authorship Using Radial Basis Function, Proceedings of the National Conference on Artificial Intelligence and Neural Networks, Department of Computer Applications, SRM University, Kattankulathur, Tamil Nadu, pp.246-256.*

3. *Holmes, D. I. (1995), The Analysis of Literary Style: A Review, Journal of Royal Statistical Society, Series A, Vol. 148, 328-334, England.*
4. *Holmes D.I. and Forsyth, R.S. (1995), The Federalist Revisited: New Directions in Authorship Attribution, Literary and Linguistic Computing, 10, 111-127, England.*
5. *Kjell, B. (1994), Authorship Determination Using Letter-pair Frequency Features with Neural Network Classifiers, Literary and Linguistic Computing, Vol.9, 119-124, England.*
6. *Merriam, T. and Mathews, R. (1993), Neural Computation in Stylometry I: An Application to the Works of Shakespeare and Fletcher, Literary and Linguistic Computing, Vol.8, 203-209, England.*
7. *Merriam, T. and Mathews, R. (1994), Neural Computation in Stylometry II: An Application to the Works of Shakespeare and Marlowe, Literary and Linguistic Computing, Vol.9, 1-6, England.*
8. *Mosteller, F. and Wallace, D. L. (1968), Inference and Disputed Authorship: The Federalist Papers, Addision-Wesley, Messachusetts.*
9. *https://orange3.readthedocs,io/projects/orange-visual programming/en/latest/widgets/model.html*

DOI No. : 10.5958/2394-9309.2022.00030.0
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

FACTORS INFLUENCING WORD OF MOUTH IN CONSUMER BUYING BEHAVIOUR OF CARS- REFERENCE TO TAMBARAM CITY

Mrs. M. Amritha Ph.D. Research Scholar* **K. Murugan** Assistant Professor**
*Department of Commerce, Sri Sankara Arts and Science College (Autonomous), Enathur, Kanchipuram.

ABSTRACT

Word of Mouth is the complete of new ideas up to our ears in numerous ways such as; telling stories, dialogs, conversations, communal experiences radio and etc., and its impact remains on our souls until we hear another story, thus, word of mouth helps us to know about what is new, and what is the latest. These days we often hear people talking about hybrid cars, some of them focus on vigor saving as an advantage, and others are focusing on it's environmentally care, because it does not contaminate the air as a result of fuel combustion, whereas, others warn of the high cost compared to regular cares. In this study the researcher finds out that various factors influencing the word of in consumer buying behaviour of cars in Tambaram City.

Key words: Buying behaviour, Word of mouth, factors, motivation, cars, Chi square Test, Purchase intention etc

1. INTRODUCTION

Now a day's buyer makes an informed decision by analysing the consumer analyses written by consumers of the product on forums, websites and other social broadcasting channels, popularly known as e-WOM (electronic Word Of Mouth).

Communication play a vital role for influencing and forming consumer outlooks and behavioural intents i.e. Word of Mouth

In present day the style of consumers shopping was changed for the technological development. With the advancement of internet, the technological developments like smart phone, IPhone and 4G and 5G internet have opened new paths for digital shopping. The consumers purchasing habits are shifted from offline mode to online mode through e-commerce.

Marketing impress the word of mouth for the results of more sales than a paid media impression and search engine optimization and most of the people are likely to trust and buy from a brand recommended by a friends and relatives or third persons.

Word of Mouth is the sound of new ideas up to our ears in several ways such as; telling stories, dialogues, conversations, shared experiences radio and etc., and its impact remains on our personalities until we hear another story, thus, word of mouth helps us to know about what is new, and what is the latest one in the world.

Word-of-mouth marketing (WOM marketing) is when a consumer's attention in a company's product or service is replicated in their daily interchanges. Fundamentally, it is free advertising motivated by customer experiences—and usually, approximately consumers go beyond what they expected in the market.

The haste of everything has become one of the features of this epoch; feast of knowledge and data, communications, e-management, e-marketing and selling and even everything, the whole practice became faster and more useful through the means of various communications. Now a day we often hear people talking about hybrid cars, some of them focus on energy saving as an advantage, and others are focusing on its environmentally care, because it does not pollute the air as a result of fuel combustion, whereas, others caution of the high cost associated to regular maintenances.In recent years, an incomparable development and improvement can be witnessed in the number of passenger cars on the roads all around the world. In this paper the researcher finds out the various factors influencing word of mouth in consumer buying behaviour of cars- reference to Tambaram city.

2. OBJECTIVES OF THE STUDY

- To study the demographic profile of the customers buying behaviour.
- To explore the marketing strategies to improve the buying habits of the consumers.
- To analyse the factors influencing the buying behaviour in the word of mouth.
- To give fruitful suggestions and conclusion for this study.

3. HYPOTHESIS OF THE STUDY

- Marital status does not influence the various factors for consumers buying behavior

4. LITERATURE REVIEW

Many researches use the 5-stage model described by (Kotler & Keller, 2012) to understand consumer behaviour and the process involved while buying a particular commodity. As per this model, the 5 stages that are involved in the consumer decision process are namely- problem recognition, information search, evaluation of alternatives, product choice and post purchase.

Applying this to the process of buying a car, we can establish the following: -Stage 1- Problem Recognition -here a consumer realizes the need to buy a car. Stage 2- Information Search -he uses different sources to find information about the cars available in the market. Stage 3- Evaluation of Alternatives-he compares several car models in terms of their price, features, deals etc. Stage 4- Product Choice – he chooses one particular car over another based on his own appraisal and reasoning. Stage 5- Post Purchase- consumer takes ownership and evaluates if the purchase has met his expectations or not. According to (Shen,1997) population growth and increasing living standards of people can be seen as the major causes for the rise in number of cars in the cities.

Based on the work of (Monga et. al 2012) it was seen that increase in disposable income was a contributing factor for purchasing a car. Their study also showed that growing family needs and an increase in the family size were some additional motives for buying a car.

In another study that was conducted in the state of Karnataka (Joseph & Kamble, 2011) found that the availability of auto finance or consumer credit is an important aspect that influences the purchase of passenger cars. Similarly, researchers (Sheik & Ali,2013) in Tirunelveli, Tamil Nadu analysed consumer preferences towards passenger cars. It was found that certain factors like price, technology and comfort were taken into account by buyers before selecting a brand.

A 2013 study was targeted towards the luxury car segment of India, which is currently on the rise as more luxury cars are entering into the Indian markets. It was found that such types of cars were mostly preferred by High-Net-Worth Individuals who wished to differentiate themselves from others. The study also found that factors relating to demographics, socio-culture and customer requirements impacted the choices for those types of cars (Verma & Rathore, 2013) . Various studies across India have reported that television advertising is one of the biggest influencers of car purchasing behaviour along with magazines, word of mouth and ratings.

In the study conducted by (Srivastava & Matta 2014) in Delhi NCR, results illustrated that many respondents relied upon their friends and advertisements in case of determining a pre-purchase. Their study also showed that the decisions of the consumers were influenced by factors like culture, family and lifestyle.

Another empirical study based in Hyderabad (Ravinder & Srikrishna, 2017), found that purchasing of a car is greatly influenced by car advertisements followed by the recommendations of family and friends. It was also observed that in case of the small car segment of India, Alto 800 was the most preferred car followed by Santro, Tata Indica and Spark. Researchers (Stella and Rajeshwari ,2012) focused on the relationship between customer satisfaction, brand image and information from mass media.

Their study showed that many consumers rely upon inputs from their relatives/friends/spouses etc for decision making. Through the work of (Sharma, 2010) it was seen that when a buyer thinks of buying a new car, it takes on an average 9.8 days to research about the product and those who have used the car will take about 7.7 days. Researchers (Kaur, Sandhu,2004) found that some of the factors considered by consumers while buying a car were safety & comfort, luxury, reliability, ease of finance, variety, colour, fuel efficiency, spaciousness and brand image. Another study found that majority of the respondents preferred sedans (61%) followed by SUVs and Hatchbacks. MUV was seen to be the least favoured car type that was preferred by only (6.4%) of the respondents (Mathur et. al, 2018).

5. RESEARCH METHODOLOGY

This research is descriptive quantitative one with analytical measurements to measure the relationship between the factors influencing word of mouth in the buying behaviour of cars reference to Tambaram city. The researcher has used for the convenience sampling technique for this research. The interviewer has collect the data for primary and secondary data. The primary data has been collected through interview schedule method. The secondary data has been collected through the books and articles available in libraries and the internet. In order to collect the primary data, the researcher has designed a questionnaire distributed for consumers in Tambaram city. The study population is the Tambaram city who are buying the cars in the word of mouth. The researcher has collected the 150 interview schedule but only using 120 interview schedules are used remaining interview schedules were not fulfilling the data.

5. LIMITATIONS OF THE STUDY

- ✓ The researcher collecting the data in the Tambaram city only.
- ✓ The researcher has collect only 210 respondents are taken for this study.
- ✓ Due to time and cost constraints the data will be collected for 210 respondents.

Analysis and interpretation

Table 1: Socio Economic conditions of the Respondents

Sl. No.	Parameters	Variables	Frequency	Percentage
1.	**Age**	Below 20	4	3.33

		20-30	38	31.67
		30-40	**52**	**43.33**
		40-50	23	19.17
		Above 50	3	2.5
2.	**Gender**	Male	**62**	**51.67**
		Female	58	48.33
3.	**Educational status**	Primary	6	5.00
		High school	41	34.17
		Intermediate	28	23.33
		Graduate and above	**45**	**37.5**
4.	**Family type**	Nuclear	**86**	**71.67**
		Joint	34	28.33
5.	**Occupation type**	Govt.Service	22	18.33
		Private Service	28	23.33
		Business	**45**	**37.5**
		House wives	14	11.67
		Students	11	9.17
6.	**Locality**	Urban	**45**	**37.5**
		Semi Urban	40	33.33
		Rural	35	29.17
7	**Income (Monthly in Rs)**	Below Rs. 20,000	25	20.83
		Rs.20,000-Rs. 40,000	**45**	**37.50**
		Rs.40,000– Rs.60,000	35	29.17
		Above Rs.60,000	15	12.50
8	**Marital status**	Married	**95**	**79.17**
		Single	25	20.83

Source: Primary Data

The Socio-Economic conditions were studied in terms of age, gender, education level, family type, occupation, Income and locality (Table 1)

Age of the respondents was recorded and it was observed that maximum respondents (43.33%) belonged to the age group of 30-40 years and minimum respondents (2.5%) belonged to the age group of above 50 years. It was observed that 51.67 per cent of the respondents were male and only 48.33per cent were female in respondents. The data pertaining to education revealed that all the respondents were educated. Maximum number of respondents (37.5%) was graduate or post graduate, 23.33 per cent respondents had qualification up to intermediate, 34.17 per cent respondents were high school pass and only 5 percent respondents had education up to primary level only. The data also revealed that 71.67 per cent respondents belonged to nuclear family and only 28.33per cent belonged to joint family. It is clear from Table 1 that maximum respondent 37.5 percent were belonged to business class, 23.33 percent belonged to private service, 18.33 percent belonged to government service, 11.67 percent belonged to house wives and minimum 9.17 percent belonged to students. It is also clear from Table 1, that 37.5 per cent respondent lived in urban areas and only 29.17 percent in rural area. The researcher further analysed that income level of the respondents most of the 37.5 per cent of the respondents earn the monthly income of Rs.20,000 to Rs. 40,000.

Table 2 - Factors Influencing The Consumer Behaviour in Word of Mouth

Sl no	Factors	No of consumers	Percentage
1	Psychological	35	29.17
2	Personal	60	50.00
3	Social	15	12.50
4	Cultural	10	8.33
Total		**120**	**100.00**

Source: Primary Data

Table 2 clearly shows that various factors influencing the word of mouth for the advertisement and purchase of goods. Out of 120 respondents 50 per cent of the respondents more influence the personal factors like lifestyle, economic circumstances etc. 29.17 per cent of the respondents influencing the psychological factors like beliefs, motivation etc followed by 12.50 per cent of the respondents influenced for the social factors like family members and the rest 8.33 per cent of the respondents say that cultural factors.

Table 3- Purchase intention for the word of mouth

Sl.no	Factors	No of consumers	Percentage
1	Purchase intention	50	41.67
2	After sales service	25	20.83
3	Price and product	15	12.50
4	Brand quality and loyalty	20	16.67
Total		**120**	**100.00**

Source: Primary data

The table 3 clearly exhibits that purchase intention of the customers for the word of mouth 41.67 per cent of the customers are purchase intention is major factor followed by 20.83 per cent of the respondents say that after sales service was good; 16.67 per cent of the customers say that brand quality and loyalty of the product and the rest 12.50 per cent of the customers said that price and product (Cheaper price).

6. VERIFICATION OF HYPOTHESIS

The researcher further analysed that and test the hypothesis that marital status does not influence the various factors for buying for consumer behaviour. The researcher using chi square test for the analysis.

Calculated Chi Square Value	P value	Degrees of freedom	Significant level	Result
4.041	7.815	5 % level	Significant	Accepted

Source: Computed Data

In this research 3 d.f at 5 % level of significance the table value is higher than the calculated value so the hypothesis is accepted. Hence concluded that Marital status does not influence the various factors for buying behaviour of the cars.

7. FINDINGS OF THE STUDY

- Most of the (43.33 per cent) respondents are come under the age group of 30 to 40 years.
- 51.67 per cent of the customers are male of these mostly married.
- 37.50 per cent of the respondents are studied graduate level, business people and earned income of Rs 20,000 to Rs.40,000 respectively.
- 71.67 per cent of the families were lived in separate i.e individual or nuclear.
- 79.17 per cent of the respondents are married those married peoples are come under the age group of 30 to 40 years.
- Majority (50 per cent) of the customers are mostly influencing for the personal factors.
- Many of the buyers are intention to purchase the goods for the necessity.

8. RECOMMENDATIONS AND CONCLUSION

Companies of all types have to be aware that word of mouth is an important part of the marketing strategy. It must build good relations with its customers to earn their loyalty, and ensure their commitment to talk about the virtues of the company and its products in all forums. Relying on word of mouth reduces marketing expenses. Managers must understand that the satisfied customer can attract new customers through word of mouth.

The researcher concluded that The impact of word of mouth on brand image was found to be very strong, therefore marketers should monitor the consumers reviews/ comments on the social media, consumer forums and blogs. The marketer can create a positive image of the brand in the mind of consumer, although the same may not lead to purchase and Also the word of mouth does not lead to purchase intentions, therefore the marketers should try to create a positive image of the brand by controlling the word of mouth.

REFERENCE

1. *Aaker, D.A. (1991), Managing Brand Equity: Capitalizing on the Value of a Brand Name, The Free Press, New York, NY.*
2. *ACNielsen (2007), Trust in Advertising: A Global Nielsen Consumer Report, October, ACNielsen, New York, NY.*
3. *Bambauer-Sachse, S. and Mangold, S. (2011), "Brand equity dilution through negative online word-of-mouth communication", Journal of Retailing and Consumer Services, Vol. 18, pp. 38-45.*

4. *Bian, X. and Moutinho, L. (2011), "The role of brand image, product involvement, and knowledge in explaining consumer purchase behaviour of counterfeits: direct and indirect effects",European Journal of Marketing, Vol. 45 Nos 1/2, p. 191.*
5. *Bickart, B. and Schindler, R.M. (2001), "Internet forums as influential sources of consumer information", Journal of Interactive Marketing, Vol. 15 No. 3, pp. 31-40.*
6. *Chatterjee, P. (2001), "Online reviews: do consumers use them?",Advances in Consumer Research, Vol. 28 No. 1, pp. 129-33.*
7. *Chevalier, J.A. and Mayzlin, D. (2006), "The effect of word of mouth on sales: online book reviews", Journal of Marketing Research, Vol. 43 No. 3, pp. 345-54.*
8. *Davis, D.F., Golicic, S.L. and Marquardt, A. (2009), "Measuring brand equity for logistics services", International Journal of Logistics Management, Vol. 20 No. 2, pp. 201-12.*

DOI No. : 10.5958/2394-9309.2022.00031.2
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ON STRONGLY*-2 DIVISOR CORDIAL LABELLING

Dr. V. M. Arul Flower Mary*, V.G. Michael Florance**
*Department of Mathematics, Holy Cross College(Autonomous), Nagercoil – 629004, Tamilnadu, India.
**Department of Mathematics, Holy Cross College(Autonomous), Nagercoil – 629004, Tamilnadu, India.
E-mail: arulflowermary@gmail.com, miclaelflorance@gmail.com

ABSTRACT
A Srongly* -2 divisor cordial labelling of a graph G with the vertex set $V(G)$ is a bijection $f: V(G) \rightarrow \{1, 2, 3, \ldots\ldots\ldots |V(G)|\}$ such that each edge uv assigned the label 1 if $\left\lfloor\frac{f(u)+f(v)+f(u)f(v)}{2}\right\rfloor$ is odd and 0 if $\left\lfloor\frac{f(u)+f(v)+f(u)f(v)}{2}\right\rfloor$ is even; then the number of edges labeled with 0 and the number of edges labeled with 1 differs by atmost 1. A graph which admits a Strongly*-2 divisor cordial labelling is called a Strongly*-2 divisor cordial graph. In this paper, we investigate the Strongly*-2 Divisor Cordial labelling of Ladder graph L_n, Slanting Ladder SL_n, Double Triangular Snake$D(T_n)$, Triple Triangular Snake $T(T_n)$, Four Triangular Snake $F(T_n)$, Double Quadrilateral Snake $D(Q_n)$ and $P_n \odot K_2$

Keywords: Function, Bijection, Cordial labelling, Strongly*-2 divisor cordial labelling.
AMS Subject Classification number (2020): 05C78

1. INTRODUCTION

Graph labelling was introduced in 1960's. We refer J. A. Gallian [5] for looking over the detailed view of graph labelling. We persueHarary [4] for fundamental terms and notations. We refer [3], [7] and [6] for further classifications of cordial labelling. In [1] and [2] we study about Strongly*- graph.
Stimulated by these, we introduce Strongly*-2 divisor Cordial labelling. In this paper, we investigate the Strongly*-2 Divisor Cordial labelling of Ladder graph L_n, Slanting Ladder SL_n, Double Triangular Snake$D(T_n)$, Triple Triangular Snake $T(T_n)$, Four Triangular Snake $F(T_n)$, Double Quadrilateral Snake $D(Q_n)$ and $P_n K_2$.

We have stated the definitions which is required for our present work.

The Ladder graph is defined as $P_n \times P_2$. Let $c_1, c_2, c_3, \ldots . c_n$ and $d_1, d_2, d_3, \ldots . d_n$ be two paths. Then the graph attained by connecting every edge c_i with d_{i+1} is called Slanting ladder graph SL_n. A double triangular snake $D(T_n)$ containing two triangular snakes that have a common path. A triple triangular snake $T(T_n)$ consists of three triangular snakes that have a common path. A Four triangular snake $F(T_n)$ consists of four triangular snakes that have a common path. A Double Quadrilateral Snake $D(Q_n)$ containing two Quadrilateral Snakes that have a common path.

2. MAIN RESULTS

Definition 2.1. A strongly*-2 divisor cordial labelling of a graph G with the vertex set $V(G)$ is a bijection $f : V(G) \to \{1,2,3 \ldots . |V(G)|\}$ such that each edge uv assigned the label 1 if $\left\lfloor \frac{f(u)+f(v)+f(u)f(v)}{2} \right\rfloor$ is odd and 0 if $\left\lfloor \frac{f(u)+f(v)+f(u)f(v)}{2} \right\rfloor$ is even;then the number of edges labeled with 0 and the number of edges labeled with 1 differs by atmost 1. The number of edgeslabeled with 0 is denoted as$e_f(0)$ and the number of edges labeled with 1 is denoted as $e_f(1)$. A graph which admits a Strongly*-2 divisor cordial labelling is called a Strongly*-2 divisor cordial graph.

Example 2.1. A Strongly*-2 Divisor Cordial labelling of the graph K_3 is shown in Figure 1

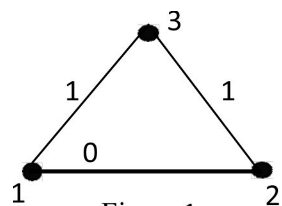


Figure 1.

Theorem 2.2. The Ladder graph L_n is a Strongly*-2 divisor cordial graph.

Proof. Let $V(L_n) = \{c_s, d_s, 1 \leq s \leq n\}$ and $E(L_n)) = \{c_s d_s \,; 1 \leq s \leq n\} \cup \{c_s c_{s+1} \,; 1 \leq s \leq n-1\} \cup \{d_s d_{s+1};\ 1 \leq s \leq n-1\}$. Then L_n has $2n$vertices and$3n-2$ edges.

Case 1: n is odd

Define $\alpha : V(L_n) \to \{1,2,3 \ldots ,2n\}$ by

$\alpha(c_{2s}) = 4s, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(c_{2s+1}) = 4s+1, 0 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(d_{2s-1}) = 4s-2, 1 \leq s \leq \left\lceil \frac{n}{2} \right\rceil$

$\alpha(d_{2s}) = 4s-1, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

Case 2: n is even

Define $\alpha : V(L_n) \to \{1,2,3 \ldots ,2n\}$ by

$\alpha(c_{2s}) = 4s, 1 \leq s \leq \frac{n}{2}$

$\alpha(c_{2s+1}) = 4s+1, 0 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s-1}) = 4s-2, 1 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s}) = 4s-1, 1 \leq s \leq \frac{n}{2}$

The values of $e_\alpha(0)$ and $e_\alpha(1)$are shown in Table 1.

Nature of n	$e_\alpha(0)$	$e_\alpha(1)$
n is odd	$\frac{3n-1}{2}$	$\frac{3n-1}{2} - 1$
n is even	$\frac{3n-2}{2}$	$\frac{3n-2}{2}$

Table 1.

By Table 1, Clearly $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Ladder graph L_n is a Strongly*-2 divisor cordial graph.

Example 2.3. A Strongly*-2 Divisor Cordial labelling of the graph L_6 is shown in Figure 2.

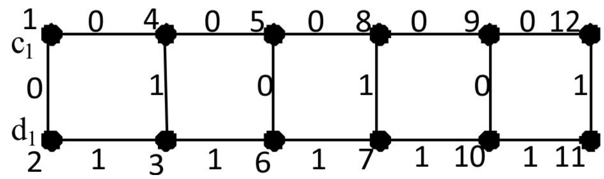


Figure 2

Theorem 2.4. The Slanting Ladder SL_n is a Strongly*-2 divisor cordial graph.

Proof. Let $V(SL_n) = \{c_s, d_s, 1 \leq s \leq n\}$ and $E(SL_n)) = \{c_s d_{s+1}\,; 1 \leq s \leq n-1\} \cup \{c_s c_{s+1}\,; 1 \leq s \leq n-1\} \cup \{d_s d_{s+1};\ 1 \leq s \leq n-1\}$. Then SL_n has $2n$vertices and$3n-3$ edges.

Case 1: n is odd

Define $\alpha : V(SL_n) \to \{1,2,3 \ldots, 2n\}$ by

$\alpha(c_{2s}) = 4s, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(c_{2s-1}) = 4s - 1, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(c_n) = 2n$

$\alpha(d_{2s+1}) = 4s + 1, 0 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(d_{2s}) = 4s - 2, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

Case 2: n is even

Define $\alpha : V(SL_n) \to \{1,2,3 \ldots, 2n\}$ by

$\alpha(c_{2s-1}) = 4s - 1, 1 \leq s \leq \frac{n}{2}$

$\alpha(c_{2s}) = 4s, 1 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s+1}) = 4s + 1, 0 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s}) = 4s - 2, 1 \leq s \leq \frac{n}{2}$

The values of $e_\alpha(0)$ and $e_\alpha(1)$ are shown in Table 2.

Nature of n	$e_\alpha(0)$	$e_\alpha(1)$
n is odd	$\frac{3n-3}{2}$	$\frac{3n-3}{2}$
n is even	$\frac{3n-2}{2}$	$\frac{3n-2}{2} - 1$

Table 2.

By Table 1, Clearly $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Slanting Ladder SL_n is a Strongly*-2 divisor cordial graph.

Example 2.5. A Strongly*-2 Divisor Cordial labelling of the graph SL_7 is shown in Figure 3

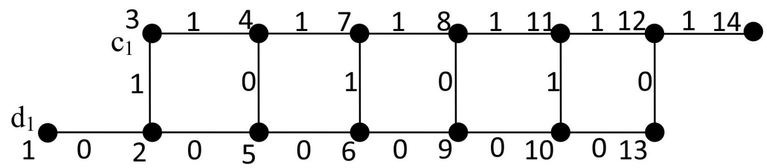


Figure 3

Theorem 2.6.The Double Triangular Snake graph $D(T_n)$ is a Strongly*-2 divisor cordial graph.

Proof. Let $V(D(Q_n)) = \{c_1, c_2, c_3, \dots . c_n\} \cup \{d_1, d_2, d_3, \dots . d_n\} \cup \{e_1, e_2, e_3, \dots . e_n\}$and

$E(D(T_n)) = \{c_s d_s\,; 1 \le s \le n-1\} \cup \{c_s c_{s+1}\,; 1 \le s \le n-1\} \cup \{c_s e_s;\ 1 \le s \le n-1\}$
$\cup \{d_s c_{s+1};\ 1 \le s \le n-1\} \cup \{e_s c_{s+1}; 1 \le s \le n-1\}$.

Then Double Triangular Snake graph has $3n-2$ vertices and $5(n-1)$ edges.

Case 1: n is odd

Define $\alpha : V(D(T_n)) \to \{1,2,3 \dots ,3n-2\}$ by

$\alpha(c_{2s}) = 4s-2, 1 \le s \le \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(c_{2s+1}) = 4s+1, 0 \le s \le \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(d_{2s-1}) = 4s-1, 1 \le s \le \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(d_{2s}) = 4s, 1 \le s \le \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(e_s) = 3n-1-s, 1 \le s \le n-1$

Case 2: n is even

Subcase 2.1: $n = 2$

Define $\alpha : V(D(T_n)) \to \{1,2,3 \dots ,3n-2\}$ by

$\alpha(c_1) = 1;\ \alpha(c_2) = 3$

$\alpha(d_1) = 2;\ \alpha(e_1) = 4$

Subcase 2.1: $n \ne 2$

Define $\alpha : V(D(T_n)) \to \{1,2,3 \dots ,3n-2\}$ by

$\alpha(c_1) = 1$

$\alpha(c_{2s}) = 4s, 1 \le s \le \frac{n}{2} - 1$

$\alpha(c_{2s+1}) = 4s+1, 1 \le s \le \frac{n}{2} - 1$

$\alpha(c_n) = 2n-1$

$\alpha(d_{2s-1}) = 4s-2, 1 \le s \le \frac{n}{2}$

$\alpha(d_{2s}) = 4s-1, 1 \le s \le \frac{n}{2} - 1$

$\alpha(e_s) = 3n-1-s, 1 \le s \le n-1$

The values of $e_\alpha(0)$ and $e_\alpha(1)$ are given in Table 3.

Nature of n	$e_\alpha(0)$	$e_\alpha(1)$
n is odd	$\frac{5n-5}{2}$	$\frac{5n-5}{2}$
n is even	$\frac{5n}{2}-3$	$\frac{5n}{2}-2$

Table 3.

From Table 3, Clearly $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Double Triangular Snake graph $D(T_n)$ is a Strongly*-2 divisor cordial graph.

Example 2.7. A Strongly*-2 Divisor Cordial labelling of the graph $D(T_6)$ is shown in Figure 4

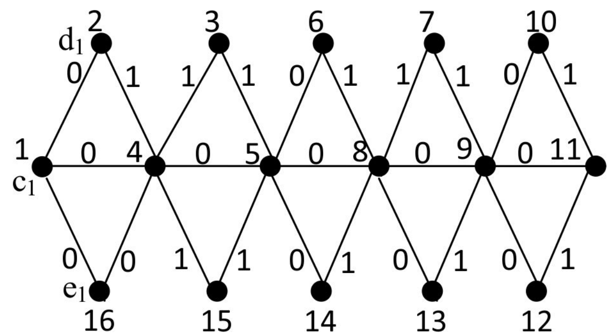


Figure 3

Theorem 2.8.The Triple Triangular Snake graph $T(T_n)$ is a Strongly*-2 divisor cordial graph.

Proof. Let $V(T(T_n)) = \{c_1, c_2, c_3, \ldots . c_n\} \cup \{d_1, d_2, d_3, \ldots . d_n\} \cup \{e_1, e_2, e_3, \ldots . e_n\} \cup \{f_1, f_2, f_3, \ldots . f_n\}$ and

$E(T(T_n)) = \{c_s d_s \,; 1 \leq s \leq n-1\} \cup \{c_s c_{s+1} \,; 1 \leq s \leq n-1\} \cup \{c_s e_s;\ 1 \leq s \leq n-1\}$
$\cup \{c_s f_s;\ 1 \leq s \leq n-1\} \cup \{d_s c_{s+1};\ 1 \leq s \leq n-1\} \cup \{e_s c_{s+1}; 1 \leq s \leq n-1\}$
$\cup \{f_s c_{s+1}; 1 \leq s \leq n-1\}$.

Then Triple Triangular Snake graph has $4n-3$ vertices and $7(n-1)$ edges.

Case 1: n is odd

Define $\alpha : V(T(T_n)) \to \{1,2,3 \ldots, 4n-3\}$ by

$\alpha(c_{2s}) = 4s-2, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(c_{2s+1}) = 4s+1, 0 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(e_{2s-1}) = 4s-1, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(e_{2s}) = 4s, 1 \leq s \leq \left\lfloor \frac{n}{2} \right\rfloor$

$\alpha(f_s) = 3n-1-s, 1 \leq s \leq n-1$

$\alpha(d_s) = 3n-2+s, 1 \leq s \leq n-1$

By this labelling , we get $e_\alpha(0) = e_\alpha(1) = \left\lfloor \frac{7n}{2} \right\rfloor - 3$

Case 2: n is even

Define $\alpha : V(T(T_n)) \to \{1,2,3 \ldots, 4n-3\}$ by

$\alpha(c_{2s}) = 4s-2, 1 \leq s \leq \frac{n}{2}$

$\alpha(c_{2s+1}) = 4s + 1, 0 \leq s \leq \frac{n}{2}$

$\alpha(e_{2s-1}) = 4s - 1, 1 \leq s \leq \frac{n}{2}$

$\alpha(e_{2s}) = 4s, 1 \leq s \leq \frac{n}{2} - 1$

$\alpha(f_s) = 3n - 1 - s, 1 \leq s \leq n - 1$

$\alpha(d_s) = 3n - 2 + s, 1 \leq s \leq n - 1$

By this labelling , we get $e_\alpha(1) = \frac{7n}{2} - 3$ and $e_\alpha(0) = \frac{7n}{2} - 4$

Clearly $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Triple Triangular Snake graph $T(T_n)$ is a Strongly*-2 divisor cordial graph.

Example 2.9. A Strongly*-2 Divisor Cordial labelling of the graph $T(T_6)$ is shown in Figure 5

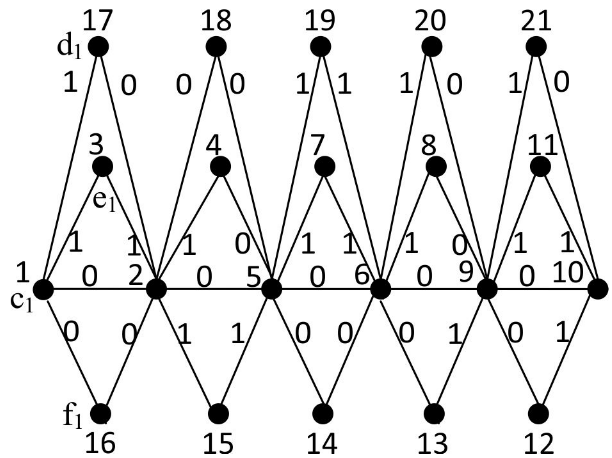


Figure 5

Theorem 2.10.The Four Triangular Snake graph $F(T_n)$ is a Strongly*-2 divisor cordial graph.

Proof. Let

$V(F(T_n)) = \{c_1, c_2, c_3, \dots . c_n\} \cup \{d_1, d_2, d_3, \dots . d_{n-1}\} \cup \{e_1, e_2, e_3, \dots . e_{n-1}\} \cup \{f_1, f_2, f_3, \dots . f_{n-1}\} \cup \{g_1, g_2, g_3, \dots . g_{n-1}\}$ and

$E(F(T_n)) = \{c_s d_s ; 1 \leq s \leq n - 1\} \cup \{c_s c_{s+1} ; 1 \leq s \leq n - 1\} \cup \{c_s e_s ; 1 \leq s \leq n - 1\}$
$\cup \{c_s f_s ; 1 \leq s \leq n - 1\} \cup \{d_s c_{s+1} ; 1 \leq s \leq n - 1\} \cup \{e_s c_{s+1} ; 1 \leq s \leq n - 1\}$
$\cup \{f_s c_{s+1} ; 1 \leq s \leq n - 1\} \cup \{c_s g_s ; 1 \leq s \leq n - 1\} \cup \{g_s c_{s+1} ; 1 \leq s \leq n - 1\}$.

Then the Four Triangular Snake graph has $5n - 4$ vertices and $9(n - 1)$ edges.

Define a function $\alpha : V(T(T_n)) \to \{1,2,3 \dots ,5n - 4\}$

and assign the labels for the vertices c_s for $1 \leq s \leq n$, d_s, e_s and f_s for $1 \leq s \leq n - 1$ as in the cases 1 and 2 in theorem 2.8.

Now assign the label to the vertices $g_s, 1 \leq s \leq n - 1$ as follows

$\alpha(g_s) = 5n - 3 - s, 1 \leq s \leq n - 1$ for all n.

The values of $e_\alpha(0)$ and $e_\alpha(1)$ are shown in Table 4.

Nature of n	$e_\alpha(0)$	$e_\alpha(1)$
n is odd	$\frac{9n-9}{2}$	$\frac{9n-9}{2}$
$n \equiv 0(mod\ 4)$	$\frac{9n}{2}-5$	$\frac{9n}{2}-4$
$n \equiv 2(mod\ 4)$	$\frac{9n}{2}-4$	$\frac{9n}{2}-5$

Table 4.

Clearly, $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Four Triangular Snake graph $F(T_n)$ is a Strongly*-2 divisor cordial graph.

Example 2.11. A Strongly*-2 Divisor Cordial labelling of the graph $F(T_6)$ is shown in Figure 6

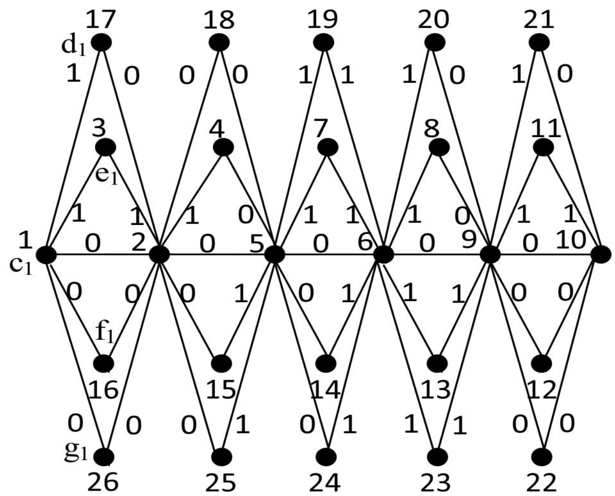


Figure 6

Theorem 2.12. The Double Quadrilateral Snake graph $D(Q_n)$ is a Strongly*-2 divisor cordial graph for n is even.

Proof. Let $V(D(T_n)) = \{c_1, c_2, \dots, c_n\} \cup \{d_1, d_2, \dots ., \} \cup \{e_1, e_2, \dots, e_{n-1}\} \cup \{\acute{d}_1, \acute{d}_2, \dots, \acute{d}_{n-1}\} \cup \{\acute{e}_1, \acute{e}_2, \dots, \acute{e}_{n-1}\}$ and

$E(D(T_n)) = \{c_s d_s\,; 1 \leq s \leq n-1\} \cup \{c_s c_{s+1}\,; 1 \leq s \leq n-1\} \cup \{d_s e_s;\ 1 \leq s \leq n-1\}$
$\cup \{e_s c_{s+1};\ 1 \leq s \leq n-1\} \cup \left\{c_s \acute{d}_s; 1 \leq s \leq n-1\right\} \cup \left\{\acute{d}_s \acute{e}_s; 1 \leq s \leq n-1\right\}$
$\cup \{c_{s+1} \acute{e}_s; 1 \leq s \leq n-1\}$.

Then Double Quadrilateral Snake graph has $5n-4$ vertices and $7(n-1)$ edges.

Define $\alpha : V(D(Q_n)) \to \{1,2,3, \dots, 5n-4\}$ by

$\alpha(d_1) = 1;\ \alpha(e_1) = 2$

$\alpha(d_s) = 5s-3, 2 \leq s \leq n-1$

$\alpha(e_s) = 5s-2, 2 \leq s \leq n-1$

$\alpha(c_1) = 6;\ \alpha(c_2) = 3$

$\alpha(c_{s+1}) = 5s-1, 2 \leq s \leq n-1$

$\alpha\left(\acute{d}_s\right) = 5s+1, 2 \leq s \leq n-1$

$\alpha\left(\acute{d}_1\right) = 5\,;\ \alpha(\acute{e}_1) = 4$

$\alpha(\acute{e}_s) = 5s, 2 \leq s \leq n-1$

The values of $e_\alpha(0)$ and $e_\alpha(1)$are tabulated in Table 5.

Nature of n	$e_\alpha(0)$	$e_\alpha(1)$
$n = 2$	4	3
$n \equiv 0(mod\ 4)$	$\frac{7n}{2} - 3$	$\frac{7n}{2} - 4$
$n \neq 2$ and $n \equiv 2(mod\ 4)$	$\frac{7n}{2} - 4$	$\frac{7n}{2} - 3$

Table 5.

Clearly, $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the Double Quadrilateral Snake graph $D(Q_n)$ is a Strongly*-2 divisor cordial graph.

Theorem 2.13. The graph $P_n \odot K_2$ is a Strongly*-2 divisor cordial graph for n is even.

Proof.$V(P_n \odot K_2) = \{c_1, c_2, c_3, \dots . c_n\} \cup \{d_1, d_2, d_3, \dots . d_{n-1}\} \cup \{e_1, e_2, e_3, \dots . e_{n-1}\}$and

$$E(P_n \odot K_2) = \{c_s d_s\, ; 1 \leq s \leq n-1\} \cup \{c_s c_{s+1}\, ; 1 \leq s \leq n-1\} \cup \{d_s e_s;\ 1 \leq s \leq n\} \cup \{c_s e_s;\ 1 \leq s \leq n\}$$

Then $P_n \odot K_2$ has $5n$vertices and$4n - 1$ edges.

Define $\alpha : V(P_n \odot K_2) \to \{1,2, \dots ,3n\}$ by

$\alpha(c_{2s}) = 4s, 1 \leq s \leq \frac{n}{2}$

$\alpha(c_{2s+1}) = 4s + 1, 0 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s-1}) = 4s - 2, 1 \leq s \leq \frac{n}{2}$

$\alpha(d_{2s}) = 4s - 1, 1 \leq s \leq \frac{n}{2}$

$\alpha(e_s) = 2n + s, 1 \leq s \leq n - 1$

$\alpha(e_n) = 3n$

From the above labeling, we get $e_\alpha(0) = 2n - 1$ and $e_\alpha(1) = 2n$

Clearly, $|e_\alpha(0) - e_\alpha(1)| \leq 1$.

Hence the graph $P_n \odot K_2$ is a Strongly*-2 divisor cordial graph for n is even.

REFERENCES

1. *C. Adiga and D. Somashekara, Strongly*- graphs,Math. Forum, Volume 13, 31-36,1999/2000.*
2. *J.BaskarBabujeeandC.Beaula,OnvertexStrongly*-graphs,ProceedingsofInternational Conference on Mathematics and Computer Science, 25-26 July (2008),Layola College, Chennai(2008).*
3. *I. Cahit, Cordial graphs: a weaker version of graceful and harmonious graph, Ars Combin.,23(1987), 201-207.*
4. *F.Harary,GraphTheory,NarosaPublishingHouse,NewDelhi,1988.*
5. *Joseph A. Gallian, A Dynamic survey of Graph Labeling, The Electronic JournalofCombinatorics,24(2021),#DS6.*
6. *A.LourdusamyandF.Patrick,Sumdivisorcordialgraphs,ProyeccionesJournalofMathematics, 2016;35(1):115-132.*
7. *R.Varatharajan,S.NavaneethaKrishnanandK.Nagarajan,Divisorcordialgraph,Internationalj ournalofMathematicsandcombinatorics,2011;4:15-25.*

DOI No. : 10.5958/2394-9309.2022.00031.2
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Online) : 2394-9309

APPLICATION OF MACHINE LEARNING ALGORITHMS IS USING FOR PREDICTION AND CLASSIFICATION OF COVID-19 CASES IN INDIA

M. Aniji*, S. Saraswathy, P. Murugaian***, Manimannan G******
*,**,***,**** Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai

ABSTRACT
This research paper is attempted to identify the hidden patterns of COVID-19 affected population using k Means++ and Machine Learning (ML) algorithms are of States and Union Territories with their parameters. Prediction, mapping, and classifying the data of COVID-19 database with help of seven parameters are used in the Machine Learning Process (ML), and they are total cases, active cases, discharged cased, death cases, active ratio percentage, discharge ratio percentage and death ratio percentage. The secondary sources of database were collected from Ministry of Health and Family Welfare Department (MHFWD), from Indian State and Union Territories during March 2021 to up to February, 2022 Initially, the k-Means++ machine learning clustering algorithm identified four cluster based on Silhouette Scores 0.632, 0.644, 0.657and 0.620 using preprocessing initialization method of k-Means++. Secondly, k-Means++ trained data used for identify in the following ML model. ML algorithms used to identify the States and UTs classification, prediction and their AUC, CA, F1, Precision and Recall of COVID-19 parameters. Finally, four clusters are visualized in GIS map. The SGD Model and other ML model accuracy values are 100% in SGD, 99.1 in kNN and 100% in Naïve Bayes. Almost all the states in UTs stable condition of COVID-19 parameters Except, Maharashtra, Kerala and Mizoram. The Maharashtra State is Very Highly Affected cases, Kerala State is Highly Affected States, Mizoram is very low Affected States and rest of the States and Union Territories (UTs) low affected and stable in India, COVID-19 parameters, it is based on 2011 population census
Keywords: COVID-19, k-Means++. SGD, kNN, Naïve Bayes, Predictions, Classification, GIS mapping and Scatter Plot.

1. INTRODUCTION

In third wave of COVID-19 spread in Indian states and UTs. The second largest population in the world and first and seconds waves, India lost many populations. Third wave virus is called delta declared by World Health Organization. In this connection, The COVID-19 pandemic disease caused by SARS-CoV-2 virus and this virus was identified from Wuhan, China in the year 2019. World Health Organization declared world pandemic situation on 11th March, 2020, its spread over the world in a short period. Many people were affected by this virus and lost their lives, economy, jobs, Education, etc. In India, the first case is recorded from Kerala and it has spread over the Indian states and union territories.

Recently, the second wave of COVID-19 spread is exponentially increasing in all states and union territories. The Indian government launched vaccine camps for the age group of above 45 to upper age people. The vaccine raises immunity in our human bodies and second dosage of vaccine is from 28 to 42 days. The main objective of this research paper is to predict probabilities, classification and visualization of COVID-19 affected states and UTs using SGD and other ML algorithm.

2. BACKGROUND OF THE STUDY

In Data Mining the KDD (Knowledge Discovery in a Database) is the iterative process that uses to discover novel information and knowledge from large amounts of database. According to Han and Kamber, data mining software allows end-users to analyze data from different dimensions Also, categorize the details and summarize the relationships which are identified during the mining process Kamber and Han, classification falls into a supervised data mining technique. This process consists of two steps, the first step is learning setup, in which the model is constructed and trained with the help of a predetermined database with class labels. The second step is the place where the trained model is consumed to perform the predictions for given database and measure the accuracy level of the classifier algorithm model [1].

Most of the researches are focusing on behavior of predictive models in data mining (PadmavathiJanardhana). With the help of modern technologies, the researcher can collect a large number of various types of data with different types of parameters in relevant field. Then apply the data mining techniques very correctly and effectively to mine and absorb meaningful interpretations, predictions, etc. when comes to the medical science field, the above said process can be applied to many database like, predicting, classification and visualization of breast cancer, heart attack, oral diseases, diabetes, etc. [2] \

Manimannan G. *et. al.* classification is defined as a process that gives the model to describe and differentiate database classes or concepts to predict the class of objects whose class label is not known. Partitioning clustering algorithm, artificial neural networks, etc. are the major tools used for constructing these models in COVID-19 of Indian States [3]. In recent days, data mining has become an attractive discipline that is used in business, medical science, engineering, text mining and other professional field as well in information technology community. Data mining strategies can provide useful answers to a problem. The methods are Classification, Association, Clustering, Estimation, Novelty detection, sequencing deduction, etc. [4]

Arumugam P. *et. al.* (2020), has used ML methods of Silhouette distance measure for k-Means++ clustering algorithm, ML Methods of like SVM, kNN, Random Forest and Logistic Regression. K-Means++ clustering analysis is used to identify five meaningful clusters and are labeled as Very Low, Low, Moderate, High and Very High of four major parameters based on their average values. Five clusters are cross validated using four ML algorithm and affected states are visualized with help of prediction and probabilities. The different ML models cross validation and classification accuracy are

88%, 97%, 91% and 91%. The Classification of States and Union Territories were named as Very Low Affected (VLA), Low Affected (LA), Moderately Affected (MA), Highly Affected (HA) and Very Highly Affected (VHA) States and UTs of India by COVID-19 cases. Maharashtra is correctly classified and Very High Affected States, Delhi, Uttar Pradesh and West Bengal falls in Moderately Affected States, Assam, Bihar, Chhattisgarh, Haryana, Gujarat, Madhya Pradesh, Odisha, Punjab, Rajasthan and Telangana falls in Low Affected States and Tamilnadu, Kerala Andhra Pradesh and Karnataka forms a group of Highly Affected States. Remaining States and Union Territories falls in Very Low affected by Covid-19 Cases [5].

Rajkumar and G.S Reena carried out research using machine learning algorithms (such as K- nearest neighbor, Naive Bayes) for heart disease prediction. The data set consists of 3000 instances with 14 attributes. Dataset was divided as 70% for training and 30% for testing. According to the test results, Naïve Bayes algorithm was selected as the algorithm with better performances when compared with KNN and Decision List [6].

Halgurd S. Maghdid*et al*. have projected a new framework to detect corona virus disease using the inboard smartphone sensors. The designed AI (Artificial Intelligence) framework collects data from various sensors to predict the grade of pneumonia as well as predicting the infection of the disease [7]. The proposed framework takes uploaded CT scan images as the key method to predict COVID-19 [8]

Manimannan G. *et. al.* has used Silhouette distance measure for k- means clustering algorithm. It produced effective results and visualized their result in a simple manner. This technique achieved three meaningful groups and is labeled as C1, C2 and C3. C1 represents highly affected, C2 Moderate affected and C1 Low affected States and Union Territories [9].

Machine learning (ML) is a branch of knowledge that is involved with the design and implementation of algorithms that allow computers to adjust their behavior according to data [10]. ML automatically learns to identify composite patterns and makes intelligent judgment based on data. The development of the data mining applications such as classification and clustering led to the need for the ML algorithms to be applied for huge scale data. The aim of ML is to resolve problems in intelligent ways by enhancing the performance of computer programs [11].

3. DATABASE AND METHODOLOGY

The secondary source of database was collected from Ministry of Health and Family Welfare Department (MHFWD), from Indian State and Union Territories during March 2021 to up to February, 2022 (Table 1). The database consists of seven parameters namely total cases, active cases, discharged cased, death cases, active ratio percentage, discharge ratio percentage and death ratio percentage. Initially all parameters are clustered using k-means algorithm and additionally tested and cross validated 10 folds stratified sampling methods with help of testing database 30 and training database 70 percent.

Subsequently, the machine learning algorithm of kNN, SVM, Random Forest classification and neural network cross validates the k-mean++ clustered database and gives better results.

Table 1. Sample Data and Parameters

	State/UTs	Total Cases	Active	Discharged	Deaths	Active Ratio (%)	Discharge Ratio (%)	Death Ratio (%)
1	Andaman and Nicobar	7539	6	7404	129	0.08	98.21	1.71
2	Andhra Pradesh	1970008	20582	1936016	13410	1.04	98.27	0.68
3	Arunachal Pradesh	48565	3508	44823	234	7.22	92.29	0.48
4	Assam	568257	12429	550534	5294	2.19	96.88	0.93
5	Bihar	724917	401	714872	9644	0.06	98.61	1.33
6	Chandigarh	61960	33	61116	811	0.05	98.64	1.31
7	Chhattisgarh	1002458	1918	987012	13528	0.19	98.46	1.35
8	Dadra and Nagar Haveli and Daman and Diu	10650	15	10631	4	0.14	99.82	0.04
9	Delhi	1436401	538	1410809	25054	0.04	98.22	1.74
10	Goa	171295	1027	167118	3150	0.60	97.56	1.84
11	Gujarat	824922	251	814595	10076	0.03	98.75	1.22
12	Haryana	769956	703	759614	9639	0.09	98.66	1.25
13	Himachal Pradesh	206369	1304	201543	3522	0.63	97.66	1.71
14	Jammu and Kashmir	321725	1254	316090	4381	0.39	98.25	1.36
15	Jharkhand	347223	239	341855	5129	0.07	98.45	1.48
16	Karnataka	2908284	24045	2847627	36612	0.83	97.91	1.26
17	Kerala	3425473	165834	3242684	16955	4.84	94.66	0.49
18	Ladakh	20345	57	20081	207	0.28	98.70	1.02
19	Lakshadweep	10207	79	10078	50	0.77	98.74	0.49
20	Madhya Pradesh	791862	132	781217	10513	0.02	98.66	1.33
21	Maharashtra	6315063	78700	6103325	133038	1.25	96.65	2.11
22	Manipur	99872	9814	88480	1578	9.83	88.59	1.58
23	Meghalaya	65939	5843	58987	1109	8.86	89.46	1.68
24	Mizoram	40111	12316	27642	153	30.70	68.91	0.38
25	Nagaland	28004	1300	26130	574	4.64	93.31	2.05
26	Odisha	979737	13318	960386	6033	1.36	98.02	0.62
27	Puducherry	121059	944	118320	1795	0.78	97.74	1.48
28	Punjab	599162	473	582395	16294	0.08	97.20	2.72
29	Rajasthan	953704	241	944509	8954	0.03	99.04	0.94
30	Sikkim	26880	3323	23211	346	12.36	86.35	1.29
31	Tamil Nadu	2563544	20385	2509029	34130	0.80	97.87	1.33
32	Telengana	645997	8819	633371	3807	1.37	98.05	0.59
33	Tripura	79026	3104	75167	755	3.93	95.12	0.96
34	Uttar Pradesh	1708500	646	1685091	22763	0.04	98.63	1.33
35	Uttarakhand	342198	574	334261	7363	0.17	97.68	2.15
36	West Bengal	1529295	10803	1500331	18161	0.71	98.11	1.19

The following section describes various machine learning algorithms and Workflow (Figure 1 and 2):

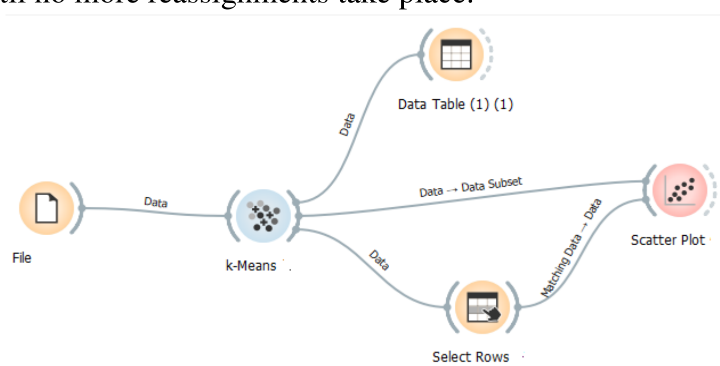
3.1 kMean++ Clustering Algorithm

MacQueen [12] suggests the term k-means for describing an algorithm of his that assigns each item to the cluster having the nearest centroids. The process composed of these three steps:

Step 1: Partition the items (input database) into k-initial clusters.

Step 2: Proceed through the list of items, assigning an item to the cluster whose centroid is nearest (using Euclidean distance measure). Recalculate the centroid for the cluster receiving the new item and for the cluster losing the item.

Step 3: Repeat Step 2 until no more reassignments take place.



3.2 Stochastic Gradient Decent (SGD) Model

Stochastic Gradient Descent often abbreviated SGD. It is an iterative method for optimizing an objective function with suitable smoothness properties. It can be regarded as a stochastic approximation of gradient descent optimization, since it replaces the actual gradient calculated from the entire data set by an estimate calculated from a randomly selected subset of the data. Exclusively in high-dimensional optimization problems this reduces the computational burden, achieving faster iterations in trade for a lower convergence rate.[1]

The basic idea behind stochastic approximation can be traced back to the Robbins–Monro algorithm of the 1950s, SGD has become an important optimization method in machine learning [13]

The Stochastic Gradient Descent widget uses stochastic gradient descent that minimizes a chosen loss function with a linear function. The algorithm approximates a true gradient by considering one sample at a time, and simultaneously updates the model based on the gradient of the loss function. For regression, it returns predictors as minimizers of the sum, M-estimators, and is especially useful for large-scale and sparse dataset. It has been around in the Machine Learning (ML) society for a long time, and has been given a great amount of notice just recently in the context of large-scale learning. SGD has been famously applied to large-scale and sparse machine learning obstacle frequently experienced in text classification and natural language processing [14].

Step 1: Specify the name of the model (SGD widget) in name box.

Step 2: in this model classification loss function Hinge and Regression squared loss function are used and it is like linear Support Vector Machine (SVM) Model.

Step 3: The standard Ridge (L2) regularization method with help of regularization strength α=0,00001 is used.

Step 4: The learning rate is constant with initial learning rate $\eta_0 = 0{,}0100$

Step 5: Shuffle after each iteration=5.

Step 6: This widget connect to prediction widget and it will be produce results.

3.3 k-Nearest Nabors (k-NN))

This algorithm of k-Nearest Nabors (k-NN) in statistics is a non-parametric classification method first developed by Evelyn Fix *et. al.* 1951 [15] and later expanded by Thomas Cover [16]. It is used for classification and regression. In both the cases, input consists of the *k* closest training examples in data set. The output depends on whether *k*-NN is used for classification or regression: The Orange kNN algorithm is:

Step 1: A name under which it will appear in other widgets. The default name is “kNN”.

Step 2: Set the number of nearest neighbors, the distance parameter and weights as model criteria. In this paper, the researcher used Euclidean distance between two points.

Step 3: The *Weights* you can use are:

Uniform: all points in each neighborhood are weighted equally.

Distance: closer neighbors of a query point have a greater influence than the neighbors further away.

Step 4: Produce a report in the Test and Score window.

3.4 Naïve Bayes Model

The Naïve Baye’s algorithm can be trained very efficiently in large database, classification, sentiment analysis, text analysis, etc. The Naïve Baye’s algorithm offers fast model building and scoring both for binary and multiclass situation for high volume of data. The algorithm makes predictions using Baye's Theorem, which incorporate evidence or prior knowledge in its prediction (Hand. e*t. al.,* 2001) [17] Bayes theorem relates the conditional and marginal probabilities of stochastic events *C* and *A*, which when statistically stated are:

$$P\left(\frac{C}{A}\right) = \frac{P\left(\frac{A}{C}\right)P(C)}{P(A)} = L\left(\frac{A}{C}\right)P(C)$$

Where *P* stands for the probability of the variable within parenthesis, and *L(A/C)* is referred to as likelihood of *A* given fixed *C*, the conventional names of each terms are:

P(C) is the *prior probability or marginal probability* of *C*. It is prior in the sense that it has not yet accounted for the information available in *A*.

P(C/A) is the *conditional probability* of *C,* It is also called the posterior probability because it has already incorporated the out come of event *A*.

P(A/C) is the conditional probability of *A* give*n* *C*.

P(A) is the prior or marginal probability of *A*, which is normally the evidence. With this terminology, the theorem may be put in word as:

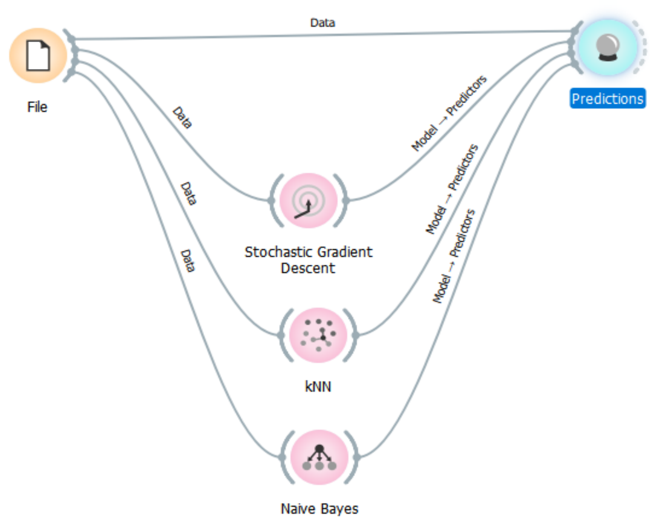
$$Posterior = \frac{Likelihood * Prior}{Normalizing\ Constant}$$

The ratio *P(A/C)/P(A)* is sometimes called the standardized likelihood.

Step 1: The Naïve Baye's Widget applies automatically and Produce the result in Prediction Widget.

Step 2: Naïve Baye's widget remove the empty columns in the preprocessing by default.

Figure 2. Workflow of various SGD and other Machine Learning Algorithms



Reddy Prasad suggests a system at his paper to classify the patient with heart disease based on some features. Therefore, the proposed system can predict the presence of heart problems on a person based on the given data. He used logistic regression technology to perform classification and prediction. Further used sigmoid function for the representation processed [18].

4. RESULT AND DISCUSSION

The k-Means++ algorithms achieved five meaningful clusters and are labeled as four categories of States and Union Territories of India affected by COVID-19 (Table 1 and Figure 3). The k-means++ clustering algorithm identified four cluster based on Silhouette distance Scores 0.632, 0.644, 0.657and 0.620 with preprocessing initialization method. k-means++ trained data used is for SGD and other ML method prediction and classification of COVID-19 cases of Indian states and UTs.

Table 1: Final Cluster of k-means Algoritm

	State/UTs	Cluster	Silhouette	Total Cases	Active	Discharged	Deaths	Active Ratio (%)	ischarge Ratio (%)	Death Ratio (%)
1	Andaman and ...	C1	0.705419	7539	6	7404	129	0.08	98.21	1.71
2	Andhra Pradesh	C1	0.657742	1970008	20582	1936016	13410	1.04	98.27	0.68
3	Arunachal Prad...	C1	0.665502	48565	3508	44823	234	7.22	92.29	0.48
4	Assam	C1	0.703006	568257	12429	550534	5294	2.19	96.88	0.93
5	Bihar	C1	0.711409	724917	401	714872	9644	0.06	98.61	1.33
6	Chandigarh	C1	0.708937	61960	33	61116	811	0.05	98.64	1.31
7	Chhattisgarh	C1	0.707602	1002458	1918	987012	13528	0.19	98.46	1.35
8	Dadra and Nag...	C1	0.671351	10650	15	10631	4	0.14	99.82	0.04
9	Delhi	C1	0.69095	1436401	538	1410809	25054	0.04	98.22	1.74
10	Goa	C1	0.704689	171295	1027	167118	3150	0.60	97.56	1.84
11	Gujarat	C1	0.710119	824922	251	814595	10076	0.03	98.75	1.22
12	Haryana	C1	0.710918	769956	703	759614	9639	0.09	98.66	1.25
13	Himachal Prad...	C1	0.707718	206369	1304	201543	3522	0.63	97.66	1.71
14	Jammu and Kas...	C1	0.711824	321725	1254	316090	4381	0.39	98.25	1.36
15	Jharkhand	C1	0.711231	347223	239	341855	5129	0.07	98.45	1.48
16	Karnataka	C1	0.613987	2908284	24045	2847627	36612	0.83	97.91	1.26
17	Kerala	C4	0.5	3425473	165834	3242684	16955	4.84	94.66	0.49
18	Ladakh	C1	0.705477	20345	57	20081	207	0.28	98.70	1.02
19	Lakshadweep	C1	0.691184	10207	79	10078	50	0.77	98.74	0.49
20	Madhya Pradesh	C1	0.71082	791862	132	781217	10513	0.02	98.66	1.33
21	Maharashtra	C2	0.5	6315063	78700	6103325	133038	1.25	96.65	2.11
22	Manipur	C1	0.652154	99872	9814	88480	1578	9.83	88.59	1.58
23	Meghalaya	C1	0.663365	65939	5843	58987	1109	8.86	89.46	1.68
24	Mizoram	C3	0.5	40111	12316	27642	153	30.70	68.91	0.38
25	Nagaland	C1	0.689412	28004	1300	26130	574	4.64	93.31	2.05
26	Odisha	C1	0.689422	979737	13318	960386	6033	1.36	98.02	0.62
27	Puducherry	C1	0.709983	121059	944	118320	1795	0.78	97.74	1.48
28	Punjab	C1	0.674187	599162	473	582395	16294	0.08	97.20	2.72
29	Rajasthan	C1	0.703692	953704	241	944509	8954	0.03	99.04	0.94
30	Sikkim	C1	0.605822	26880	3323	23211	346	12.36	86.35	1.29
31	Tamil Nadu	C1	0.638464	2563544	20385	2509029	34130	0.80	97.87	1.33
32	Telengana	C1	0.693837	645997	8819	633371	3807	1.37	98.05	0.59
33	Tripura	C1	0.701409	79026	3104	75167	755	3.93	95.12	0.96
34	Uttar Pradesh	C1	0.689124	1708500	646	1685091	22763	0.04	98.63	1.33
35	Uttarakhand	C1	0.696119	342198	574	334261	7363	0.17	97.68	2.15
36	West Bengal	C1	0.69189	1529295	10803	1500331	18161	0.71	98.11	1.19

Figure 3. Scatter diagram of kMean++ Clustering of COVID-19 cases in Indian States and ITs

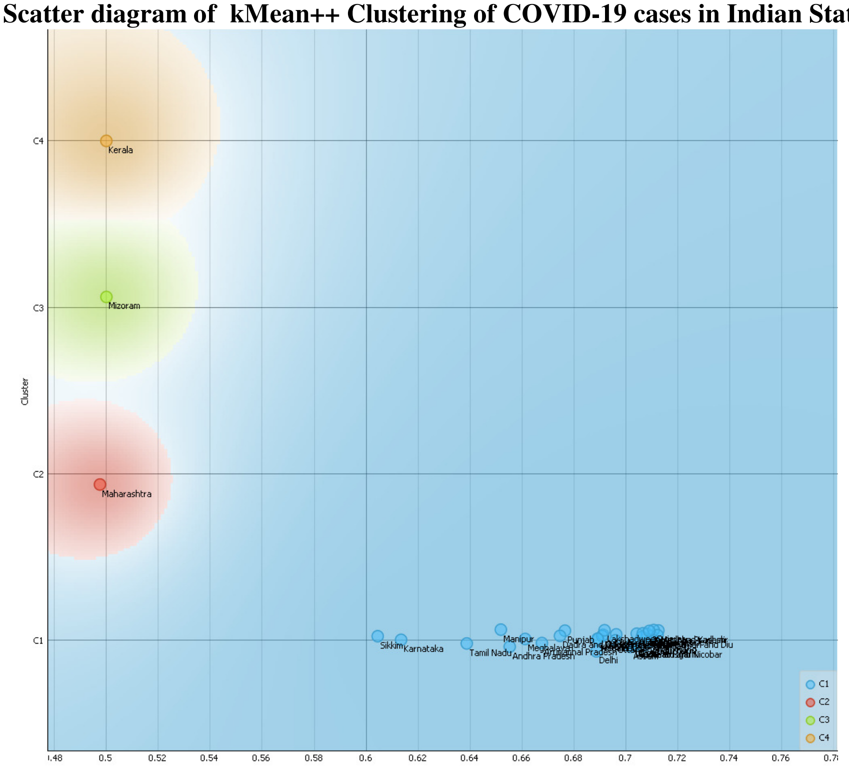


Table 2 result is shows that SGD and other Machine Learning (ML) model accuracy of training data. This is a very high value of accuracy and it is reasonable to expect that the model would be useful to predict new, previously unknown, instance of COVID-19 problem. The prediction results is achieved closer to hundred percent.

The Test scores of various algorithms can be used to define a number of performance criteria commonly used in model evaluation suggested by Ali, 2005 [19]. The following section describes various measures of the test score:

4.1 AUC (Area under Curve)

AUC stands for "Area under the ROC Curve." That is, AUC measures the entire two-dimensional area underneath the entire ROC curve from (0,0) to (1,1).

An **ROC** curve (Receiver Operating characteristic Curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameter, they are True Positive Rate and False Positive Rate.

True Positive Rate (**TPR**) is a synonym for recall and is therefore defined as follows:

$$TPR = \frac{TP}{TP + FN}$$

False Positive Rate (**FPR**) is defined as follows:

$$FPR = \frac{FP}{FP + TN}$$

4.2 Classification Accuracy (CA)

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions and formally has the following definition:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

Accuracy=Number of correct predictions shared by the Total number of predictions. For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

4.3 F1 Score

The F1 score is calculated by using the following formula:

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} = \frac{TP}{TP + \frac{1}{2}\ (FP + FN)}$$

4.4 Precision

Precision is given by:

$$\frac{True\ Positive}{True\ Positive\ \ + False\ Positive}$$

4.5 Recall

Recall Measure is computed by using the formula:

$$\frac{True\ Positive}{True\ Positive\ \ + False\ Negative}$$

The Models of AUC, CA, F1, Precision and Recall measure accuracy value closer to 1 is the best fitted model and closer to 0 is not a good fitted model . In this study all the measures are closter to 1 and these models are best fitted models using SGD and ML algorithms. The prediction and probabilitiy results of SGD and other ML algorithms of Indian states and UTs are in Table 3,4 and Figure 2. Maharastra is the Very High COVID-19 cases is predicted and visualized in the spatial map (Figure 4). Moderate state is Kerala, Very Lowest cases of COVID-19 occurred in Mizoram state and Rest of the States and UTs in lowest cases of COVID-19 cases using SGD and other ML methods

Table 2: Test Score and Classification accuracy of Various ML algorithms

Scores					
Model	AUC	CA	F1	Precision	Recall
SGD	1.000	1.000	1.000	1.000	1.000
kNN	0.997	0.995	0.978	0.958	0.997
Naïve Bayes	1.000	1.000	1.000	1.000	1.000

Figure 2. Classification Accuracy and Test Score of Covid-19 Cases using ML algorithms

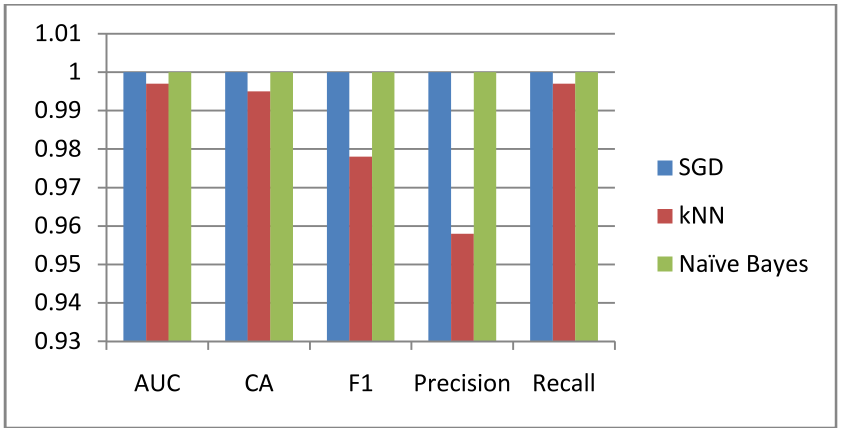


Table 3. Prediction and Probilities of States and Union Teritories

Data & Predictions

	SGD	kNN	Naive Bayes	Cluster	State/UTs
1	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.18 : 0.72 : 0.09 → C3	C1	Andaman and Nicobar
2	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.00 : 0.44 : 0.11 : 0.44 → C2	C1	Andhra Pradesh
3	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.20 : 0.00 → C1	0.00 : 0.03 : 0.86 : 0.11 → C3	C1	Arunachal Pradesh
4	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.11 : 0.44 : 0.44 → C3	C1	Assam
5	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.07 : 0.31 : 0.31 : 0.31 → C2	C1	Bihar
6	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.20 : 0.00 → C1	0.05 : 0.24 : 0.48 : 0.24 → C3	C1	Chandigarh
7	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.47 : 0.06 : 0.47 → C2	C1	Chhattisgarh
8	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.05 : 0.84 : 0.10 → C3	C1	Dadra and Nagar Haveli and Daman and Diu
9	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.00 : 0.64 : 0.04 : 0.32 → C2	C1	Delhi
10	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.05 : 0.47 : 0.24 : 0.24 → C2	C1	Goa
11	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.07 : 0.31 : 0.31 : 0.31 → C2	C1	Gujarat
12	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.08 : 0.31 : 0.31 : 0.31 → C2	C1	Haryana
13	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.03 : 0.65 : 0.16 : 0.16 → C2	C1	Himachal Pradesh
14	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.08 : 0.31 : 0.31 : 0.31 → C2	C1	Jammu and Kashmir
15	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.07 : 0.31 : 0.31 : 0.31 → C2	C1	Jharkhand
16	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.00 : 0.20 → C1	0.00 : 0.64 : 0.04 : 0.32 → C2	C1	Karnataka
18	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.02 : 0.10 : 0.79 : 0.10 → C3	C1	Ladakh
19	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.10 : 0.79 : 0.10 → C3	C1	Lakshadweep
20	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.07 : 0.31 : 0.31 : 0.31 → C2	C1	Madhya Pradesh
22	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.20 : 0.40 : 0.40 → C3	C1	Manipur
23	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.20 : 0.00 → C1	0.00 : 0.25 : 0.50 : 0.25 → C3	C1	Meghalaya
25	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.20 : 0.00 → C1	0.00 : 0.10 : 0.80 : 0.10 → C3	C1	Nagaland
26	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.33 : 0.33 : 0.33 → C2	C1	Odisha
27	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.05 : 0.47 : 0.24 : 0.24 → C2	C1	Puducherry
28	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.02 : 0.56 : 0.14 : 0.28 → C2	C1	Punjab
29	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.07 : 0.31 : 0.31 : 0.31 → C2	C1	Rajasthan
30	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.20 : 0.00 → C1	0.00 : 0.05 : 0.84 : 0.11 → C3	C1	Sikkim
31	1.00 : 0.00 : 0.00 : 0.00 → C1	0.80 : 0.00 : 0.00 : 0.20 → C1	0.00 : 0.64 : 0.04 : 0.32 → C2	C1	Tamil Nadu
32	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.02 : 0.11 : 0.44 : 0.44 → C3	C1	Telengana
33	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.14 : 0.57 : 0.28 → C3	C1	Tripura
34	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.01 : 0.47 : 0.06 : 0.47 → C2	C1	Uttar Pradesh
35	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.05 : 0.47 : 0.24 : 0.24 → C2	C1	Uttarakhand
36	1.00 : 0.00 : 0.00 : 0.00 → C1	1.00 : 0.00 : 0.00 : 0.00 → C1	0.00 : 0.64 : 0.04 : 0.32 → C2	C1	West Bengal
21	0.00 : 1.00 : 0.00 : 0.00 → C2	0.60 : 0.20 : 0.00 : 0.20 → C1	0.00 : 0.78 : 0.02 : 0.20 → C2	C2	Maharashtra
24	0.00 : 0.00 : 1.00 : 0.00 → C3	0.80 : 0.00 : 0.20 : 0.00 → C1	0.00 : 0.03 : 0.86 : 0.11 → C3	C3	Mizoram
17	0.00 : 0.00 : 0.00 : 1.00 → C4	0.80 : 0.00 : 0.00 : 0.20 → C1	0.00 : 0.18 : 0.09 : 0.73 → C4	C4	Kerala

Table 4. Prediction and Probilities of States and Union Teritories

Cluster	State/UTs	Silhouette	Total Cases	Active	Discharged	Deaths	Active Ratio Percentage	Discharge Ratio Percentage	Death Ratio Percentage
C1	Andaman and Nicobar	0.705419	7539	6	7404	129	0.08	98.21	1.71
C1	Andhra Pradesh	0.657742	1970008	20582	1936016	13410	1.04	98.27	0.68
C1	Arunachal Pradesh	0.665502	48565	3508	44823	234	7.22	92.29	0.48
C1	Assam	0.703006	568257	12429	550534	5294	2.19	96.88	0.93
C1	Bihar	0.711409	724917	401	714872	9644	0.06	98.61	1.33
C1	Chandigarh	0.708937	61960	33	61116	811	0.05	98.64	1.31
C1	Chhattisgarh	0.707602	1002458	1918	987012	13528	0.19	98.46	1.35
C1	Dadra and Nagar Haveli and Daman and Diu	0.671351	10650	15	10631	4	0.14	99.82	0.04
C1	Delhi	0.69095	1436401	538	1410809	25054	0.04	98.22	1.74
C1	Goa	0.704689	171295	1027	167118	3150	0.60	97.56	1.84
C1	Gujarat	0.710119	824922	251	814595	10076	0.03	98.75	1.22
C1	Haryana	0.710918	769956	703	759614	9639	0.09	98.66	1.25
C1	Himachal Pradesh	0.707718	206369	1304	201543	3522	0.63	97.66	1.71
C1	Jammu and Kashmir	0.711824	321725	1254	316090	4381	0.39	98.25	1.36
C1	Jharkhand	0.711231	347223	239	341855	5129	0.07	98.45	1.48
C1	Karnataka	0.613987	2908284	24045	2847627	36612	0.83	97.91	1.26
C1	Ladakh	0.705477	20345	57	20081	207	0.28	98.70	1.02
C1	Lakshadweep	0.691184	10207	79	10078	50	0.77	98.74	0.49
C1	Madhya Pradesh	0.71082	791862	132	781217	10513	0.02	98.66	1.33
C1	Manipur	0.652154	99872	9814	88480	1578	9.83	88.59	1.58
C1	Meghalaya	0.663365	65939	5843	58987	1109	8.86	89.46	1.68
C1	Nagaland	0.689412	28004	1300	26130	574	4.64	93.31	2.05
C1	Odisha	0.689422	979737	13318	960386	6033	1.36	98.02	0.62
C1	Puducherry	0.709983	121059	944	118320	1795	0.78	97.74	1.48
C1	Punjab	0.674187	599162	473	582395	16294	0.08	97.20	2.72
C1	Rajasthan	0.703692	953704	241	944509	8954	0.03	99.04	0.94
C1	Sikkim	0.605822	26880	3323	23211	346	12.36	86.35	1.29
C1	Tamil Nadu	0.638464	2563544	20385	2509029	34130	0.80	97.87	1.33
C1	Telengana	0.693837	645997	8819	633371	3807	1.37	98.05	0.59
C1	Tripura	0.701409	79026	3104	75167	755	3.93	95.12	0.96
C1	Uttar Pradesh	0.689124	1708500	646	1685091	22763	0.04	98.63	1.33
C1	Uttarakhand	0.696119	342198	574	334261	7363	0.17	97.68	2.15
C1	West Bengal	0.69189	1529295	10803	1500331	18161	0.71	98.11	1.19
C2	Maharashtra	0.5	6315063	78700	6103325	133038	1.25	96.65	2.11
C3	Mizoram	0.5	40111	12316	27642	153	30.70	68.91	0.38
C4	Kerala	0.5	3425473	165834	3242684	16955	4.84	94.66	0.49

Figure 4. Visualization States and UTs based on Machine learning Prediction

The SGD and other ML models algorithm of classification accuracy I00%, 99% and 100% respectively Prediction and classification of States and UTs of COVID-19 cases in India, were named as Very Highly Affected Sates (VAS), Highly Affected States (HAS), Very Low Affected States (VLAS) and Fourth Low Affected States (FHAS). Maharashtra is Very Highly Affected State, Kerala is Higly Affected States, Very LowThird Affected State is Mizoram and Rest of the States and UTs are falls in Low Affected States and UTs by Covid-19 parameters (Table 2 to 3). The second wave of COVID-19 also started from March 2021 to February 2022. The government is taking necessary action to prevent and control the spread of COVID-19. Also our government provides COVAXIN and COVISHIELD vaccination to 18-45, 45-60 and above 60 age groups.

5. CONCLUSION

This research paper is attempted to identified the hidden patterns using k=Means++ clustering algorithm in the initial ML process. It is identified four cluster and using this database for SGD and other ML algorithms for prediction, mapping and classification of COVID-19 database in Indian states and UTs with help of seven parameters, they are total cases, active cases, discharged cased, death cases, active ratio percentage, discharge ratio percentage and death ratio percentage. The secondary sources of database were collected from Ministry of Health and Family Welfare Department (MHFWD), from Indian State and Union Territories during March 2021 to February 2022/. k-Means++ clustering algorithm identified four cluster based on Silhouette Scores 0.632, 0.644, 0.657and 0.620 with preprocessing initialization method of k-Means++ to cross validated using Stochastic Gradient Decent (SGD) Model and other ML Algorithms used to identify the States and UTs classification, AUC, CA, F1, Precision and Recall of COVID-19 database. Four clusters are visualized in GIS map. The SGD Model and other ML algorithm accuracy values are 100% in SGD, 99.1 in kNN, 100% in Naïve Bayes Algorithm. Almost all the states in UTs stable condition of COVID-19 cases Except, Maharashtra is very highly affected state, Kerala is highly affected stateand Very low affected state is Mizoram and rest of the states and UTs lowest and stable conditions occurred in India, it is based on 2011 population census.

REFERENCES

1. *Han J. Pei and M. Kamber (2011), Data Mining: Concepts and Techniques. Elsevier, 2011.*
2. *PadmavathiJanardhanaet.al, (2015), Effectiveness of Support Vector Machines in Medical Data mining Journal Of Communications Software And Systems, Vol. 11, No. 1, pp. 25-30.*
3. *Manimannan G. et. al (2020), Clustering Study of Indian States and Union Territories by coronavirus (COVID-19) using k-mean algorithm, International Journal of Data Mining and Emerging Technologies, Vol.9 Issue-2, pp.43-51.*
4. *A.B. M. Shawkat Ali and Saleh A. Wasimi (2016). Data Mining: Methods and Techniques, Cengage Learning India private Imitated, New Delhi.*
6. *P. Arumugamet.al.(2021), International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-10 Issue-1, pp.16-20*
7. *AshaRajkumar, and Mrs G. Sophia Reena, 2010, "Diagnosis of Heart Disease using Data Mining Algorithms",Global Journal of Computer Science and Technology,Vol. 10,Issue 10, pp.38-43, September*

8. *Halgurd S. Maghdid, KayhanZrarGhafoor, Ali SafaaSadiq, Kevin Curran, and KhaledRabie (2020). A novel AI-enabled framework to diagnose coronavirus covid 19 using smartphone embedded sensors: Design study. Computer Science, Cornel University.pp.1-7.*
9. *Manimannan G, et.al. (2020),Prediction of Affected and Deceased Population Trend of COVID-19 in India using Statistical Analysis, International Journal of Scientific and Innovative Mathematical Research(IJSIMR) Volume 8, Issue 2, pp. 37-43.*
10. *T.Subramaniam, H.A.Jalab, and A.Y.Taqa (2010), "Overview of textual antispam filtering techniques," International Journal of Physical Sciences, vol. 5, no. 12, pp. 1869-1882, 2010.*
11. *N.O.F.Elssied and O.Ibrahim, "K-means clustering scheme for enhanced spam detection," Research Journal of Applied Sciences, Engineering and Technology, vol. 7, no. 10, pp. 1940-1952, 2014.*
12. *MacQueen (2015), Applied Multivariate Statistical Analysis, Prentice Hall, New Delhi.*
13. *Bottou, Léon; Bousquet, Olivier (2012). "The Tradeoffs of Large Scale Learning". In Sra, Suvrit; Nowozin, Sebastian; Wright, Stephen J. (eds.). Optimization for Machine Learning. Cambridge: MIT Press. pp. 351–368. ISBN 978-0-262-01646-9.*
14. *Bottou, Léon (1998). "Online Algorithms and Stochastic Approximations". Online Learning and Neural Networks.Cambridge University Press. ISBN 978-0-521-65263-6.*
15. *Evelyn Fix and Joseph L Hodges Jr (1951). Discriminatory analysis-nonparametric discrimination: consistency properties. Technical report, DTIC Document.*
16. *Thomas M Cover and Peter E Hart (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1):21-27.*
17. *David J. Hand and Keming Yu (2001), Idiots Baye's- Not so stupid after All?, International Statistical Review, Vo.69, Issue-3, pp. 385-398.*
18. *Reddy Prasad,Pidaparthi Anjali, S.Adil, N.Deepa (2019),"Heart Disease Prediction using Logistic Regression Algorithm using Machine Learning, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, p.659-662.*
19. *16. S. Ali. And S. Wasimi (2005). Data Mining: Methods and Techniques, Cengage Learning India.*

DOI No. : 10.5958/2394-9309.2022.00033.6
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A FUZZY RULE-BASED INFERENCE SYSTEM FOR DECISION MAKING IN REAL TIME APPLICATIONS

M. Vijaya*, A. Aruna**
*Head and Prof., PG and Research Dept of Mathematics, Marudupandiyar college, Thanjavur
**Ph.D. Scholar, Dept of Mathematics, Marudupandiyar college, Thanjavur
(Affiliated to Bharathidasan University, Trichy)
E-mail: mathvijaya23@gmail.com*, arunjo.jo21@gmail.com**

ABSTRACT

Fuzzy rule-based systems are one of the most important areas of applications of fuzzy sets and fuzzy logic. Fuzzy linguistic descriptions are formal representations of systems made through IF-THEN rules. In this paper, we are going to study fuzzy inference systems(FIS) and solve real time application problems using Fuzzy rule-based inference systems.

Keywords : Fuzzy rule-based, fuzzy inference system, fuzzy set

1. INTRODUCTION

Decision-making is an essential activity for an individual or in any organization. It can be considered as a procedure of choosing the best alternatives by considering every benefit and cost factor. As the problem of making decisions becomes complex, a more detailed process is necessary to find its solution. Fuzzy Rule Based systems are widely used in various real applications such as in anomaly intrusion detection, image processing, medical applications and in many other applications. In such applications, a large number of patterns and variables are present in the useful data or in the available data. This condition leads to a rapid increase of the fuzzy rule search space which in turn affects the inductive learning of FRBC systems. With this rapid growth, the learning process becomes further complicated, which results in scalability and complexity problems. In this paper we are going to solve one real time application using a fuzzy rule based inference system.

2. FUZZY OVERVIEW

2.1.Fuzzy Modeling and Fuzzy Systems

The main concepts in fuzzy logic are linguistic variables, linguistic modifiers, propositional fuzzy logic, deductive inference rules and approximations. Zadeh proposed the concept of linguistic variables for modeling human way of thinking, based on the following principle.With the increasing complexity of a system, our ability to make accurate and at the same time rational conclusions about its behavior is reduced till its reaches a threshold after which the accuracy and consistency are almost mutually exclusive requirements[2].The process of creating a fuzzy system, called fuzzy modeling .The structure of the fuzzy system has to be designed in such a way , so that the experience of the experts can be easily implemented in it. In case the input and output data are known , the possibility of identifying the system using standard techniques has to exist[3].

There are two milestones in fuzzy modeling:

i. Identification of the overall structure. Here information from experts or information obtained after "trials and errors" is used. The steps are

(a) Selecting appropriate input and output variables

(b) Choosing the type of fuzzy inference system.

(c) Determining the number of linguistic terms of input and output fuzzy variables.

(d) Designing the set of if-then rules.

ii. Identification of the structure in depth. Here a detailed description of linguistic terms is made.

(a) Choosing the family of membership functions.

(b) Choosing the parameter values for each of the membership functions.

(c) Adjusting these parameter values.

2.2 Membership Function

A graph that defines now each point in the input space is mapped to membership value between 0 and 1. Input space is often referred to as the universe of discourse or universal set,which contains all the possible elements of concern in each particular application.

There are three types of fuzzifier

(1) Singleton fuzzifier.

(2) Gaussian fuzzifier.

(3) Trapezoidal or triangular fuzzifier.

2.3 Fuzzy inference systems

Fuzzy Inference Systems (FIS) are popular computing frameworks based on the concept of fuzzy set theory. Their success is mainly due to their closeness to human perception and reasoning, as well as their intuitive handling and simplicity, which are important factors for acceptance and usability of the systems. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping provides a basis from which decisions can be made. The process of fuzzy inference involves formulating Membership Functions, Logical Operations, and IF–THEN rules. Because of its multidisciplinary nature, FIS are associated with a number of names, such as fuzzy rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and fuzzy systems. It provides a natural framework to incorporate fuzzy IF–THEN rules from human experts. There is much freedom in the choices of fuzzifier, fuzzy inference engine, and defuzzifier, so that we may obtain the most suitable fuzzy logic system for a particular problem. Figure 1, illustrates the basic architecture of an FIS. The main components are a fuzzification interface, a fuzzy rule base (knowledge base), an inference engine (decision-making unit), and a defuzzification interface. The input variables are fuzzified whereby the membership functions defined on the input variables are applied to their actual values, to determine the degree of truth for each rule antecedent. Fuzzy if-then rules and fuzzy reasoning are the backbone of FIS. The fuzzy rule base is characterized in the form of if-then rules in which the antecedents and consequences involve linguistic variables. The collection of these fuzzy rules forms the rule base for the fuzzy logic system. Using suitable inference procedure, the truth-value for the antecedent of each rule is computed, and applied to the consequent part of each rule. This results in one fuzzy subset to be assigned to each output variable for each rule. Usually, the rule base and the database are jointly referred to as the knowledge base. Again, by using suitable composition procedure, all the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable. Finally, defuzzification is applied to convert the fuzzy output set to a crisp output.

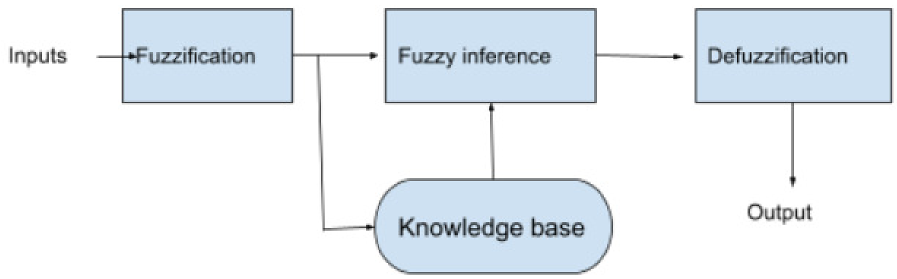


Figure.1. Fuzzy system

2.4 Fuzzy Logic

Fuzzy logic was first proposed by Lotfi Zadeh in 1965. The term fuzzy refers to things that are not clear or are vague. In the real world many times we encounter a situation when we can't determine whether the state is true or false, their fuzzy logic provides very valuable flexibility for reasoning. In this way , we consider the inaccuracies and uncertainties of any situation.

In the boolean system truth value, 1.0 represents the absolute truth value and 0.0 represents the absolute false value. But in the fuzzy system, there is no logic for the absolute truth and absolute false value. But in fuzzy logic, there is an intermediate value, to present which is partially true and partially false.

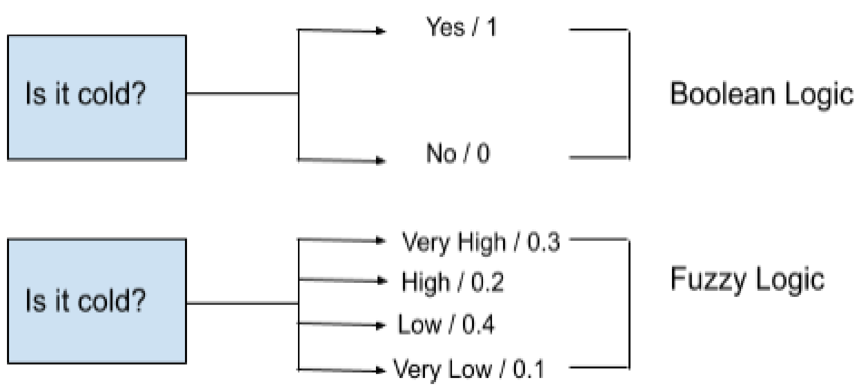


Figure.2. Fuzzy vs boolean representation

3. FUZZY RULE BASED SYSTEM

Fuzzy linguistic descriptions are formal representations of systems made through fuzzy IF-THEN rules. They encode knowledge about a system in statements of the form IF(a set of conditions) are satisfied THEN(a set of consequences) can be inferred. Fuzzy IF-THEN rules are coded in the form

IF (x_1 is A1, x_2 is A2 ,....,x_n is An) THEN (y_1 is B1, y_2 is B2 ,..., y_n is Bn).

Where linguistic variables x_i , y_j take the values of fuzzy sets Ai and Bj respectively.

A collection of rules referring to a particular system is known as a fuzzy rule base. If the conclusion C to be drawn from a rule base R is the conjunction of all the individual consequents Ci of each rule, then

$$C = C1 \cap C2 \cap \ldots. \cap Cn$$

Where

$$\mu_c(y) = \min(\mu_{c1}(y), \mu_{c2}(y), \ldots., \mu_{cn}(y)), \qquad \forall\, y \in Y$$

Where Y is the universe of discourse.

On the other hand, if the conclusion C to be drawn from a rule base R is the disjunction of the individual consequences of each rule , then

$$C = C1 \cup C2 \cup \ldots. \cup Cn$$

where

$$\mu_c(y) = \max(\mu_{c1}(y), \mu_{c2}(y), \ldots., \mu_{cn}(y)), \qquad \forall\, y \in Y$$

Where Y is the universe of discourse.

3.1 Application:

Deciding the tip amount to the server in the hotel using a Fuzzy rule based system.

The rules are framed under the quality of the food and service:

Rule 1: If (quality is good) and (service is poor) then MT

Rule 2: If (quality is good) and (service is average) then MT

Rule 3: If (quality is good) and (service is good) then HT

Rule 4: If (quality is average) and (service is poor) then MT

Rule 5: If (quality is average) and (service is average) then MT

Rule 6: If (quality is average) and (service is good) then MT

Rule 7: If (quality is poor) and (service is poor) then LT

Rule 8: If (quality is poor) and (service is average) then MT

Rule 9: If (quality is poor) and (service is good) then MT

Key:

HT - High Tip

MT - Medium Tip

LT - Less Tip

Computation of fuzzy membership value

For the fuzzification of inputs, that is , to compute the membership for the antecedents, the formula is illustrated in Figure 3.

Delta 1 = x-point 1

Delta 2 = point 2-x

If (Delta 1 ≤ 0) or (Delta 2 ≤ 0) then degree of membership = 0

Else degree of membership = min{Delta 1 ∗ slope 1, Delta 2 ∗ slope 2, max }

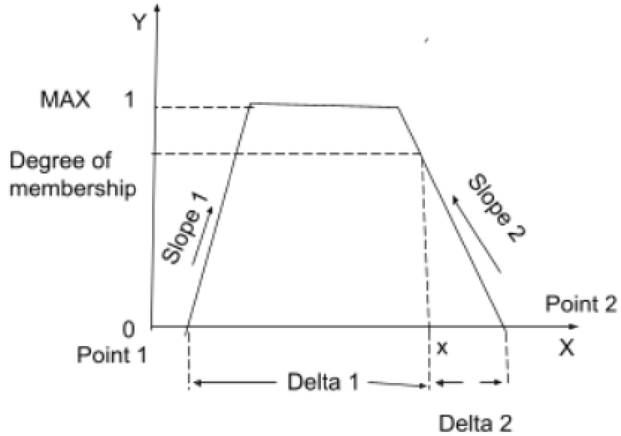


Figure 3. Computation of fuzzy membership value

Here ,x which is the system input has its membership function values computed for all fuzzy sets. For example, the system input quality of the food deals with three fuzzy sets, namely G - Good, A - Average and P - Poor. The other input service deals with three fuzzy sets, namely G - Good, A - Average and P - Poor.

The computation of the fuzzy membership values for the quality of food (x = 70), the qualifying fuzzy sets are shown in the figure 4.

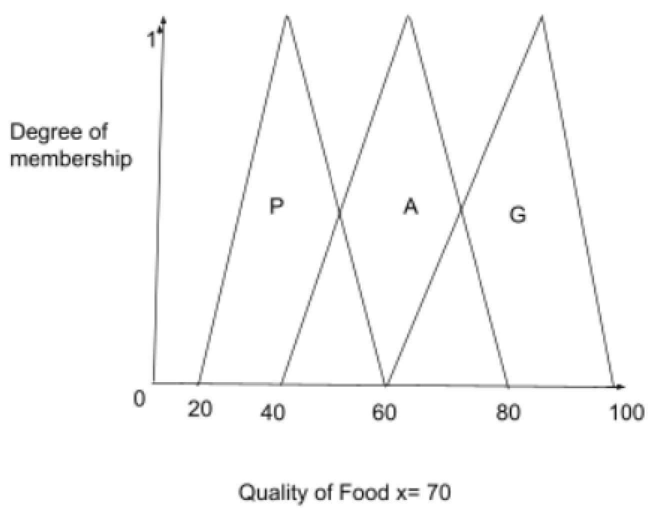


Figure 4.

Fuzzy membership functions of x = 70 for the fuzzy set G-Good.

Delta 1 = 70 - 60 = 10

Delta 2 = 100 - 70 = 30

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(10 ∗ 0.05, 30 ∗ 0.05, 1)

= min(0.5, 1.5, 1)

= 0.5

Fuzzy membership functions of x = 70 for the fuzzy set A - Average.

Delta 1 = 70 - 40 = 30

Delta 2 = 80 - 70 = 10

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_A(x) = min(30 ∗ 0.05, 10 ∗ 0.05, 1)

= min(1.5, 0.5, 1)

= 0.5

Fuzzy membership functions of x = 70 for the fuzzy set P - Poor.

Delta 1 = 70 - 20 = 50

Delta 2 = 60 - 70 = -10

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_P(x) =0

The computation of the fuzzy membership values for the food service (x = 65), the qualifying fuzzy sets are shown in the figure 5.

Fuzzy membership functions of x = 65 for the fuzzy set G-Good.

Delta 1 = 65 - 60 = 5

Delta 2 = 100 - 65 = 35

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(5 ∗ 0.05, 35 ∗ 0.05, 1)

= min(0.25, 1.75, 1) = 0.25

Fuzzy membership functions of x = 65 for the fuzzy set A - Average.

Delta 1 = 65 - 40 = 25

Delta 2 = 80 - 65 = 15

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(25 ∗ 0.05, 15 ∗ 0.05, 1)

= min(1.25, 0.75, 1)

= 0.25

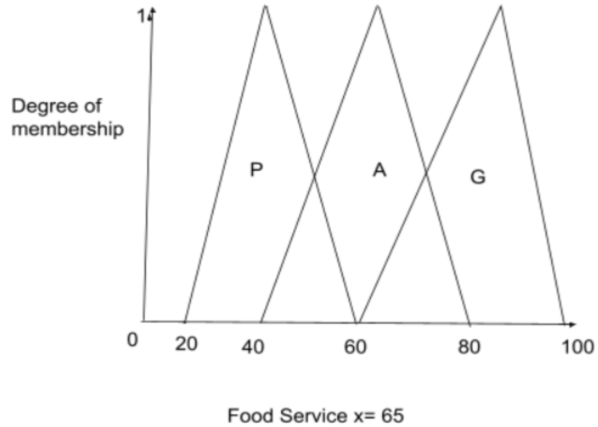


Figure 5.

Fuzzy membership functions of x = 65 for the fuzzy set P - Poor.

Delta 1 = 65 - 20 = 45

Delta 2 = 40 - 65 = -25

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_P(x) =0

For the Fuzzy rule base, the fuzzy membership values are

Rule 1 : min(0.5, 0) = 0

Rule 2 : min(0.5, 0.25) = 0.5

Rule 3 : min(0.5, 0.25) = 0.5

Rule 4 : min(0.5, 0) = 0

Rule 5 : min(0.5, 0.25) = 0.5

Rule 6 : min(0.5, 0.25) = 0.5

Rule 7 : min(0, 0) = 0

Rule 8 : min(0, 0.25) = 0

Rule 9 : min(0, 0.25) = 0

The fuzzy output is max(0.5, 0.5, 0.5, 0.5) = 0.5

The fuzzy outputs are those of rules 2, 3, 5 and 6 with strengths of 0.5

That is,

1. Quality is good and service is average.
2. Quality is good and service is good.
3. Quality is average and service is average.
4. Quality is average and service is good.

4. CONCLUSION

In this paper the general description and requirements for designing and creating a decision support system based on fuzzy logic are presented. Fuzzy inference systems can be applied in a vast number of meteorological application areas. An important advantage of the fuzzy expert system is that the knowledge is expressed in the form of an easy to understand linguistic fuzzy model, while maintaining the approximation accuracy at a reasonable level. This goal is achieved by modifying the rule antecedents to produce a flexible and interpretable output space. This model demonstrated that decision making on this method provides a better evaluation. Results indicate that fuzzy rule- based modeling is a promising alternative to the traditional approach.

REFERENCES

1. *J. Alcalá-Fdez and J.M. Alonso. A survey of fuzzy systems software: taxonomy, current research trends, and prospects. IEEE Transactions on Fuzzy Systems, 24(1):40–56, 2016.*
2. *R.E. Bellman and L.A. Zadeh. Decision-making in a fuzzy environment. Management Science, 17(4):B–141–B–164, 1970.*
3. *A. Kendall. Fuzzy Expert Systems. CRC Press, Inc., Boca Raton, FL - USA, 1991. Did not find it on WEB.*
4. *A. Konar. Artificial Intelligence and Soft Computing: Behavioral and Cognitive Modeling of the Human brain. CRC Press, Inc., Boca Raton, FL - USA, 2000.*
5. *Penka V . Georgieva. Fuzzy Rule - Based System For Decision- Making , Engineering Sciences , LIII,2016, No. 1.*
6. *A.K. Mitra . Sankar Nath . A.K. Sharma. Fog Forecasting using Rule-based Fuzzy Inference System, J. Indian Soc. Remote Sens. (September 2008) 36:243–253.*
7. *S.Rajasekaran and G. A. Vijayalakshmi Pai. Neural Networks, Fuzzy logic, And Genetic Algorithms: Synthesis and Applications. (December 2009)187-221.*

DOI No. : 10.5958/2394-9309.2022.00034.8
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A FUZZY RULE-BASED INFERENCE SYSTEM FOR DECISION MAKING IN REAL TIME APPLICATIONS

M. Vijaya*, A. Aruna**
*Head and Prof., PG and Research Dept of Mathematics, Marudupandiyar college, Thanjavur
**Ph.D. Scholar, Dept of Mathematics, Marudupandiyar college, Thanjavur
(Affiliated to Bharathidasan University, Trichy)
E-mail: mathvijaya23@gmail.com*, arunjo.jo21@gmail.com**

ABSTRACT

Fuzzy rule-based systems are one of the most important areas of applications of fuzzy sets and fuzzy logic. Fuzzy linguistic descriptions are formal representations of systems made through IF-THEN rules. In this paper, we are going to study fuzzy inference systems(FIS) and solve real time application problems using Fuzzy rule-based inference systems.

Keywords : Fuzzy rule-based, fuzzy inference system, fuzzy set

1. INTRODUCTION

Decision-making is an essential activity for an individual or in any organization. It can be considered as a procedure of choosing the best alternatives by considering every benefit and cost factor. As the problem of making decisions becomes complex, a more detailed process is necessary to find its solution. Fuzzy Rule Based systems are widely used in various real applications such as in anomaly intrusion detection, image processing, medical applications and in many other applications. In such applications, a large number of patterns and variables are present in the useful data or in the available data. This condition leads to a rapid increase of the fuzzy rule search space which in turn affects the inductive learning of FRBC systems. With this rapid growth, the learning process becomes further complicated, which results in scalability and complexity problems. In this paper we are going to solve one real time application using a fuzzy rule based inference system.

2. FUZZY OVERVIEW

2.1.Fuzzy Modeling and Fuzzy Systems

The main concepts in fuzzy logic are linguistic variables, linguistic modifiers, propositional fuzzy logic, deductive inference rules and approximations. Zadeh proposed the concept of linguistic variables for modeling human way of thinking, based on the following principle.With the increasing complexity of a system, our ability to make accurate and at the same time rational conclusions about its behavior is reduced till its reaches a threshold after which the accuracy and consistency are almost mutually exclusive requirements[2].The process of creating a fuzzy system, called fuzzy modeling .The structure of the fuzzy system has to be designed in such a way , so that the experience of the experts can be easily implemented in it. In case the input and output data are known , the possibility of identifying the system using standard techniques has to exist[3].

There are two milestones in fuzzy modeling:

i. Identification of the overall structure. Here information from experts or information obtained after "trials and errors" is used. The steps are

(a) Selecting appropriate input and output variables

(b) Choosing the type of fuzzy inference system.

(c) Determining the number of linguistic terms of input and output fuzzy variables.

(d) Designing the set of if-then rules.

ii. Identification of the structure in depth. Here a detailed description of linguistic terms is made.

(a) Choosing the family of membership functions.

(b) Choosing the parameter values for each of the membership functions.

(c) Adjusting these parameter values.

2.2 Membership Function

A graph that defines now each point in the input space is mapped to membership value between 0 and 1. Input space is often referred to as the universe of discourse or universal set,which contains all the possible elements of concern in each particular application.

There are three types of fuzzifier

(1) Singleton fuzzifier.

(2) Gaussian fuzzifier.

(3) Trapezoidal or triangular fuzzifier.

2.3 Fuzzy inference systems

Fuzzy Inference Systems (FIS) are popular computing frameworks based on the concept of fuzzy set theory. Their success is mainly due to their closeness to human perception and reasoning, as well as their intuitive handling and simplicity, which are important factors for acceptance and usability of the systems. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping provides a basis from which decisions can be made. The process of fuzzy inference involves formulating Membership Functions, Logical Operations, and IF–THEN rules. Because of its multidisciplinary nature, FIS are associated with a number of names, such as fuzzy rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and fuzzy systems. It provides a natural framework to incorporate fuzzy IF–THEN rules from human experts. There is much freedom in the choices of fuzzifier, fuzzy inference engine, and defuzzifier, so that we may obtain the most suitable fuzzy logic system for a particular problem. Figure 1, illustrates the basic architecture of an FIS. The main components are a fuzzification interface, a fuzzy rule base (knowledge base), an inference engine (decision-making unit), and a defuzzification interface. The input variables are fuzzified whereby the membership functions defined on the input variables are applied to their actual values, to determine the degree of truth for each rule antecedent. Fuzzy if-then rules and fuzzy reasoning are the backbone of FIS. The fuzzy rule base is characterized in the form of if-then rules in which the antecedents and consequences involve linguistic variables. The collection of these fuzzy rules forms the rule base for the fuzzy logic system. Using suitable inference procedure, the truth-value for the antecedent of each rule is computed, and applied to the consequent part of each rule. This results in one fuzzy subset to be assigned to each output variable for each rule. Usually, the rule base and the database are jointly referred to as the knowledge base. Again, by using suitable composition procedure, all the fuzzy subsets assigned to each output variable are combined together to form a single fuzzy subset for each output variable. Finally, defuzzification is applied to convert the fuzzy output set to a crisp output.

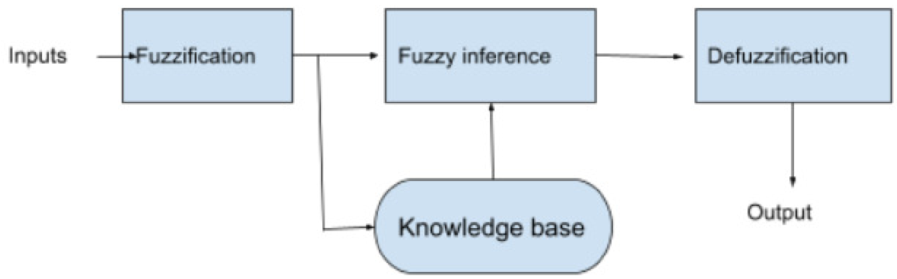


Figure.1. Fuzzy system

2.4 Fuzzy Logic

Fuzzy logic was first proposed by Lotfi Zadeh in 1965. The term fuzzy refers to things that are not clear or are vague. In the real world many times we encounter a situation when we can't determine whether the state is true or false, their fuzzy logic provides very valuable flexibility for reasoning. In this way , we consider the inaccuracies and uncertainties of any situation.

In the boolean system truth value, 1.0 represents the absolute truth value and 0.0 represents the absolute false value. But in the fuzzy system, there is no logic for the absolute truth and absolute false value. But in fuzzy logic, there is an intermediate value, to present which is partially true and partially false.

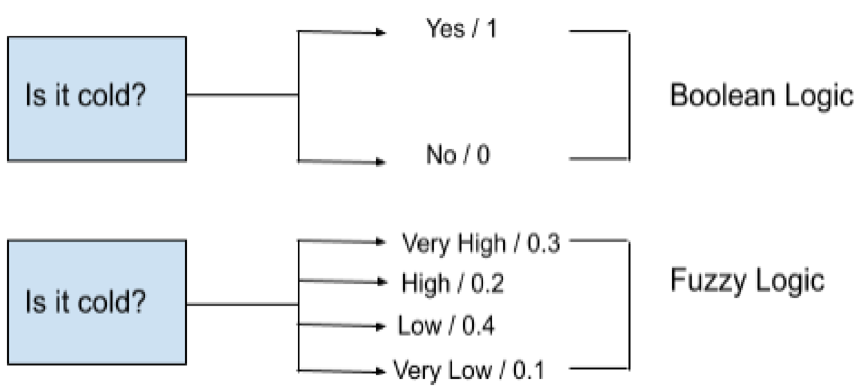


Figure.2. Fuzzy vs boolean representation

3. FUZZY RULE BASED SYSTEM

Fuzzy linguistic descriptions are formal representations of systems made through fuzzy IF-THEN rules. They encode knowledge about a system in statements of the form IF(a set of conditions) are satisfied THEN(a set of consequences) can be inferred. Fuzzy IF-THEN rules are coded in the form

IF (x_1 is A1, x_2 is A2 ,....,x_n is An) THEN (y_1 is B1, y_2 is B2 ,..., y_n is Bn).

Where linguistic variables x_i , y_j take the values of fuzzy sets Ai and Bj respectively.

A collection of rules referring to a particular system is known as a fuzzy rule base. If the conclusion C to be drawn from a rule base R is the conjunction of all the individual consequents Ci of each rule, then

$$C = C1 \cap C2 \cap \ldots. \cap Cn$$

Where

$$\mu_c(y) = \min(\mu_{c1}(y), \mu_{c2}(y), \ldots, \mu_{cn}(y)), \qquad \forall\, y \in Y$$

Where Y is the universe of discourse.

On the other hand, if the conclusion C to be drawn from a rule base R is the disjunction of the individual consequences of each rule , then

$$C = C1 \cup C2 \cup \ldots. \cup Cn$$

where

$$\mu_c(y) = \max(\mu_{c1}(y), \mu_{c2}(y), \ldots, \mu_{cn}(y)), \qquad \forall\, y \in Y$$

Where Y is the universe of discourse.

3.1 Application:

Deciding the tip amount to the server in the hotel using a Fuzzy rule based system.

The rules are framed under the quality of the food and service:

Rule 1: If (quality is good) and (service is poor) then MT

Rule 2: If (quality is good) and (service is average) then MT

Rule 3: If (quality is good) and (service is good) then HT

Rule 4: If (quality is average) and (service is poor) then MT

Rule 5: If (quality is average) and (service is average) then MT

Rule 6: If (quality is average) and (service is good) then MT

Rule 7: If (quality is poor) and (service is poor) then LT

Rule 8: If (quality is poor) and (service is average) then MT

Rule 9: If (quality is poor) and (service is good) then MT

Key:

HT - High Tip

MT - Medium Tip

LT - Less Tip

Computation of fuzzy membership value

For the fuzzification of inputs, that is , to compute the membership for the antecedents, the formula is illustrated in Figure 3.

Delta 1 = x-point 1

Delta 2 = point 2-x

If (Delta 1 ≤ 0) or (Delta 2 ≤ 0) then degree of membership = 0

Else degree of membership = min{Delta 1 ∗ slope 1, Delta 2 ∗ slope 2, max }

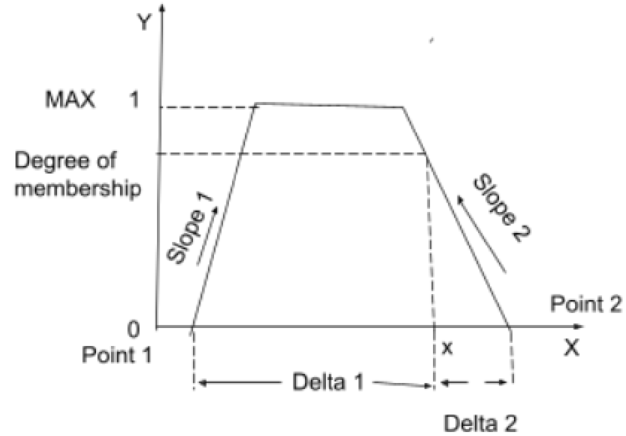


Figure 3. Computation of fuzzy membership value

Here ,x which is the system input has its membership function values computed for all fuzzy sets. For example, the system input quality of the food deals with three fuzzy sets, namely G - Good, A - Average and P - Poor. The other input service deals with three fuzzy sets, namely G - Good, A - Average and P - Poor.

The computation of the fuzzy membership values for the quality of food (x = 70), the qualifying fuzzy sets are shown in the figure 4.

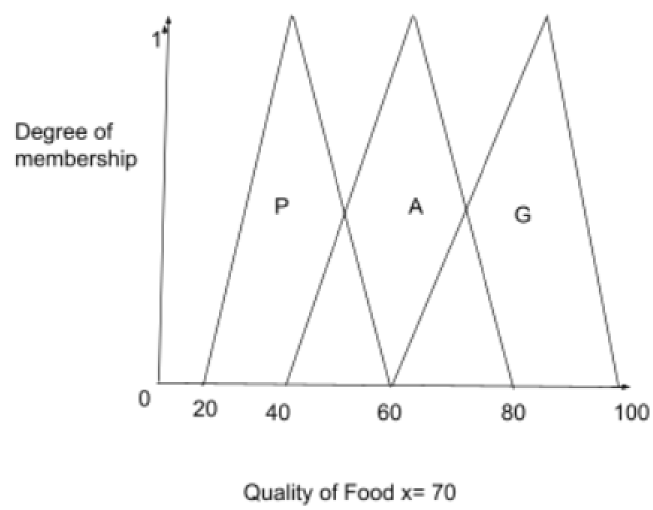


Figure 4.

Fuzzy membership functions of x = 70 for the fuzzy set G-Good.

Delta 1 = 70 - 60 = 10

Delta 2 = 100 - 70 = 30

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(10 ∗ 0.05, 30 ∗ 0.05, 1)

= min(0.5, 1.5, 1)

= 0.5

Fuzzy membership functions of x = 70 for the fuzzy set A - Average.

Delta 1 = 70 - 40 = 30

Delta 2 = 80 - 70 = 10

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_A(x) = min(30 ∗ 0.05, 10 ∗ 0.05, 1)

= min(1.5, 0.5, 1)

= 0.5

Fuzzy membership functions of x = 70 for the fuzzy set P - Poor.

Delta 1 = 70 - 20 = 50

Delta 2 = 60 - 70 = -10

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_P(x) =0

The computation of the fuzzy membership values for the food service (x = 65), the qualifying fuzzy sets are shown in the figure 5.

Fuzzy membership functions of x = 65 for the fuzzy set G-Good.

Delta 1 = 65 - 60 = 5

Delta 2 = 100 - 65 = 35

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(5 ∗ 0.05, 35 ∗ 0.05, 1)

= min(0.25, 1.75, 1) = 0.25

Fuzzy membership functions of x = 65 for the fuzzy set A - Average.

Delta 1 = 65 - 40 = 25

Delta 2 = 80 - 65 = 15

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_G(x) = min(25 ∗ 0.05, 15 ∗ 0.05, 1)

= min(1.25, 0.75, 1)

= 0.25

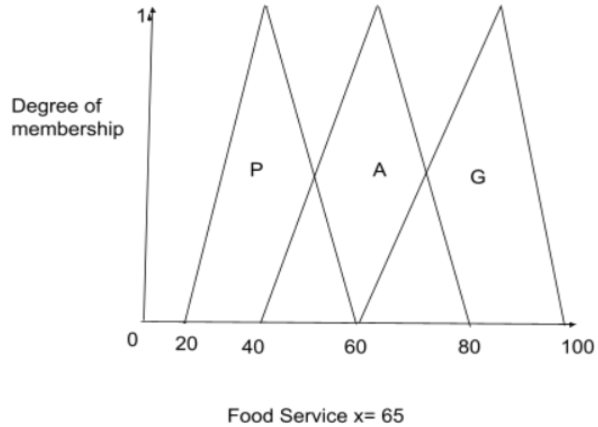


Figure 5.

Fuzzy membership functions of x = 65 for the fuzzy set P - Poor.

Delta 1 = 65 - 20 = 45

Delta 2 = 40 - 65 = -25

Slope 1 = 1/20 = 0.05

Slope 2 = 1/20 = 0.05

μ_P(x) =0

For the Fuzzy rule base, the fuzzy membership values are

Rule 1 : min(0.5, 0) = 0

Rule 2 : min(0.5, 0.25) = 0.5

Rule 3 : min(0.5, 0.25) = 0.5

Rule 4 : min(0.5, 0) = 0

Rule 5 : min(0.5, 0.25) = 0.5

Rule 6 : min(0.5, 0.25) = 0.5

Rule 7 : min(0, 0) = 0

Rule 8 : min(0, 0.25) = 0

Rule 9 : min(0, 0.25) = 0

The fuzzy output is max(0.5, 0.5, 0.5, 0.5) = 0.5

The fuzzy outputs are those of rules 2, 3, 5 and 6 with strengths of 0.5

That is,

1. Quality is good and service is average.
2. Quality is good and service is good.
3. Quality is average and service is average.
4. Quality is average and service is good.

4. CONCLUSION

In this paper the general description and requirements for designing and creating a decision support system based on fuzzy logic are presented. Fuzzy inference systems can be applied in a vast number of meteorological application areas. An important advantage of the fuzzy expert system is that the knowledge is expressed in the form of an easy to understand linguistic fuzzy model, while maintaining the approximation accuracy at a reasonable level. This goal is achieved by modifying the rule antecedents to produce a flexible and interpretable output space. This model demonstrated that decision making on this method provides a better evaluation. Results indicate that fuzzy rule- based modeling is a promising alternative to the traditional approach.

REFERENCES

1. *J. Alcalá-Fdez and J.M. Alonso. A survey of fuzzy systems software: taxonomy, current research trends, and prospects. IEEE Transactions on Fuzzy Systems, 24(1):40–56, 2016.*
2. *R.E. Bellman and L.A. Zadeh. Decision-making in a fuzzy environment. Management Science, 17(4):B–141–B–164, 1970.*
3. *A. Kendall. Fuzzy Expert Systems. CRC Press, Inc., Boca Raton, FL - USA, 1991. Did not find it on WEB.*
4. *A. Konar. Artificial Intelligence and Soft Computing: Behavioral and Cognitive Modeling of the Human brain. CRC Press, Inc., Boca Raton, FL - USA, 2000.*
5. *Penka V . Georgieva. Fuzzy Rule - Based System For Decision- Making , Engineering Sciences , LIII,2016, No. 1.*
6. *A.K. Mitra . Sankar Nath . A.K. Sharma. Fog Forecasting using Rule-based Fuzzy Inference System, J. Indian Soc. Remote Sens. (September 2008) 36:243–253.*
7. *S.Rajasekaran and G. A. Vijayalakshmi Pai. Neural Networks, Fuzzy logic, And Genetic Algorithms: Synthesis and Applications. (December 2009)187-221.*

Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
DOI No. : 10.5958/2394-9309.2022.00035.X
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

DATA MINING APPLICATIONS AND VISUALIZATION OF HEALTHCARE DATABASE

G. Manimannan[*], N. Paranjothi[**] and A. Poongothai[***]
[*]Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai.
[**]Assosicate Professor, Department of Statistics, Annamalai University, Chidambaram.
[***]Research Scholar, Department of Statistics, Annamalai University, Chidambaram.

ABSTRACT
This paper is attempted to identify hidden structure, classification and clustering of Healthcare data. In healthcare, data mining is becoming more and more popular, if not increasingly essential. Data mining applications can significantly benefit all parties involved in the healthcare industry. The huge amounts of data generated by healthcare transactions are too complex and voluminous to be processed and analyzed by traditional methods. Data mining provides the methodology and technology to transform these mounds of data into useful information for decision making. This study explores data mining applications in healthcare. In particular, this area was chosen as a model to study Master Health Checkup (MHC). The data were collected from secondary source containing 295 patients in St. Johns Hospital, Bangalore. The case sheet deals with socio demographic characteristic, Blood Pressure, Fat, Liver and diabetic related parameters. The salient feature of this study is the application of Factor Analysis, K-means clustering and Self Organizing Map (SOM) as data mining tools to develop the hidden structure present in the data. The scores from extracted factors are used to find initial groups by K-means clustering algorithm. A few outlier health care profiles, which could not be classified to any of the larger groups, are discarded as some of the parameters possessed higher values. Finally, SOM is applied and the groups are identified as MHC patients belonging to O-Class (Obesity), N-Class (Normal) and UW-Class (Under Weight) in that order. The results of the study indicate that SOM can be a feasible tool for the health care analysis of large amounts of master health checkup data.
Keywords: Master Health Check (MCH), Data mining, Factor Analysis, k-means Clustering and Self Organizing Map (SOM)

1. BACKGROUND OF THE STUDY

The Master Health Check-up (MHC) offered by various hospitals and medical research institutes is a programme that attempts to reduce health care costs by prevention and early diagnosis. A variety of chronic diseases afflict us, most of which take their toll after the fifth decade of life. Diabetes, hypertension, heart attacks, stroke and cancer are some of the more common examples. Almost all of these problems first go through a long quiescent phase where they produce no symptoms. This period can be as long as 10 - 20 years. It makes sense, therefore, that a programme that attempts to detect and correct these problems during this silent phase, will decrease the ultimate morbidity from these diseases.

In the early days of preventive health check-ups, every conceivable test and technology was ordered in the hope that some would be abnormal and provide an avenue of approach. A handful of items, mostly simple, appear to provide the greatest value. The MHC offered at various hospitals and institutes is a carefully constructed programme that offers a panel of tests that are proven to be valuable.

As an incentive to those who have taken the efforts to control their health problems, the programme also includes two or more follow up visits within a year of the MHC and the physician in change of the check up. Good health is by itself of great value. It enhances market earnings by increasing the number of healthy days an individual has available for work (Grossman 1972) and increases nonmarket productivity, allowing more time for household production (Becker 1976). Health checkups help to secure and maintain good health.

The Master Health Check (MHC) is a series of tests to screen each functional area closely to detect even the smallest symptom of a major illness. It also helps to identify the reason for minor ailments, which are constant. MHC is considered to be the most comprehensive prevention check. Master Health Check consists of five permanent packages, which are as follows: Master Health Check, Executive Health Check, Heart Check, Whole Body Check and Well Women Check (B. Krishan Reddy, G.V.R.K. Acharyulu , 2002).

In the present context the problem of MHC patients has been studied, without making any assumptions with regard to the number of groups or any other structural patterns in advance, which reflected the classification of patients based on certain medical observations (G. Manimannan, S. Hari and G. Vijaythiraviyam). The main objective of this paper is to investigate whether;

(i) Data mining paradigms together with well known *unsupervised* learning neural net model *Self Organizing Map* (SOM) can be used to exhibits the classification of MHC patients.

(ii) A visual representation reveals the clusters on a two-dimensional topological map.

The rest of the paper is organized as follows. Section 2 describes the methodology we have used, the database and the choice of MHC parameters. Section 3 presents the proposed algorithm which is used as a benchmark to achieve the objective on applying one of the well known neural network model SOM and Section 4 presents the empirical results. The conclusions of our study are presented in Section 5.

2. DATABASE

This section brings out the discussion of the database, the MHC (Master Health Checkup) parameters selected and the Data Mining Techniques. The MHC data were collected from secondary source of OPD (Out Patients Department) containing 295 patients in St. Johns Hospital, Bangalore was considered as the database. The data mainly consists of five major categories, such as socio economic and demographic characteristic, Blood Pressure, Fat, Liver and diabetic related parameters. Among the listed patients, number of patients varied over the study period owing to removal of those patients for which the required data are not available or outliers.

A. Selection of Variables

In this study, 28 medical observations (parameters) were chosen among the many that had been used in MHC case sheets. These 28 medical observations were chosen to assess socio economic and demographic characteristic, Blood Pressure, Fat, Liver and diabetic. Some of them are given below.

Table I. MHC Medical Observations during the Study Period

Parameters	Description
BP_Syst	Blood Pressure Systolic
BP_Dias	Blood Pressure diastolic
Blood_Hb	haemoglobin
Blood_PCV	Packed Cell Volume
Blood_TC	Total Count
Diabetess_Fasting	Diabetes Fasting
Diabetes_Post Pran	Diabetes Post Prandial
Cholesterol	Cholesterol
FAT_VLD	High-density lipoprotein
FAT_LDL	Low-density lipoprotein
Liver_SAP	Alkaline phosphate
Liver_ALT	Alanine transaminase

3. METHODOLOGIES

Data Mining or Knowledge Discovery in Databases (KDD) is the process of discovering previously unknown and potentially useful information from the data in databases. In the present context data mining exhibits the patterns by applying few techniques namely, factor analysis, **k**-means clustering and Self Organizing Map (SOM).

As such KDD is an iterative process, which mainly consist of the following steps on the data collected;

Step 1: Data cleaning

Step 2**:** Data Integration

Step 3: Data selection and transformation

Step 4: Data Mining

Step 5: Knowledge representation

Of these above iterative process Steps 4 and 5 are most important. If suitable techniques are applied in Step 5, it provides potentially useful information that explains the hidden structure. This structure discovers knowledge that is represented visually to the user, which is the final phase of data mining.

3.1 Factor Analysis

Factor analysis provides the tools for analyzing the structure of the interrelationships (correlations) among the large number of variables by defining sets of variables known as factors. In the present study, factor analysis is initiated to uncover the patterns underlying MHC medical observations. Orthogonal rotations such as Varimax and Quartimax rotations are used to measure the similarity of a variable with a factor by its factor loading.

3.2. k-Means Clustering Methods

McQueen (1967) suggests the term k-means for describing an algorithm of his that assigns item to the cluster having the nearest centroid (mean). Generally this technique uses Euclidean distances measures computed by variables. Since the group labels are unknown for the data set, k-means clustering is one such technique in applied statistics that discovers acceptable meaningful classes.

3.3 Self Organizing Map

SOM is an unsupervised learning procedure based on artificial neural network. It is very effective and frequently used network popularly known as ***Kohonen's*** neural network. These networks have only two layers, a standard input layer and an output layer known as the *Competitive* (Kohonen) layer. Each input neuron is connected to each and every neuron on the competitive layer which are organized as a two dimensional grid. Each MHC medical observations is associated with exactly one neuron whereas each neuron may have one or more medical observations attributed to it. This grid map enables to discover statistical regularities in its input space and develops different modes of behaviour to represent different classes of input databases.

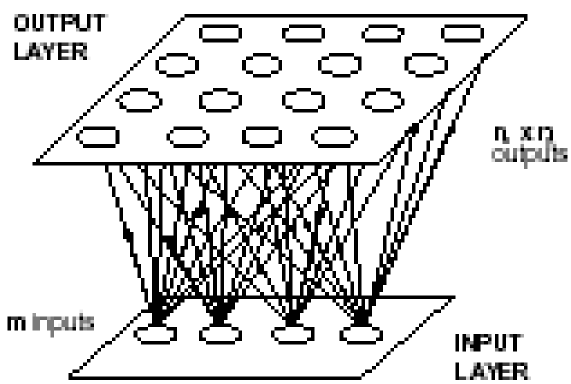


Figure 1. SOM Network

3.4 Proposed Algorithms

A brief step-by-step algorithm to classify the MHC patients during the study period based on their overall MHC is described below:

For the pruned data set the following algorithm is proposed to scale the MHC patients and visualize them on a two-dimensional map during each of the study period based on their overall medical observations (*Table 1*).

Step 1: A map of weight vectors with 295 x 10 neurons having hexagonal topology for neighborhood is obtained using SOMPAK.

Step 2: Factor analysis is initiated to find the structural pattern underlying the data set.

Step 3: **k** –means analysis is used to partition the data set into **k**-clusters using the factor scores obtained in **Step 2** as input.

Step 4: Construct a SOM using the prototype vector with appropriate hits of the companies in the data set that are assigned group labels in step 3.

4. RESULTS AND DISCUSSION

Factor analysis is extended with the techniques of Varimax and Quartimax criterion for orthogonal rotation. Even though the results obtained by both the criterions were very similar, the varimax rotation provided relatively better clustering of MHC medical observations.

Consequently, only the results of varimax rotation are reported here. We have decided to retain 73 percent of total variation in the data, and thus accounted consistently ten factors for MHC medical observations with eigen values little less than or equal to unity. *Table 2* shows variance accounted for each factors

Table 2. Percentage of Variance Explained by Factors

Factors	Variance Explained
1	17.149
2	9.256
3	6.592
4	6.168
5	5.962
6	5.776
7	5.760
8	5.304
9	5.050
10	4.599
Total	72.616

From the above table we observe that the total variances explained by the extracted factors are over 65 percent, which are relatively higher.

After performing factor analysis, the next stage is to assign initial group labels to MHC patients. Step 3 of the algorithm is explored with factor score extracted by Step 2, by conventional **k**-means clustering analysis. Formations of clusters are explored by considering 2-clusters, 3-clusters, 4-cluster and so on. Isolated groups with some MHC patients are discarded from the analysis as outliers. A few MHC for these outlier patients are comparatively high or low to those excelled in the analysis. Out of all the possible trials, 3-cluster exhibited meaningful interpretation than two, four and higher clusters. Having decided to consider only 3 clusters, it is possible to classify MHC patients as Cluster **N**, Cluster **UW** or Cluster **O** depending on whether the MHC patients belonged to Cluster 1, Cluster 2 or Cluster 3 respectively.

Cluster 1 (Cluster **N**) is a group of MHC patients that have high values for the MHC parameters, indicating that these patients are normal. The **O** with lower values for the MHC medical observations are grouped into Cluster 3 (Cluster O). This suggested that Cluster 3 is a group of patients with low-profile. Cluster 2 (Cluster **UW**) are those patients which perform moderately well as compared to the Cluster 1 and Cluster 3.

Inspite of incorporating the results for MHC patients, only the summary statistics are reported in *Table 3*. The first column in *Table 3* provides the groupings done by cluster analysis.

Table 3. Number of MHC Patients in Three Clusters

MHC Patients	k-Means			Self-Organizing Map		
	1	2	3	1	2	3
295	115	50	131	117	49	129

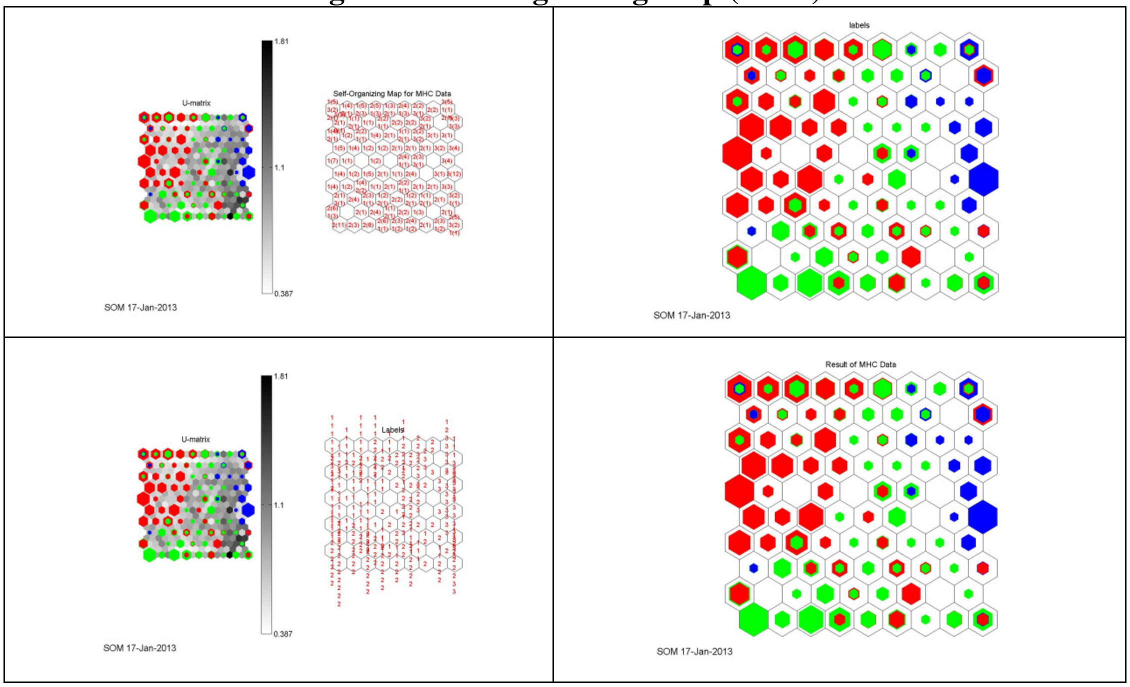
1– Scale **N** 2 – Scale **UW** 3 – Scale **O**

Figures 2 show the groupings of MHC patients into 3 clusters for the study period. Patients in cluster 1 tend to be normal, cluster 2 tends to be under weight and cluster 3 tends to be obesity. We classify the members in the first cluster as Cluster **N**, the second as Cluster **UW** and the third as Cluster **O** in terms of MHC medical parameters.

The pruned data set is then subjected to the main algorithm as in *Section 3.4* to assign appropriate classes to the MHC patients. Initially, a SOM was trained separately, with the sequential procedure algorithm, using the SOM Toolbox version 2 for MATLAB. In the construction process several maps are initialized and trained by considering the Gaussian neighbourhood function and a map with hexagonal topology. Among the prototype vector maps, the best ones in respect of average quantization error are carefully chosen to explore the data in different dimensions. From the study it is found that a weight vector map units of size 295 by 10 neurons is well spanned within the data set as in *Figure 2*.

In addition, SOM is used efficiently in data visualization due to its ability to represent the input data in two dimensions. Among the various visualization techniques the most widely used method for visualizing the cluster structure of the SOM is the distance matrix technique, especially the unified distance matrix (U-Matrix). For the present study, the method of displaying the number of hits in each map unit is well-thought off. *Figure* 2 show the groupings of MHC patients into 3 clusters over the SOM grid using the visualization method. In the following *Figures*, each colour (shades) represents classes of MHC patients.

Figure 2. Self-Organizing Map (SOM)



5. CONCLUSIONS

The purpose of this paper is to explore the possibility to identify the meaningful groups of MHC patients that are scaled as the best with respect to their medical observations (parameters) using SOM and few related classification techniques. Initially, factor analysis is used to identify the underlying structure based on 29 medical observations. The factor scores are used to partition the MHC patients into different clusters by using k-means clustering algorithm.

MHC patient's data is mined by choosing random weight vectors in a map of size 295 x 10 neurons with different training parameters and a final map with the least quantization error is obtained using SOM Toolbox version 2. Scaling of MHC patients is attempted based on certain medical observations applying the proposed algorithm. The ability of SOM provides several methods of representing the high-dimensional vectors to be projected to a low dimension preserving the inter-sample distances as faithfully as possible and essential topological properties. In the present study, unified distance matrix with corresponding hits of the data to the prototype vectors are illustrated in the above Figures 2. This enables us to visualize the clustering of MHC patients grouped into different categories of patients on a two dimensional map unambiguously. Hence, the members of Cluster 1 are labeled as Score N (Normal). Similarly, the Cluster 2 includes patients which performed moderately well being scored as UW (Under Weight) and the Cluster 3 with low-profile MHC patients with score O (Obesity).

The present analysis has shown that only 3 groups could be meaningfully formed for all the data. This indicates that only 3 types of patients existed over a study period. Further, the MHC patients find themselves classified into Normal (Score B), Under Weight (Scale UW) and Obesity (Scale O) categories depending on certain medical observations. A generalization of the results is under investigation to obtain an incorporated class of 3 groups of MHC patients for any study period.

REFERENCES

1. *Becker, Gary S., ed. 1976. A Theory of the Allocation of Time. In The economic approach to human behavior, 89-114. Chicago: University of Chicago Press.*
2. *Grossman, Michael. 1972. On the concept of health capital and the demand for health. Journal of Political Economy 80 (2): 223-55.*
3. *Hian Chye Koh and Gerald Tan, Data Mining Application in Health Care Journal of Healthcare Information Management - Vol. 19, No. 2*
4. *Juha Vesanto (1999), SOM - Based Visualization Methods, http://www.cis.hut.fi/ projecs/ide/publications.*
5. *Juha Vesanto (2000), Using SOM in Data Mining. http://www.cis.hut.fi/projecs/ide/publications.*
6. *B.Krishan Reddy, G.V.R.K. Acharyulu, (2002) Customer Relationship Management (CRM) in Health Care Sector - A Case Study on Master Health Check, Journal of the Academy of Hospital Administration, Vol. 14, No. 1 . (2002-01 - 2002-06).*
7. *J. B. MacQueen (1967): "Some Methods for Classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, 1:281-297.*
8. *G. Manimannan, S. Hari and G. Vijay Thiraviyam (2012), Data Mining Applications in Master Health Checkup: a Statistical Exploration, Paper presented at National Conference on Statistics for Twenty First Century-2012 (NCSTC-2012), Department of Statistics, University of Kerala, Trivandrum, Kerala, India.*
9. *Richard A Johnson and Dean W Wichern (1992), Applied Multivariate Statistical Analysis, 3/ed, Prentice-Hall of India Private Limited, New Delhi.*

DOI No. : 10.5958/2394-9309.2022.00036.1
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

A COMPARATIVE STUDY ON VARIOUS TIME SERIES MODEL AND ITS VISUALIZATION FOR THE SPREAD OF COVID-19 - TAMILNADU

Prathibaa Sreenivasan
Department of Biostatistics, SDNB Vaishnav College for Women, ChromepetChennai.
Email : prathibaa.ps@gmail.com

ABSTRACT

The coronavirus (COVID-19), which has been a pandemic for the past two years, is infecting around 212 countries and territories around the world. The World Health Organization (WHO) and the Ministry of Health and Family Welfare of India have released an overall overview and state-by-state status for COVID-19 across the world, which includes Active/Vaccinated/Recovered/Death counts. COVID-19 is highly contagious, and research suggests that it is transmitted from person to person by "spreading by touching surfaces, intimate contact, air, or items" containing viral particles. In the following study, I sought to forecast COVID-19's future conditions in order to mitigate its impact on the state of Tamil Nadu. From March 2020, data on COVID-19 confirmed, recovered, and death cases has been collected.

The representation of COVID cases, as well as their trend, is depicted in this study. Different regression and time series models are performed with the help of these trends to offer a statistical forecast for the confirmed cases, and their respective performances are compared to identify the best fit in order to calculate the forecasted trend line for the next 31 days. These time series models are malleable to model-dependent data and have been used to estimate and forecast a wide range of real-world issues.

According to the findings, the moving average model has the lowest root mean squared error value, which can provide a decent look at overall trends for projecting future values. Also included is an attempt to estimate the case fatality and recovery rates for Tamil Nadu's districts. This visualisation will help us to see how the epidemic has progressed over time and what the trend is. The forecasted line will show the direction of the trend line, allowing us to take precautionary measures to flatten the rising growth.

Keywords : Covid-19, Pandemic, Coronavirus, Trend, Plot, Time series model, Regression models, MA model, Forecast.

1. INTRODUCTION:

Coronavirus illness (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus family. Its signs and symptoms might range from a simple cold to a more serious disease. It started at a wholesale fish and seafood market in Wuhan, China, and quickly spread over the world, turning into a global pandemic. There have been over **33765502** confirmed cases of COVID-19 in India and 2663789 confirmed cases of COVID-19 in Tamil Nadu as of March-05,2022. Forecasting the pandemic trend plays a major role in controlling its future impact. While forecasting, it is very important to analyse its past patterns, as we know that "history repeats itself. “Several researchers have been applying various forecasting models like ARIMA time series model (ARIMA modelling & forecasting of COVID-19 in top five affected countries by Alok Kumar Sahai,NamitaRath 2,Vishal Sood,ManvendraPratap Singh);(Time series modelling to forecast the confirmed and recovered cases of COVID-19 by Mohsen Maleki,a Mohammad Reza Mahmoudi,b,c,DarrenWraith,d,Kim-Hung Phoe); Exponential forecasting method (Forecasting the novel coronavirus covid-19. PLoS One. 2020;15 by Petropoulos F.,Makridakis S.); Simulation optimization approach-retrospective forecast(Forecasting Peaks of Seasonal Influenza Epidemics by Elaine Nsoesie,Madhav Mararthe,John Brownstein, 2013 June 21.); SIRD(Susceptible-Infectious-Recovered-Dead) model to forecast evolution of outbreak (Data-based analysis, modeling and forecasting of the COVID-19 outbreak by Cleo Anastassopoulou, LuciaRusso, Athanasios Tsakris, Constantinos Siettos). But the accuracy of the forecasting is uncertain as it depends on various influences which remain unpredictable. Hence, several epidemiologists suggest comparative study, that is, comparing various models and choosing the best out of them, in order to get better accuracy, like comparison between MEM, RNN, and statistical models (Comparative study of a mathematical epidemic model, statistical modelling, and deep learning for COVID-19 forecasting and management by Mohammad Masum, M A Masud, HussainShahriar, and Sangil Kim); and comparison between deep learning models-RNN, LSTM, BiLSTM, GRU, and VAE (Deep learning methods for forecasting covid-19 time-series data: a comparative study-Solit Fractals, Chaos. 2020;140:110121., by Zeroual A., Harrou F., DairiA,Sun Y.)Therefore, in this study, I have attempted to compare the machine learning models (linear regression, polynomial regression, AR, MA, ARIMA, SARIMA, and Fb Prophet) and the best model is identified based on the model accuracy.

1.2 OBJECTIVE

- To visualize the trend of Covid-19 so far (from February 2020 to September 2021)
- To fit a predicted time series model for the covid-19 data.
- To compare various time series model and select the best model based on root mean squared error value.
- To project a forecasted trend line for next 31 days based on the best fitted model.

2. DATA DESCRIPTION:

The dataset was obtained from www.covid19india.org (which collects and compiles data from various news articles and government websites).It comprises daily covid-19 data of Tamil Nadu from February(2020) till march (2021)

- Sample size :573
- Source : secondary data
- Attributes : 3

2.1 TOOLS AND TECHNIQUES:

- Data Visualization using matplotlib and plotly python libraries
- Linear regression
- Polynomial regression
- Auto regressive model
- Moving average model
- Auto regressive integrated moving average model
- Seasonal auto regressive integrated moving average model
- Support vector regression
- Facebook prophet model
- Forecasted trend line using best fitted model

The best model is obtained using the accuracy metric, Root Mean Squared Error (RMSE). The model with the least RMSE was chosen to be the best with the least error. In this work, the time series models are used to forecast the number of new COVID-19 cases in Tamil Nadu based on the daily reported data. By empirically comparing multiple models in terms of their forecasting accuracy, I intend to suggest an appropriate model that can be utilised by society or governments to assess the near future of this outbreak. Since this is a complex forecasting problem, the fact is that the pandemic continues and there are many factors that we cannot presently control, the prediction accuracy will improve with time and more data.

S.NO	ATTRIBUTES	DESCRIPTION	DATA TYPE
1	Date	Date in the format DD/MM/YYYY	Datetime[ns]
2	Confirmed	Daily count of confirmed cases	numerical
3	Recovered	Daily count of recovered cases	numerical
3	Deceased	Daily count of deceased cases	numerical
4	Tested	Daily count of tested cases	numerical

Table 1: CATEGORISATION OF VARIABLES IN THE ANALYSIS

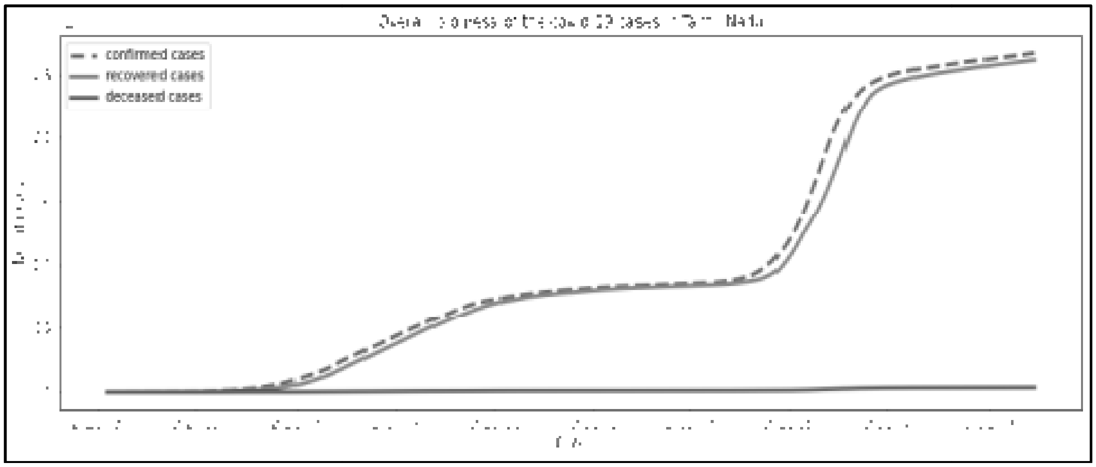


Figure 1: Overall Progress of the Covid crisis in Tamil Nadu

**Figure 1 shows the graphical representation of covid-19 trend in Tamil Nadu. It is observed that the increasing number of Covid confirmed cases and recoverednumbers simultaneously resultedin positive sign

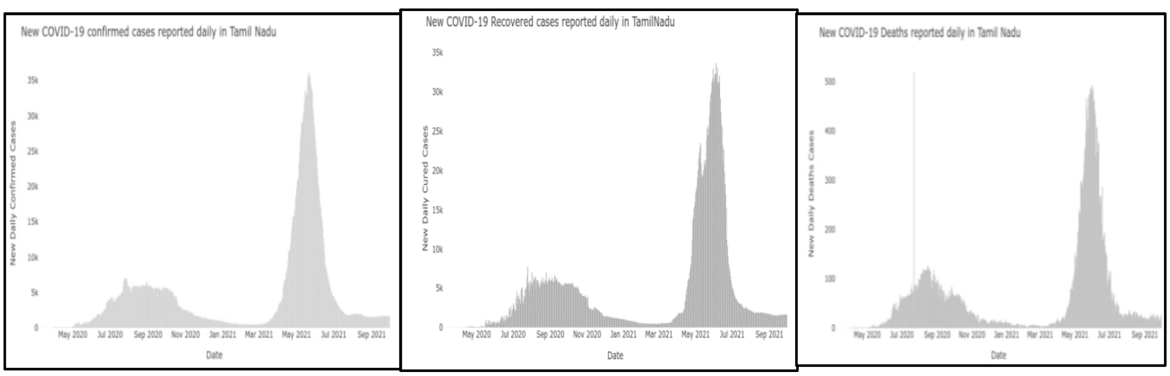


Figure 2: New Covid 19 Confirmed Cases, Recovered Cases, Death Cases

**Figure 2 shows that the Confirmed, Recovered and Death cases shown in figure 2 respectivelyis reached a peak in the month of May toJune2021;where second wave of covid-19 hits highlyin Tamil Nadu.Based on the Medical facilities and other supporting bodies has proves that the increased recovery rates, which significantly helped to control the second wave.

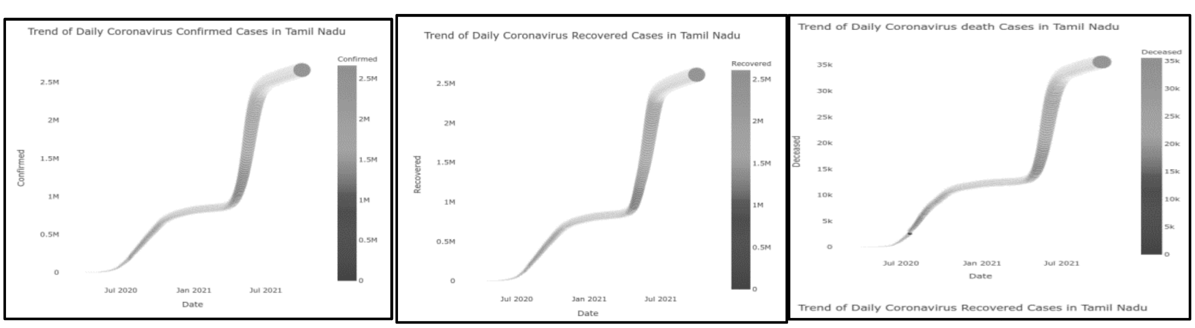


Figure 3: Cumulative growth of the covid-19 Confirmed, Recovered and Death cases till September, 2021

**Figure 3 shows that as of September30, 2021 -

- The cumulative number of confirmed cases is **2663789**
- The cumulative number of recovered cases is **261106**
- The cumulative number of death cases is reduced to **35578**.

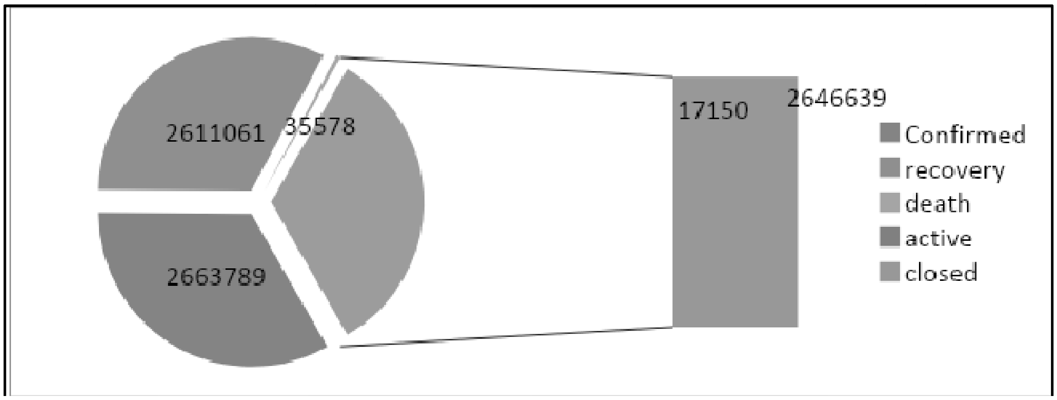


Figure 4: Basic information regarding covid-19 spread in Tamil Nadu - September-30, 2021

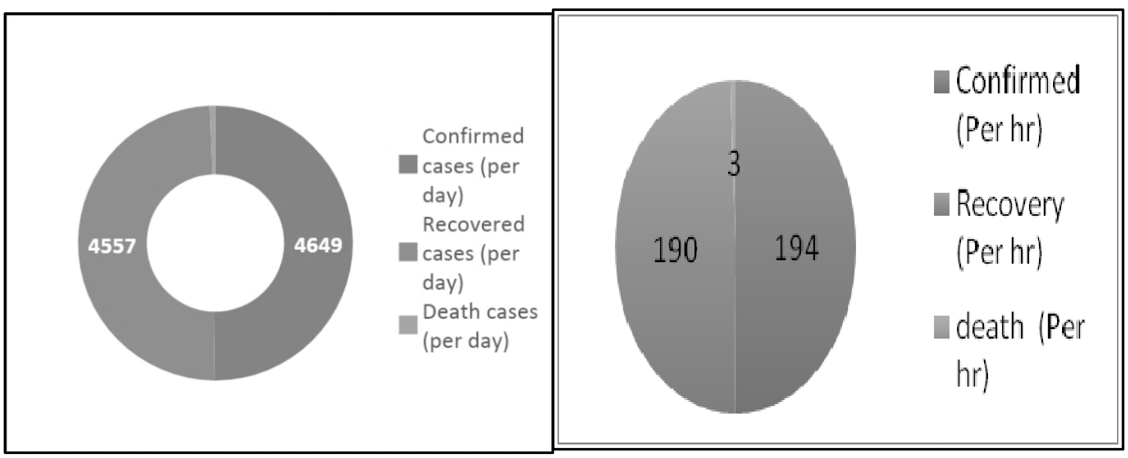


Figure 5: Average number of covid confirmed cases per day and per hour – As of September - 30, 2021

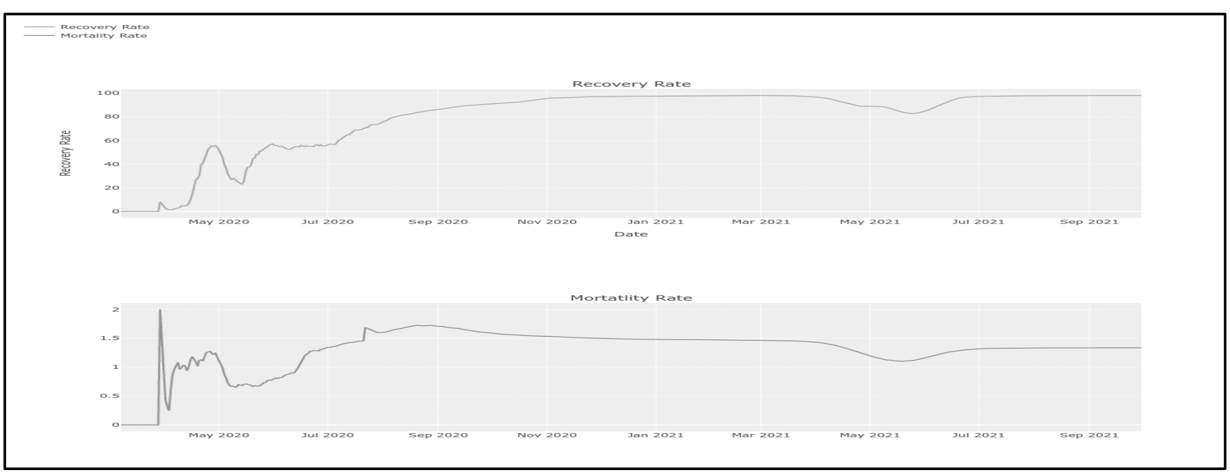


Figure 6: Recovery and Mortality rate between the Month of May 2020 -September - 30, 2021

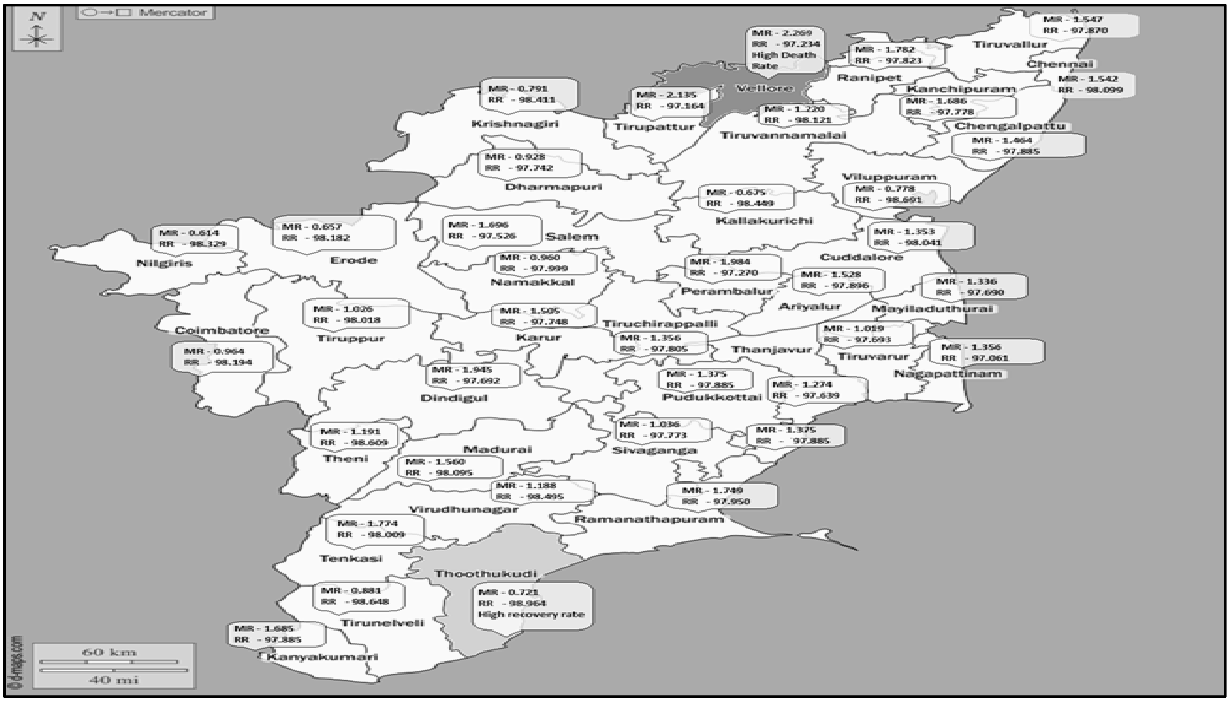


Figure 7: Recovery and Mortality rate for districts of Tamil Nadu –September – 30,21

GENERAL STATISTICS -

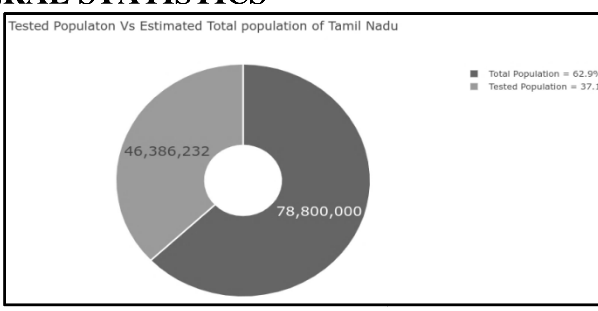


Figure 8: Tested Vs Total Population

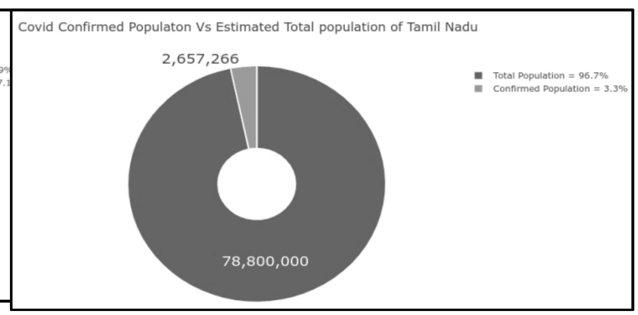


Figure 9: Confirmed Vs Total Population

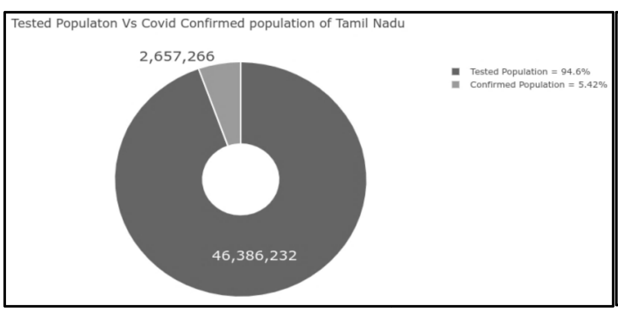


Figure 10: Tested Vs Confirmed Population

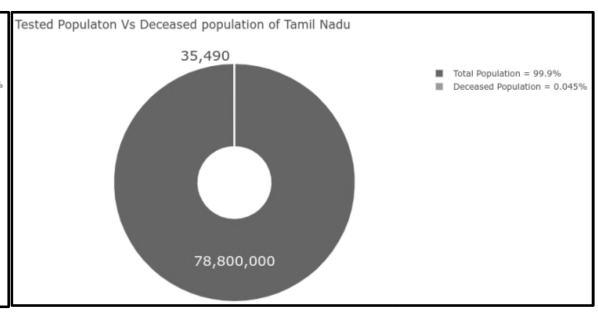
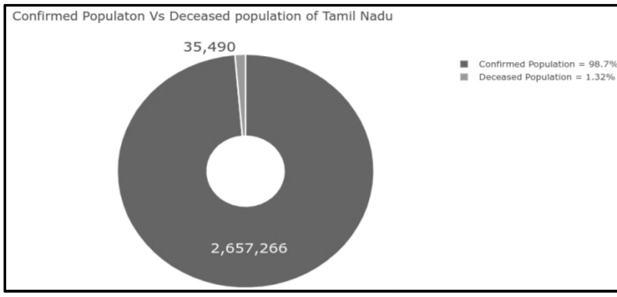


Figure 11: Tested Vs Dead Population



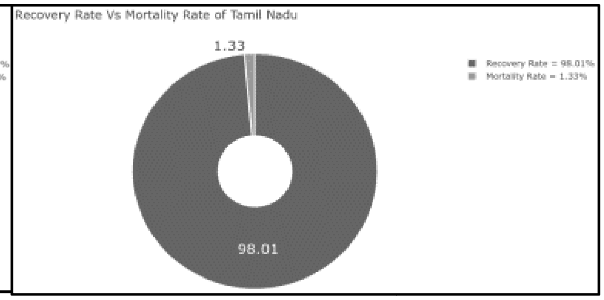


Figure 12: Confirmed Vs Dead Population **Figure 13: Recovery Vs Mortality rate**

Growth factor –

Growth factor is the factor by which a quantity multiplies itself over time.

Formula: Every day's new (Confirmed,Recovered,Death) / new (Confirmed,Recovered,Death) on the previous day.

- A growth factor above 1 indicates an increase in corresponding cases.
- A growth factor above 1 but trending downward is a positive sign, whereas a growth factor constantly above 1 is the sign of exponential growth.

A growth factor constant at 1 indicates there is no change in any kind of cases.

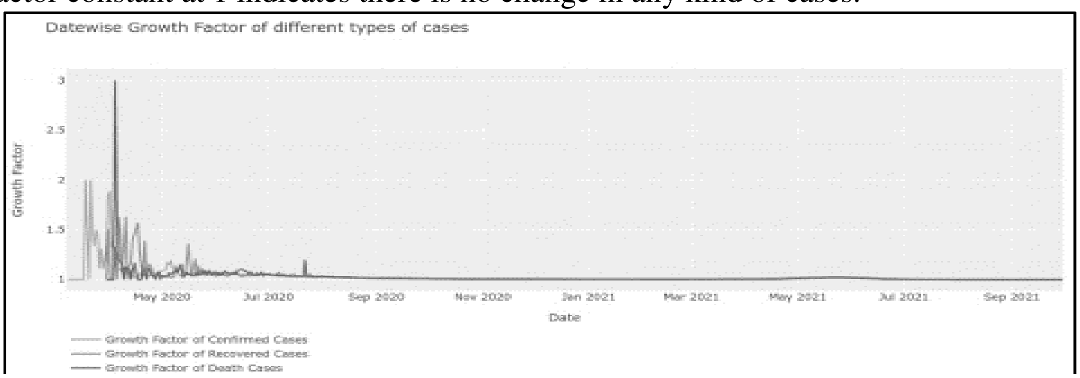


Figure 14: Growth factor of Confirmed/Recovered/Death cases of Tamil Nadu – As of September – 30,2021

The above figure shows that the growth factor for all the cases is constantly maintained at 1 after August 2020, which indicates that there is no significant change in any kind of cases and we can say that the situation is under control.

FORECAST FOR NEXT 31 DAYS -

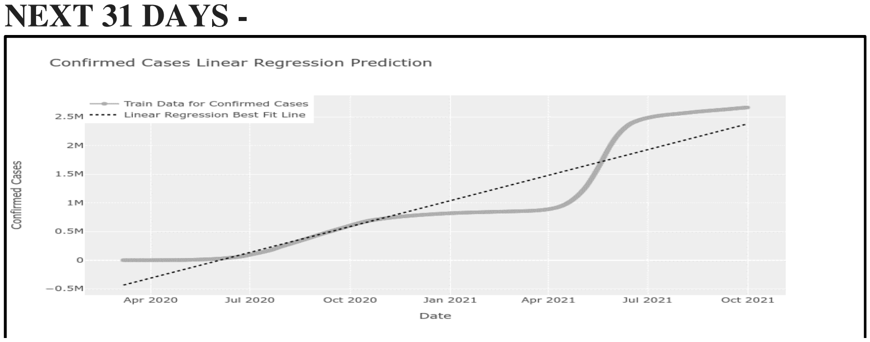


Figure 15: Linear Regression model

The above Figure shows the trend of Confirmed Cases is not Linear in the result.Hence, a linear model cannot be used to forecast the trend. The root means squared error value for linear fitting is obtained as - **334935.4082613871**

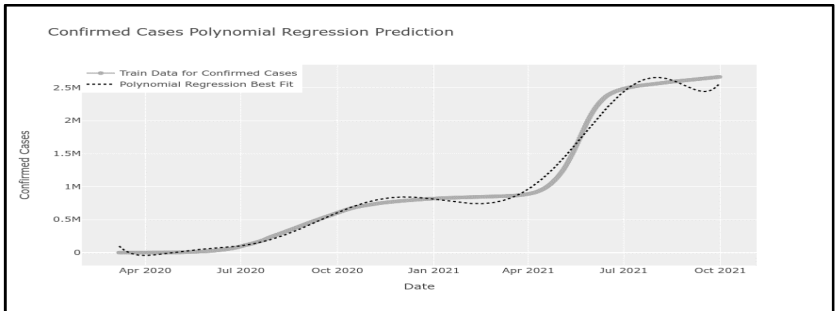


Figure 16: Polynomial Regression model

The above figure shows that the Polynomial regression has performed better than linear regression.The root mean squared error value for the polynomial fit is obtained as **- 169868.6130216721**

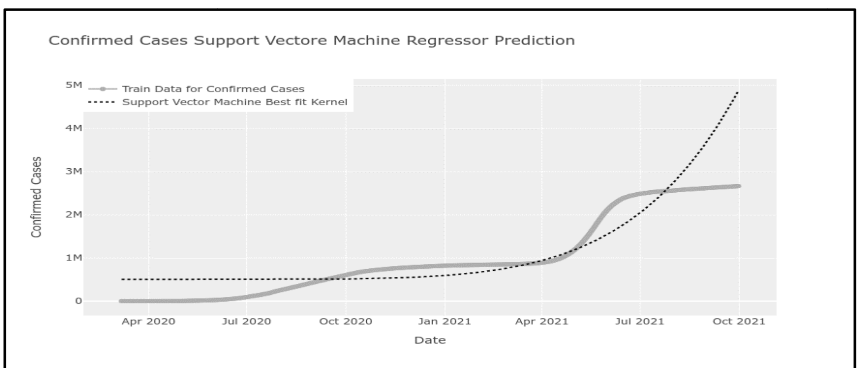


Figure 17: Support Vector Regression model

The above figure shows that the Support Vector Regression model isn't providing great results; the predictions are either overshooting or really lower than what's expected.The root mean squared error value for SVR is obtained as - **1669515.787697374**

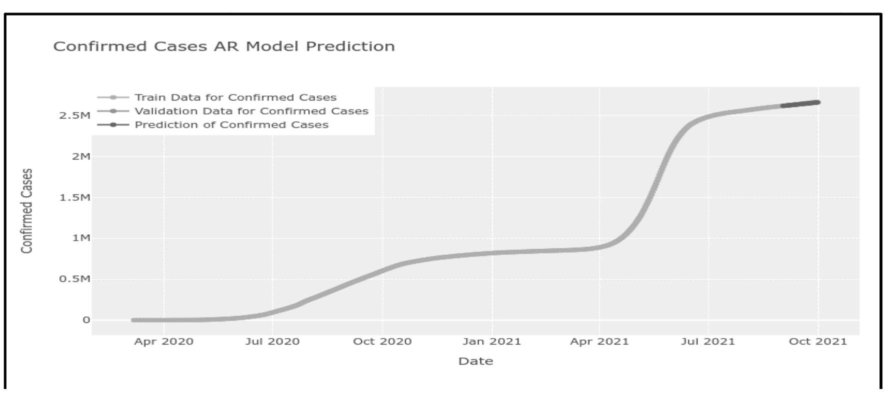


Figure 18: Auto Regressive model

The above figure shows that the Auto Regressive modelgives a good fit. The root mean squared error value for Auto Regressive model is obtained as - **2625.8196975635697**

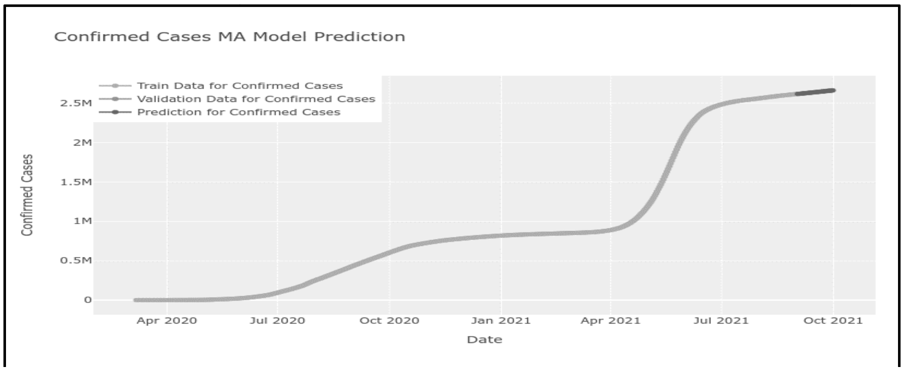


Figure 19: Moving Average model

The above figure shows that the Moving Average model gives a good fit.The root mean squared error value for Moving Average model is obtained as - **1462.6009300985552**.

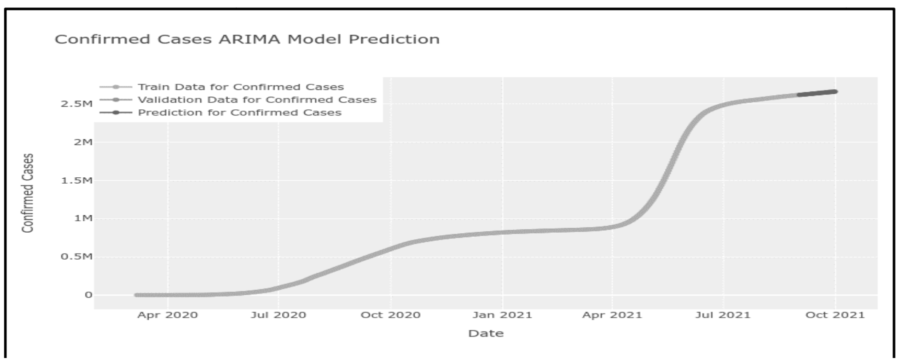


Figure 20: Auto Regressive Integrated Moving Average model (ARIMA)

The above figure shows that the ARIMA model gives a good fit.The root mean squared error value for ARIMA model is obtained as - **3056.455974558541**

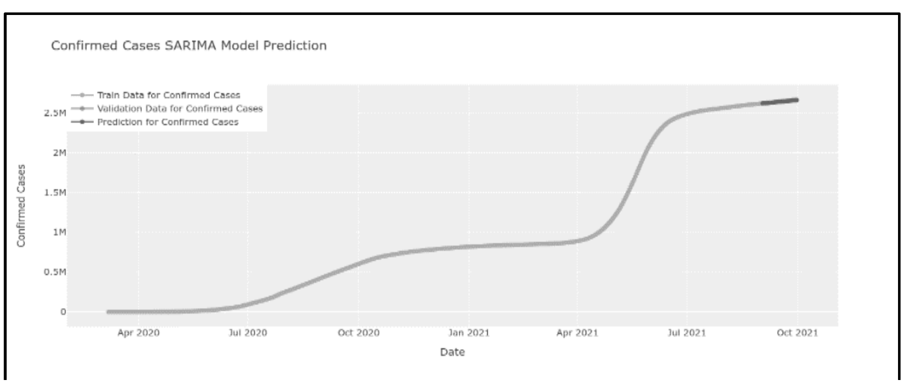


Figure 21: Seasonal Auto Regressive Integrated Moving Average model (SARIMA)

The above figure shows that the SARIMA model gives a good fit.The root mean squared error value for SARIMA model is obtained as - **3142.2053527593243**

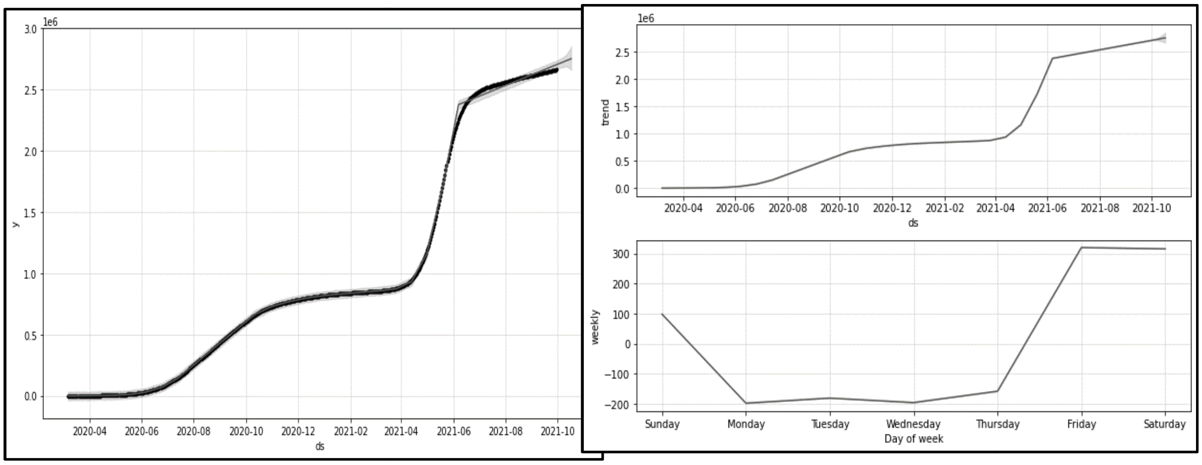


Figure 22: Facebook Prophet model

The above figure shows that the Face book Prophet Model givesan average fit.It also gives us an idea about weekly seasonality and monthly seasonality.The root mean squared error value for Facebook Prophet model is obtained as - **17388.478924680512**

RMSE COMPARISON:

Moving average model	**1462.60**	ARIMA	**3056.45**
Auto regressive model	**2625.82**	SARIMA	**3142.20**
Polynomial regression	**169868.61**	Facebook Prophet	**17388.48**
Support vector regression	**1669515.79**	Linear regression	**334935.41**

Table 2: Model along with their RMSE value

The above table shows that the Moving Average model has the least RMSE and hence it gives best fit for covid data.

FORECAST USING MOVING AVERAGE MODEL

DATE	FORECAST	DATE	FORECAST
2021-10-01	2662802	2021-10-16	2686894
2021-10-02	2664390	2021-10-17	2688521
2021-10-03	2665980	2021-10-18	2690150
2021-10-04	2667573	2021-10-19	2691783
2021-10-05	2669169	2021-10-20	2693418
2021-10-06	2670767	2021-10-21	2695055
2021-10-07	2672368	2021-10-22	2696695
2021-10-08	2673972	2021-10-23	2698338
2021-10-09	2675578	2021-10-24	2699983
2021-10-10	2677186	2021-10-25	2701631
2021-10-11	2678798	2021-10-26	2703282
2021-10-12	2680412	2021-10-27	2704935
2021-10-13	2682028	2021-10-28	2706591
2021-10-14	2683648	2021-10-29	2708249
2021-10-15	2685269	2021-10-30	2709909
-	-	2021-10-31	2711572

Table 3: Forecast for next 31 days

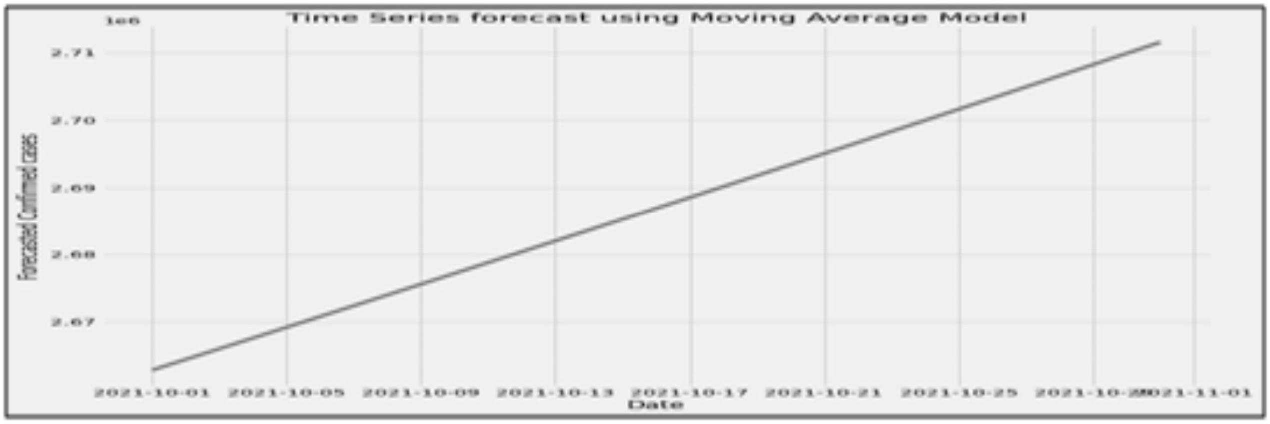


Figure 23: Forecasted trend line for next 31 days

**Forecast for the next 31 days (i.e.) for October month is obtained.There may be other factorslike temperature , climate , location , vaccination drugs , immunity of the people ,transportation , lockdown measures and so on , which affects the future cases.Those factorsare not considered here, hence there may be some deviation in the prediction.This model can be optimised in the future by considering other external factors, in order to get better results.

CONCLUSION

- The confirmed cases have reached a peak in the month of May to June 2021, this was thetime period when the second wave of covid-19 hit Tamil Nadu.This also resulted in thesimultaneous increase in recovered death cases.
- As of september-30, 2021 -

  The cumulative number of confirmed cases till date is **2663789**, that is 2.6million; The cumulative number of recovered cases till date is **2611061**, that is 2.6million; The cumulative number of death cases till date is **35578**.
- As of september-30, 2021, in Tamil Nadu --
- Total number of Confirmed Cases: **2663789**
- Total number of Recovered Cases: **2611061**
- Total number of Deaths Cases: **35578**
- Total number of Active Cases: **17150**
- Total number of Closed Cases: **2646639**
- Approximate number of Confirmed Cases per Day: **4649.0**

- Approximate number of Recovered Cases per Day: **4557.0**
- Approximate number of Death Cases per Day: **62.0**
- Approximate number of Confirmed Cases per hour: **194.0**
- Approximate number of Recovered Cases per hour: **190.0**
- Approximate number of Death Cases per hour: **3.0**
- Recovery rate is very high for Thoothukudidistrict(**98.96**), and the mortalityrate(**0.72**) isalso low , which is a good sign of control.Whereas, Mortality rate is very high forVellore district(**2.269**), which is a bad sign.Nilgiris has a very low mortality ratecomparatively(**0.61**), which is a good sign.
- As of September-30, 2021- 37.1% of Tamilnadu's total population have been tested; 3.3% of Tamilnadu's total population have been tested positive forcovid-19; Out of 100 people tested, only 5.42% of people got confirmed for thedisease; Only 0.045% of people died due to covid-19 out of total populationof Tamil Nadu;Only1.32% of people got confirmed of covid-19 out of total population of Tamil Nadu; The recovery rate of Tamil Nadu covid cases stands at 98.01,whereas mortality rate is 1.33, which is very low compared to the recovery rate.This is apositive sign, that more people are recovering from the disease.
- The growth factor for all confirmed, recovered,death cases are constantly maintained at 1after August 2020, which indicates there is no significant change in any kind of casesand we can say that the situation is under control.
- Recovery Rate of Tamil Nadu has started to pick up again after a slight drop in June, which is a good sign, a reason to support why the number of Closed Cases isincreasing.Mortality rate of Tamil Nadu is showing a steady trend for a pretty long time, which is a positive sign.
- The RMSE Values of different models are obtained as -
  Moving average model - **1462.6009300985552;** AR model - **2625.8196975635697;** ARIMA - **3056.455974558541;** SARIMA - **3142.2053527593243;**Facebook prophet - **17388.478924680512;** Polynomial regression - **169868.6130216721;** Linear regression - **334935.4082613871;** Support vector regression - **1669515.787697374**
- The Moving Average model has the least RMSE and hence it gives best fit for coviddata.We can also see that regression models are not suitable for covid prediction.
- Using this moving average predictive model, a forecasted trend line is obtained for thenext 31 days.

- It is expected from the model that the covid-19 confirmed cases would reach **2711572**cases at the end of October in Tamil Nadu.

REFRENCES:

1. *ARIMA modelling & forecasting of COVID-19 in top five affected countries by Alok Kumar Sahai,NamitaRath 2,Vishal Sood,ManvendraPratap Singh*
2. *Forecasting the novel coronavirus covid-19. PLoS One. 2020;15 by Petropoulos F.MakridakisS*
3. *M. Mofijur, I.M.R. Fattah, M.A. Alam, et al.-Impact of COVID-19 on the social, economic, environmental and energy domains: lessons learnt from a global pandemic*
4. *J. Zheng-SARS-CoV-2: an emerging coronavirus that causes a global threat*
5. *Time series modelling to forecast the confirmed and recovered cases of COVID-19 by Mohsen Maleki,a Mohammad Reza Mahmoudi,b,c,DarrenWraith,d,Kim-Hung Phoe*
6. *T. Alamo, D.G. Reina, M. Mammarella, A. Abella-Covid-19: open-data resources for monitoring, modeling, and forecasting the epidemic*
7. *Forecasting Peaks of Seasonal Influenza Epidemics by Elaine Nsoesie,MadhavMararthe,John Brownstein, 2013 June 21*
8. *(Data-based analysis, modeling and forecasting of the COVID-19 outbreak by Cleo Anastassopoulou,LuciaRusso,AthanasiosTsakris,ConstantinosSiettos*
9. *Comparative study of a mathematical epidemic model, statistical modelling, and deep learning for COVID-19 forecasting and management by Mohammad Masum, M A Masud, HussainShahriar, and Sangil Kim*
10. *Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020.*
11. *Deep learning methods for forecasting covid-19 time-series data: a comparative study-Solit Fractals, Chaos. 2020;140:110121., by Zeroual A., Harrou F., DairiA,Sun Y.*
12. *Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-2019).*
13. *WHO. Coronavirus disease 2019 (COVID-19) situation reports.*
14. *Hyndman RJ, Koehler AB, Ord JK, Snyder RD. Forecasting with Exponential Smoothing: The State Space Approach. Berlin Germany: Springer, 2008. 372 p.*

15. *De Livera A, Hyndman R, Snyder R. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. Journal of the American Statistical Association. 2011; 106(496), 1513–1527.*
16. *Springer handbook of computational statistics - handbook of data visualization*
17. *COVID-19 prediction and detection using deep learningInt. J. Comp. Info. Syst. Indust. Manag. Appl., 12 (2020), pp. 168-181*
18. *Comparing the accuracy of several network-based COVID-19 prediction algorithms Int. J. Forecast. (2020)*

DOI No. : 10.5958/2394-9309.2022.00037.3
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

GOODMAN-RONNING TYPE HARMONIC UNIVALENT FUNCTIONS DEFINED BY q- DIFFERENTIAL OPERATORS

Asha Thomas*, Thomas Rosy, Sunil Varma*****
Department of Mathematics, Madras Christian College, Tambaram, Chennai, Tamilnadu, India-600059.
E-mail: ashathomas@mcc.edu.in*, rosy@mcc.edu.in** , sunilvarma@mcc.edu.in***

ABSTRACT

The main objective of this article is to define and investigate a family of complex valued harmonic univalent functions related to Goodman-Ronning type uniformly convex functions using the q-differential operators. We consider a subclass of the aforementioned class and obtain necessary and sufficient conditions, distortion bounds, extreme points, convolution conditions and convex combination for the family of harmonic functions.

Keywords: Harmonic, Univalent, Uniformly convex, q-differential operator

1. INTRODUCTION

The function $f=u+iv$, continuous and complex-valued, defined in the simply connected complex domain Ψ is harmonic if both u and v are real harmonic in Ψ. In any simply connected domain $\mathfrak{D}\subset \Psi$, we can write $f = h + \bar{g}$ where h and g are the analytic and the co-analytic part of f respectively. A necessary and sufficient condition for f to be locally univalent and orientation preserving in $\mathfrak{D}$ is that $|h'(\xi)| > |g'(\xi)|$ in $\mathfrak{D}$ (see [1]). Let $\mathcal{H}$ denote the family of functions $f = h + \bar{g}$ which are harmonic, univalentand orientation preserving in the open unit disc $\mathbb{U} = \{\xi : |\xi| < 1\}$ and are of the form $f(\xi) = h(\xi) + \overline{g(\xi)} = \xi + \sum_{n=2}^{\infty} a_n\xi^n + \sum_{n=1}^{\infty} \overline{b_n\xi^n}$. (1.1)

Obviously $|b_1| < 1$ and the family $\mathcal{H}$ reduces to the known class S of normalized analytic univalent functions if the co-analytic part of f is identically zero, that is $g \equiv 0$.

The Hadamard product or convolution of two power series $h_1(\xi) = \sum_{n=1}^{\infty} a_n\xi^n$ and $h_2(\xi) = \sum_{n=1}^{\infty} c_n\xi^n$ is given by $h_1(\xi) * h_2(\xi) = (h_1 * h_2)(\xi) = \sum_{n=1}^{\infty} a_n c_n\xi^n$

and the convolution of two harmonic functions $f_1(\xi) = h_1(\xi) + \overline{g_1(\xi)}$ and $f_2(\xi) = h_2(\xi) + \overline{g_2(\xi)}$ is given by $f_1(\xi) \tilde{*} f_2(\xi) = (f_1 \tilde{*} f_2)(\xi) = h_1(\xi) * h_2(\xi) + \overline{g_1(\xi) * g_2(\xi)}$

We now recall the *q-operator or q-difference operator* that plays an important role in the study of hypergeometric series, quantum physics, and operator theory. In 1908 Jackson [3] initiated the application of *q-calculus* to analytic functions. For 0 <q< 1, Jackson's *q-derivative* of the function $h(\xi) = \xi + \sum_{n=2}^{\infty} a_n \xi^n \in S$ is given by

$$D_q h(\xi) = \begin{cases} \dfrac{h(\xi) - h(q\xi)}{(1-q)\xi} & , \quad \xi \neq 0 \\ h'(0) & , \quad \xi = 0 \end{cases}$$

where $D_q h(\xi) = 1 + \sum_{n=2}^{\infty} [n]_q a_n \xi^{n-1}$ and $[n]_q = \frac{1-q^n}{1-q}$.

The Ruscheweyh type *q-differential operator,* was introduced and investigated by Kannas and Raducanu [7].$R_q^m h(\xi) = h(\xi) * F_{q,m+1}(\xi) = \xi + \sum_{n=2}^{\infty} \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)} a_n \xi^n \;\; , m > -1,$

where$F_{q,m+1}(\xi) = \xi + \sum_{n=2}^{\infty} \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)} \xi^n = \xi + \sum_{n=2}^{\infty} \frac{[m+1]_{n-1}}{(n-1)!} \xi^n$.

We note that $$R_q^0 h(\xi) = h(\xi),$$

$$R_q' h(\xi) = \xi D_q h(\xi),$$

$$\vdots$$

$$R_q^m h(\xi) = \frac{\xi D_q^m\left(\xi^{m-1} h(\xi)\right)}{(m)!}$$

where $D_q^2 h(\xi) = D_q\left(D_q\big(h(\xi)\big)\right)$and $D_q^m h(\xi) = D_q^{m-1}\left(D_q h(\xi)\right)$.

It is obvious that$R_q^m h(\xi) = \xi + \sum_{n=2}^{\infty} \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)} a_n \xi^n$,

$\lim_{q\to 1^-} R_q^m h(\xi) = R^m h(\xi) = h(\xi) * \frac{\xi}{(1-\xi)^{m+1}}$ and $\lim_{q\to 1^-} F_{q,m+1}(\xi) = F_{m+1}(\xi) = \frac{\xi}{(1-\xi)^{m+1}}$.

We remark that if $q \to 1^-$, then the Ruscheweyh q-differential operator reduces to the differential operator defined by Ruscheweyh [8],$D_q R_q^m h(\xi) = 1 + \sum_{n=2}^{\infty} [n]_q \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)} a_n \xi^{n-1}$.

Recently, Jahangiri [5] applied *q-difference operators* to classes of harmonic functions and obtained coefficient bounds for such functions.

Goodman [2] defined uniformly convex (UCV) functions so that its analytic characterization is $\varphi \epsilon UCV$if and only ifRe$\left\{1 + \frac{z\varphi''(z)}{\varphi'(z)}\right\} \geq \text{Re}\,\frac{\zeta\varphi''(z)}{\varphi'(z)}$, $\qquad z, \zeta \in \mathbb{U} \times \mathbb{U}$.

Ronning [8] defined the class $\mathcal{S}_p$consisting of functions $\psi(\xi) = \xi\varphi'(\xi)$ so that $\varphi \in UCV$. Upon choosing $\zeta = e^{i\alpha}\xi$ in a suitable way, the analytic characterization of $\mathcal{S}_p$ may be written as $\psi \in UCV$ if and only ifRe$\left\{\frac{z\psi'(z)}{\psi(z)} + e^{i\alpha}\left(\frac{z\psi'(z)}{\psi(z)} - 1\right)\right\} \geq 0$.

Generalizing the class $\mathcal{S}_p$to include harmonic functions, we let $G_{\mathcal{H}}(\vartheta)$ denote the subclass of $\mathcal{H}$consisting of functions$f = h + \bar{g} \in \mathcal{H}$that satisfy the condition

$$\text{Re}\left\{\left(1 + e^{i\alpha}\right)\frac{zf'(z)}{z'f(z)} - e^{i\alpha}\right\} \geq \vartheta, \qquad 0 \leq \vartheta < 1.$$

Motivated by [2,5,7,8], we define a class of Ruscheweyh-type q-calculus harmonic functions $\mathcal{H}_q^m(\vartheta)$ consisting of functions $f \in \mathcal{H}$ satisfying

$$\operatorname{Re}\left\{\left(1+e^{i\alpha}\right)\frac{\xi D_q R_q^m f(\xi)}{\xi' R_q^m f(\xi)} - e^{i\alpha}\right\} \geq \vartheta \tag{1.2}$$

where $\xi \in \mathbb{U}, 0 \leq \lambda \leq 1, \xi' = \frac{\partial}{\partial\theta}(\xi), \xi = re^{i\theta}$ and

$$\xi D_q\left(R_q^m f(\xi)\right) = \xi D_q R_q^m h(\xi) - \overline{\xi D_q\left(R_q^m g(\xi)\right)}.$$

We also define $\overline{\mathcal{H}}_q^m(\gamma) \equiv \mathcal{H}_q^m(\gamma) \cap \overline{\mathcal{H}}$ where $\overline{\mathcal{H}}$ is the subfamily of $\mathcal{H}$ consisting of harmonic functions of the form $f(\xi) = \xi - \sum_{n=2}^{\infty} a_n \xi^n + \sum_{n=1}^{\infty} b_n \bar{\xi}^n$, $a_n \geq 0, b_n \geq 0.$ (1.3)

Jahangiri [4], Silverman [10] and Silverman and Silvia [11] studied the harmonic starlike functions.

Jahangiri [4] proved that if $f = h + \bar{g}$ is given by (1.1) and if

$$\sum_{n=2}^{\infty}\left(\frac{n-\vartheta}{1-\vartheta}\right)|a_n| + \sum_{n=1}^{\infty}\left(\frac{n+\vartheta}{1-v}\right)|b_n| \leq 1,\ 0 \leq \vartheta < 1$$

then f is harmonic, univalent and starlike of order ϑ in $\mathbb{U}$. This condition is provedto be also necessary if h and g are of the form (1.3). The case when $\vartheta = 0$ is givenin [11] and for $\vartheta = b_1 = 0$ see [10].

The objective of this paperis to obtain the sufficient coefficient conditions for the harmonic function $f = h + \bar{g}$ to be in the class $\mathcal{H}_q^m(\vartheta)$. We also determine necessary and sufficient conditions for harmonic functions $f = h + \bar{g}$ to be in the class $\overline{\mathcal{H}_q^m}(\vartheta)$. Furthermore, distortion bounds, convolution conditions, convex combinations and extreme points for functions in $\overline{\mathcal{H}_q^m}(\vartheta)$ are also obtained.

2. MAIN RESULTS

Theorem 2.1

Let $f = h + \bar{g} \in \mathcal{H}$. If

$$\sum_{n=1}^{\infty}\left[\frac{2[n]_q - 1 - \vartheta}{1-\vartheta}|a_n| + \frac{2[n]_q + 1 + \vartheta}{1-\vartheta}|b_n|\right]\phi_q(n,m) \leq 2 \tag{2.1}$$

where $a_1 = 1$, $0 \leq \vartheta < 1$, $\phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}$ then $f \in \mathcal{H}_q^m(\vartheta)$.

Proof.

We will show, if the coefficients of the harmonic function $f = h + \bar{g} \in \mathcal{H}$ satisfies (2.1) then

$f = h + \bar{g}$ satisfies (1.2). Equivalently, we need to show that

$$\operatorname{Re}\left\{\frac{\left(1+e^{i\alpha}\right)\left(\xi DqR_q^m h(\xi) - \overline{\xi D_q R_q^m g(\xi)}\right) - e^{i\alpha}\left(R_q^m h(\xi) + \overline{R_q^m g(\xi)}\right)}{R_q^m h(\xi) + \overline{R_q^n g(\xi)}}\right\} = \operatorname{Re}\left\{\frac{A(\xi)}{B(\xi)}\right\} \geq \vartheta \tag{2.2}$$

where $\xi = re^{i\theta}$, $A(\xi) = \left(1+e^{i\alpha}\right)\left(\xi DqR_q^m h(\xi) - \overline{\xi D_q R_q^m g(\xi)}\right) - e^{i\alpha}\left(R_q^m h(\xi) + \overline{R_q^m g(\xi)}\right)$ and

$B(\xi) = R_q^m h(\xi) + \overline{R_q^n g(\xi)}$

Using the fact that $\operatorname{Re}\{\omega\} \geq \vartheta$ if and only if $|1-\vartheta+\omega| \geq |1+\vartheta-\omega|$,

it suffices to show that $|A(\xi)+(1-\vartheta)B(\xi)| - |A(\xi)-(1+\vartheta)B(\xi)| \geq 0$ (2.3)

Substituting for $A(\xi)$ and $B(\xi)$ in (2.3) we obtain

$$\left|(1-\vartheta-e^{i\alpha})R_q^m h(\xi) + (1+e^{i\alpha})\xi D_q R_q^m h(\xi) + (1-\vartheta-e^{i\alpha})\overline{R_q^m g(\xi)} + (1+e^{i\alpha})\overline{\xi Dq R_q^m g(\xi)}\right| -$$

$$\left|(1+\vartheta+e^{i\alpha})R_q^m h(\xi) - (1+e^{i\alpha})\xi Dq R_q^m h(\xi) + (1+\vartheta+e^{i\alpha})\overline{R_q^m g(\xi)} + (1+e^{i\alpha})\overline{\xi Dq R_q^m g(\xi)}\right|$$

$$=\Big|(2-\vartheta)\xi + \textstyle\sum_{n=2}^{\infty}\big([n]_q+1-\vartheta\big)+e^{i\alpha}\big([n]_q-1)\big)\phi_q(n,m)a_n\xi^n - \sum_{n=1}^{\infty}([n]_q-(1-\vartheta)) +$$

$$e^{i\alpha}\big([n]_q+1)\big)\phi_q(n,m)\overline{b_n}\,\bar{\xi}^n\Big| - \Big|-\vartheta\xi + \textstyle\sum_{n=2}^{\infty}\Big([n]_q-(1+\vartheta)\Big)+e^{i\alpha}\big([n]_q-1)\big)\phi_q(n,m)a_n\xi^n -$$

$$\textstyle\sum_{n=1}^{\infty}([n]_q+1+\vartheta)+e^{i\alpha}\big([n]_q+1)\big)\phi_q(n,m)\overline{b_n}\,\bar{\xi}^n\Big|$$

$$\geq (2-\vartheta)|\xi| - \sum_{n=2}^{\infty}\big(2[n]_q-\vartheta\big)\phi_q(n,m)|a_n||\xi^n| + \sum_{n=1}^{\infty}\big(2[n]_q+\vartheta\big)\phi_q(n,m)\left|\overline{b_n}\right|\left|\bar{\xi}^n\right|$$

$$-\vartheta\xi - \sum_{n=2}^{\infty}\big(2[n]_q-2-\vartheta\big)\phi_q(n,m)|a_n||\xi^n| \quad - \sum_{n=1}^{\infty}\big(2[n]_q+2+\vartheta\big)\phi_q(n,m)\overline{|b_n|}\,\overline{|\xi|}^n$$

$$=2(1-\vartheta)|\xi|\left(1-\left[\textstyle\sum_{n=2}^{\infty}\frac{2[n]_q-1-\vartheta}{1-\vartheta}|a_n| - \sum_{n=1}^{\infty}\frac{2[n]_q+1+\vartheta}{1-\vartheta}|b_n|\right]\phi_q(n,m)\xi^{n-1}\right)$$

$$\geq\ 2(1-\vartheta)\left(1-\left[\textstyle\sum_{n=2}^{\infty}\frac{2[n]_q-1-\vartheta}{1-\vartheta}|a_n| - \sum_{n=1}^{\infty}\frac{2[n]_q+1+\vartheta}{1-\vartheta}|b_n|\right]\phi_q(n,m)\right) \geq 0$$ by (2.1) and

so $f \in \mathcal{H}_q^m(\vartheta)$. ■

The harmonic function

$$f(\xi) = \xi + \sum\nolimits_{n=2}^{\infty}\frac{1-\vartheta}{(2[n]_q-1-\vartheta)\phi_q(n,m)}x_n\xi^n + \sum\nolimits_{n=1}^{\infty}\frac{1-\vartheta}{(2[n]_q+1+\vartheta)\phi_q(n,m)}y_n\bar{\xi}^n \qquad (2.4)$$

where $\sum_{n=2}^{\infty}|x_n| + \sum_{n=1}^{\infty}|y_n| = 1$ shows that the coefficient bound given by (2.1) is sharp.

The function of the form (2.4)are in $\mathcal{H}_q^m(\vartheta)$ because

$$\sum_{n=1}^{\infty}\left[\frac{2[n]_q-1-\vartheta}{1-\vartheta}a_n + \frac{2[n]_q+1+\vartheta}{1-\vartheta}b_n\right]\phi_q(n,m)$$

$=1 + \sum_{n=2}^{\infty}|x_n| + \sum_{n=1}^{\infty}|y_n|$=1+1=2.

Theorem 2.2

For $a_1 = 1$, $\quad 0 \leq \vartheta < 1, \phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}$, $f = h + \bar{g} \in \bar{\mathcal{H}}_q^m(\vartheta)$ and $h(\xi) = \xi - \sum_{n=2}^{\infty}|a_n|\xi^n$

if and only if $\sum\nolimits_{n=1}^{\infty}\left[\frac{2[n]_q-1-\vartheta}{1-\vartheta}a_n + \frac{2[n]_q+1+\vartheta}{1-\vartheta}b_n\right]\phi_q(n,m) \leq 2.$ (2.5)

Proof.

Since $\overline{H_q^m}(\vartheta) \subset H_q^m(\vartheta)$ the if part is from Theorem 2.1.

For the 'only if 'part we show that $f \notin \bar{\mathcal{H}}_q^m(\vartheta)$ if the condition (2.5) does not hold.

The necessary and sufficient condition for $f = h + \bar{g} \in \bar{\mathcal{H}}_q^m(\vartheta)$
is that $\operatorname{Re}\left[\left(1+e^{i\alpha}\right)\frac{\xi D_q R_q^m f(\xi)}{\xi' R_q^m f(\xi)} - e^{i\alpha}\right] \geq \vartheta$
which is equivalent to $\operatorname{Re}\left\{\frac{(1+e^{i\alpha})(\xi DqR_q^m h(\xi) - \overline{\xi D_q R_q^m g(\xi)}) - e^{i\alpha}(R_q^m h(\xi) + \overline{R_q^m g(\xi)})}{R_q^m h(\xi) + R_q^n g(\xi)} - \vartheta\right\} =$

$$\operatorname{Re}\left\{\frac{\xi - \sum_{n=2}^{\infty}[n]_q\left(1+e^{i\alpha} \ \ -e^{i\alpha}\right)\phi_q(n,m)\left|a_n\right|\xi^n - \sum_{n=1}^{\infty}[n]_q(1+e^{i\alpha}+e^{i\alpha})\phi_q(n,m)\left|b_n\right|\bar{\xi}^n}{\xi - \sum_{n=2}^{\infty}|a_n|\xi^n + \sum_{n=2}^{\infty}|b_n|\bar{\xi}^n} - \vartheta\right\} =$$

$$Re\left\{\frac{1-\vartheta-\sum_{n=2}^{\infty}([n]_q - \vartheta + ([n]_q - 1)e^{i^{\alpha}})\phi_q(n,m)\left|a_n\right|\xi^{n-1} - \frac{\bar{\xi}}{\xi}\sum_{n=1}^{\infty}([n]_q + \vartheta + ([n]_q + 1)e^{i^{\alpha}})\phi_q(n,m)\left|b_n\right|\bar{\xi}^{n-1}}{1-\sum_{n=2}^{\infty}|a_n|\xi^{n-1} + \frac{\bar{\xi}}{\xi}\sum_{n=2}^{\infty}|b_n|\bar{\xi}^{n-1}}\right\}$$

$$\geq 0$$

The above required condition must hold for all ξ.For $|\xi| = r < 1$

$$\frac{1-\vartheta-\sum_{n=2}^{\infty}([n]_q-\vartheta)\phi_q(n,m)|a_n|r^{n-1} - \sum_{n=1}^{\infty}([n]_q+\vartheta)\phi_q(n,m)|b_n|r^{n-1} - e^{i^{\alpha}}\sum_{n=2}^{\infty}([n]_q-\vartheta)\phi_q(n,m)|a_n|r^{n-1} - \sum_{n=1}^{\infty}([n]_q+\vartheta)\phi_q(n,m)|b_n|r^{n-1}}{1-\sum_{n=2}^{\infty}|a_n|\xi^{n-1} + \frac{\bar{\xi}}{\xi}\sum_{n=2}^{\infty}|b_n|\bar{\xi}^{n-1}}$$

$$\geq 0 \tag{2.6}$$

If the condition (2.5) does not hold, then the numerator in (2.6) is negative for *r* sufficiently close to 1.

Hence, there exist $\xi_0 = r_0$ in (0, 1) for which the quotient in (2.6) is negative. This contradicts the condition for $f \in \bar{\mathcal{H}}_q^m(\vartheta)$and hence the result. ■

DISTORTION BOUNDS

Theorem 2.3

If $f \in \bar{\mathcal{H}}_q^m(\vartheta),\ \phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}$ then

$|f(\xi)| \leq (1+b_1)r + \frac{1}{\phi_q(2,m)}\left(\frac{1-\vartheta}{2[2]_q-1-\vartheta} - \frac{3+\vartheta}{2[2]_q-1-\vartheta} \quad b_1\right)r^2$ and

$|f(\xi)| \geq (1-b_1)r - \frac{1}{\phi_q(2,m)}\left(\frac{1-\vartheta}{2[2]_q-1-\vartheta} - \frac{3+\vartheta}{2[2]_q-1-\vartheta} \quad b_1\right)r^2$ for $|\xi| = r < 1$.

Proof.

Let $f \in \bar{\mathcal{H}}_q^m(\vartheta)$. The proof of left-hand inequality is similar to that of the right-hand inequality, so we prove the right-hand inequality by taking the absolute value,

$$|f(\xi)| = \left|\xi + \sum_{n=2}^{\infty} a_n \xi^n + \sum_{n=1}^{\infty} b_n \bar{\xi}^n\right|$$

$$\leq (1+b_1)|\xi| + \sum_{n=2}^{\infty}(a_n + b_n)\,|\xi|^n$$

$$\leq (1+b_1)r + \sum_{n=2}^{\infty}(a_n + b_n) r^2$$

$$\leq (1+b_1)r + \frac{1-\vartheta}{(2[2]_q-1-\vartheta)\phi_q(2,m)} \sum_{n=2}^{\infty}\left(\frac{2[2]_q-1-\vartheta}{1-\vartheta}\phi_q(2,m)a_n + \frac{2[2]_q+1+\vartheta}{1-\vartheta}\phi_q(2,m)b_n\right) r^2.$$

$$\leq (1+b_1)r + \frac{1-\vartheta}{2[2]_q-1-\vartheta}\left(1-\frac{3+\vartheta}{1-\vartheta}b_1\right) r^2$$

$$\leq (1+b_1)r + \frac{1}{\phi_q(2,m)}\left(\frac{1-\vartheta}{2[2]_q-1-\vartheta} - \frac{3+\vartheta}{2[2]_q-1-\vartheta} \quad b_1\right) r^2. \qquad \blacksquare$$

Theorem 2.5

$f \in clco\bar{\mathcal{H}}_q^m(\vartheta)$ if and only if $f(\xi) = \sum(X_n h_n + Y_n g_n)$ where$h_1(\xi) = \xi$, $h_n(\xi) = \xi - \frac{1-\vartheta}{(2[n]_q-1-\vartheta)\phi_q(n,m)}\xi^n$ and $g_n(\xi)$=$\frac{1-\vartheta}{(2[n]_q+1+\vartheta)\phi_q(n,m)}\bar{\xi}^n$ for $n \geq 2$, $\phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}, \sum_{n=1}^{\infty}(X_n + Y_n) = 1, X_n \geq 0$, $Y_n \geq 0$. In particular, the extreme points of $\bar{\mathcal{H}}_q^m(\vartheta)$ are $\{h_n\}$ and$\{g_n\}$.

Proof.

$$f(\xi) = \sum(X_n h_n + Y_n g_n)$$
$$= \sum_{n=1}^{\infty}\big(X_n h_n(\xi) + Y_n g_n(\xi)\big)$$
$$= \sum_{1}^{\infty}(X_n + Y_n)\xi - \sum_{n=2}^{\infty}\frac{1-\vartheta}{\big(2[n]_q-1-\vartheta\big)\phi_q(n,m)}X_n\xi^n + \sum_{n=1}^{\infty}\frac{1-\vartheta}{\big(2[n]_q+1+\vartheta\big)\phi_q(n,m)}Y_n\bar{\xi}^n$$

then

$$\sum_{n=2}^{\infty}\frac{\big(2[n]_q-1-\vartheta\big)\phi_q(n,m)}{1-\vartheta}|a_n| + \sum_{n=1}^{\infty}\frac{\big(2[n]_q+1+\vartheta\big)\phi_q(n,m)}{1-\vartheta}|b_n|$$
$$= \sum_{2}^{\infty}X_n + \sum_{1}^{\infty}Y_n$$

$= 1 - X_1 \leq 1$ so $f \in clco\bar{\mathcal{H}}_q^m(\vartheta)$

Conversely, if $f \in clco\bar{\mathcal{H}}_q^m(\vartheta)$ set

$$X_n = \frac{\big(2[n]_q-1-\vartheta\big)\phi_q(n,m)}{1-\vartheta}|a_n| \qquad 0 \leq X_n \leq 1, n \geq 2$$

$Y_n = \frac{(2[n]_q+1+\vartheta)\phi_q(n,m)}{1-\vartheta}|b_n| 0 \le Y_n \le 1, n \ge 2$ and $X_1 = 1 - \sum_{n=2}^{\infty} X_n - \sum_{n=1}^{\infty} Y_n$ and note that by Theorem 2.2 $X_1 \ge 0$. Consequently, we get $f(\xi) = \sum_1^{\infty}(X_n h_n(\xi) + Y_n g_n(\xi))$

Using Theorem 2.2 we can observe that $\bar{\mathcal{H}}_q^m(\vartheta)$ is convex and closed so $clco\bar{\mathcal{H}}_q^m(\vartheta) = \bar{\mathcal{H}}_q^m(\vartheta)$. ∎

Theorem 2.6

The family $\bar{\mathcal{H}}_q^m(\vartheta)$ is closed under convex combination, where $\phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}$.

Proof.

Suppose $f_i \in \bar{\mathcal{H}}_q^m(\vartheta)$

$f_i(\xi) = \xi - \sum_{n=2}^{\infty} a_{i,n}\xi^n + \sum_{n=2}^{\infty} \bar{b}_{i,n}\bar{\xi}^n$ then $\sum_{n=1}^{\infty}\left[\frac{2[n]_q-1-\vartheta}{1-\vartheta}|a_{i,n}| + \frac{2[n]_q+1+\vartheta}{1-\vartheta}|b_{i,n}|\right] \le 2$, (2.7)

by Theorem 2.2.

For $\sum_{i=1}^{\infty} t_i = 1, 0 \le t_i \le 1$, the convex combinations of f_i can be written as

$$\sum_{i=1}^{\infty} t_i f_i(\xi) = \xi - \sum_{n=2}^{\infty}\left(\sum_{i=1}^{\infty} t_i |a_{i,n}|\right)\xi^n + \sum_{n=2}^{\infty}\left(\sum_{i=1}^{\infty} t_i |\bar{b}_{i,n}|\right)\bar{\xi}^n$$

Then by (2.7)

$$\sum_{n=1}^{\infty}\frac{(2[n]_q-1-\vartheta)\phi_q(n,m)}{1-\vartheta}\sum_{n=1}^{\infty} t_i|a_{i,n}| + \sum_{n=1}^{\infty}\frac{(2[n]_q+1+\vartheta)\phi_q(n,m)}{1-\vartheta}\sum_{n=1}^{\infty} t_i|\bar{b}_{i,n}|$$

$$=\sum_{n=1}^{\infty} t_i\left(\sum_{n=1}^{\infty}\frac{(2[n]_q-1-\vartheta)\phi_q(n,m)}{1-\vartheta}|a_{i,n}| + \sum_{n=1}^{\infty}\frac{(2[n]_q+1+\vartheta)\phi_q(n,m)}{1-\vartheta}|\bar{b}_{i,n}|\right)$$

$\le 2\sum_{n=1}^{\infty} t_i = 2.$ ∎

Theorem 2.7

For $0 \le \beta \le \vartheta < 1$, $\phi_q(n,m) = \frac{\Gamma_q(n+m)}{(n-1)!\Gamma_q(1+m)}$, let $f \in \bar{\mathcal{H}}_q^m(\vartheta)$ and $F \in \bar{\mathcal{H}}_q^m(\beta)$ then

$f * F \in \bar{\mathcal{H}}_q^m(\vartheta) \subset \bar{\mathcal{H}}_q^m(\beta)$

Proof.

Let $f(\xi) = \xi - \sum_2^{\infty}|a_n|\xi^n + \sum_1^{\infty}\overline{|b_n|\xi^n} \epsilon \bar{\mathcal{H}}_q^m(\vartheta)$ and

$F(\xi) = \xi - \sum_2^{\infty}|A_n|\xi^n + \sum_1^{\infty}\overline{|B_n|\xi^n} \epsilon \bar{\mathcal{H}}_q^m(\beta)$. The coefficients of f *and* F must satisfy the conditions similar to (2.1) so for the coefficients of $f * F$ we write

$$f * F = \sum_{n=1}^{\infty}\left[\frac{2[n]_q-1-\vartheta}{1-\vartheta}|a_n A_n| + \frac{2[n]_q+1+\vartheta}{1-\vartheta}|b_n B_n|\right]$$

$$\leq \sum_{n=1}^{\infty}\left[\frac{2[n]_q - 1 - \vartheta}{1-\vartheta}|a_n| + \frac{2[n]_q + 1 + \vartheta}{1-\vartheta}|b_n|\right] \leq 2$$

Thus $f * F \in \overline{\mathcal{H}}_q^m(\vartheta) \subset \overline{\mathcal{H}}_q^m(\beta)$. ■

REFERENCES

1. *J. Clunie, T. Sheil-Small, Harmonic univalent functions, Ann. Acad. Sci. Fenn. Ser. A I. Math. 9 (1984) 3–25.*
2. *A. W. Goodman, On uniformly convex functions, Annal. Polonici Math., 56(1991), 86-92.*
3. *F.H. Jackson, On q-functions and a certain difference operator, Trans. Roy. Soc. Edinburgh 46 (1909) 253–281.*
4. *J.M. Jahangiri, Harmonic functions starlike in the unit disc, J. Math. Anal. Appl. 235 (1999) 470–477.*
5. *J.M. Jahangiri, Harmonic univalent functions defined by q-calculus operators, Inter. J. Math. Anal. Appl. 5 (2018), no. 2, 39–43.*
6. *J.M. Jahangiri, G. Murugusundaramoorthy, K. Vijaya, Starlikeness of Rucheweyh type harmonic univalent functions, J. Indian Acad. Math. 26 (2004) 191–200.*
7. *S. Kanas, D. Răducanu,Some subclass of analytic functions related to conic domains, Math. Slovaca 64(5) (2014) 1183–1196.*
8. *F. Rønning, Uniformly convex functions and a corresponding class of starlike functions, Proc. Amer. Math. Soc., 118(1993), 189-196.*
9. *S. Ruscheweyh, New criteria for univalent functions, Proc. Amer. Math. Soc. 49 (1975) 109–115.*
10. *H. Silverman, Harmonic univalent function with negative coefficients, J. Math. Anal. Appl., 220(1998), 283-289.*
11. *H. Silverman and E. M. Silvia, Subclasses of harmonic univalent functions, NewZealand J. Math.,* ***28****(1999), 275-284.*
12. *K. Vijaya, Studies on certain subclasses of harmonic functions, Ph.D. Thesis, VIT University, Vellore, 2014.*

DOI No. : 10.5958/2394-9309.2022.00038.5
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

TENSOR PRODUCT TENSOR SUM OF COMPOSITION AND MULTIPLICATIONOPERATOR INDUCED DYNAMICAL SYSTEM IN ANALYTIC LIPSCHITZ SPACES

G. S. Vigneswari[*], P. Chandra Kala[**] and V. Anusuya[***]
[*]Research Scholar (Full time) (18213152092033), Department of Mathematics, S.T. Hindu college, Nagercoil-629002, Tamil Nadu, India
[**]Assistant Professor, Department of mathematics, Sree Devi KumariWomen's College, Kuzhithurai-629163, Tamil Nadu, India
[***]Assistant Professor, Department of mathematics, S.T. Hindu College, Nagercoil-629002, Tamil Nadu, India
[Affiliated to ManonmaniamSundaranar University, Abishekapatti-Tirunelveli-627012, Tamil Nadu, India]
E-mail : vigneshwari.gs.vg@gmail.com*, kala.mathchandra@gmail.com**, anusuyameenu@yahoo.com***

ABSTRACT

In this section, introducing a new concept namely tensor product tensor sum of Composition and Multiplication operator induced dynamical system on analytic lipschitz spaces.

Keywords: Locally convex space, tensor sum operator, tensor product operator, multiplication operator, composition operators, dynamical system and toeplitz operator.

2000 AMS Subject Classification: 46 M 05, 47 A 80, 47 B 37.

1. INTRODUCTION

Let X and Y be Banach spaces and let $F(X,Y)$ be the topological vector space of vector-valued continuous (or) analytic functions from $X$to Y. Let $L(X,Y)$ be the vector space of all continuous((or) analytic) from X to Y. If $X=Y$ we write $F(X)$for $F(X,Y)$. Then each mapping $\varphi: X \to X$ gives rise to a linear transformation C_φ from $F(X,Y)$ itself, defined as

$$C_\varphi f = f \circ \varphi, \text{ for every } f \in F(X,Y)$$

and it is called a composition operator on $F(X,Y)$ induced by φ. Let $\psi: X \to \mathbb{C}$ be a mapping. Then the scalar multiplication gives rise to a linear transformation M_ψ from $F(X,Y)$ itself, defined as

$$M_\psi f = \psi f, \text{ for every } f \in F(X,Y),$$

where the product of functions are defined point-wise and are called a multiplication operator on $F(X,Y)$.Let $0<\alpha<1$, and define the metric ρ_α on D by $\rho_\alpha(z,w)=|z-w|^\alpha$. Then a function f analytic in D, is in $\mathcal{A}_\alpha$ if and only if there is a $K>0$ such that$|f(z)-f(w)| \le K\rho_\alpha(z,w)$. This definition can easily be extended by replacing ρ_α with any metric on any domain. Suppose Ω is a domain in the complex plane $\mathbb{C}$ and d is metric in Ω. Define X to be the space of functions analytic in Ω satisfying a lipschitz type condition in d. In other words, a function f, analytic in Ω, is in X if and only if

$$\sup\{|f(z) - f(w)|/d(z,w) : z \neq w\} < \infty.$$

It is quite easy to see that if we define the norm

$$\|f\| = \{|f(z) - f(w)|/d(z,w) : z \neq w\},$$

then$X$ is a banach space (provided that we identity functions that differ by a constant). We refer to a space defined in this manner as a Lipchitz type space on Ω induced by d.

Given a Banach space X of analytic functions in Ω, we can also define a function d_X on $\Omega \times \Omega$ as follows: $d_X(z,w) = \ sup|f(z) - f(w)|: \|f\|_X \leq 1\}$.

$d_X$is called the induced distance on X. Observe that $d_X$is a distance on Ω if X separates points of Ω and each point evaluation is a bounded linear functional on X. Note also that if X is a lipschitz type space induced by d, then, by definition, $d_X(z,w) \leq d(z,w)$ for $z.w \in H^2$ and equality holds if and only if d is induced by a banach space Y, not necessarily X. It is easy to show that $Y \subseteq X$ and equality holds if and only if X is a lipschitz type space. However the distance that induces X may not be d_X. It is possible that it is not even comparable to d_X, as we shall see below. The relationship between X and its induced distance d_X and between d and its induced lipschitz type space are studied in [1].

2. COMPOSITION AND MULTIPLICATION OPERATOR ON ANALYTIC LIPSCHITZ SPACES

Theorem:2.1. Let $\varphi: D \to D$ and $\pi_t : D \to \mathbb{C}$ be a continuous functions. Then $(C_\varphi M_{\pi_t})(f)$ is bounded for every $t \in \mathbb{R}$, $f \in D$and $z, w \in D$.

Proof:

We shall show that $C_\varphi M_{\pi_t}$ is continuous at the origin. we claim that $(C_\varphi M_{\pi_t})(B) \subseteq B$. We have,

$$\begin{aligned}
&\left|(C_\varphi M_{\pi_t})(f(z) - f(w))\right| \\
&\qquad \leq \left|(C_\varphi M_{\pi_t})f(z) - (T_\varphi M_{\pi_t})f(w)\right| \\
&\qquad \leq |(f \circ \varphi)(z)\pi_t(z) - (f \circ \varphi)(w)\pi_t(w)| \\
&\qquad \leq \|f(z) - f(w)\| \\
&\qquad \leq \|f\||z - w| \\
&\qquad \leq K\|f\|d_X(z,w) \text{for all } z, w \in D.
\end{aligned}$$

Therefore $f \in D$, $C_\varphi M_{\pi_t} \in D$ with, $\| C_\varphi M_{\pi_t}\| \leq K\|f\|$ is bounded with $\| C_\varphi M_{\pi_t}\| \leq K$; the proof is completed.

Theorem:2.2. Let $B(E)$ be the Banach algebra of all bounded linear operators on E. Let $(D,B(E))$be the space of bounded continuous functions from $D$to $B(E)$. Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$ in $(D, B(E))$ and let f_α be a sequence converging to f in (D,E). Then the product of $f_\alpha h_\alpha(\varphi_{t\alpha})$ converges to $fh(\varphi_{t)}$in (D,E).

Proof:

Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$in $(D,B(E))$.Then

$$\|f_n h_n(\varphi_{tn}) - fh(\varphi_t)\|_E$$

$= \sup\{\left\|f_n(x)h_n(\varphi_{t_n}(x)) - f(x)h(\varphi_t(x))\right\|: x \in X\}$

$\leq \sup\{\left\|f_n(x)h_n(\varphi_{t_n}(x)) - f_n(x)h(\varphi_t(x))\right\|: x \in X\}$

$+\sup\{\|f_n(x)h(\varphi_t(x) - f(x)h(\varphi_t(x))\|: x \in X\}$

$\leq \sup\{\|f_n(x)\|\left\|h_n(\varphi_{t_n}(x) - h(\varphi_t(x))\right\| : x \in X\}$

$+\sup\{\|f_n(x) - f(x)\|\|h(\varphi_t(x))\|: x \in X\}$

$\leq \left\|h_n(\varphi_{tn}) - h(\varphi_{t)}\right\|\|f_n\|\mathrm{E} + \left\|h(\varphi_{t)}\right\|\infty\|f_n - f\|\mathrm{E} \to 0$

as$\left\|h_n(\varphi_{tn}) - h(\varphi_{t)}\right\|_\infty \to 0$ and $\|f_n - f\|_\infty \to 0$.

as$\left\|h_n(\varphi_{t_n}) - h(\varphi_{t)}\right\|_\infty \to 0$ and $\|f_n - f\|_\infty \to 0$.

Theorem: 2.3. Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \to D(\mathbb{R})$be the function defined by $\nabla(t,f) = (C_{\varphi_t} M_{\pi_t})\,(f)$ for every $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. Then ∇ is a linear dynamical system on $\mathrm{D}(\mathbb{R})$.

Proof:

Since $C_{\varphi_t} M_{\pi_t}$ is a Toeplitz operator on $D(\mathbb{R})$for every $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. It can be easily seen that $\nabla(0,f) = f$ and thus it follows that $\nabla(t,f) \in \mathrm{D}(\mathbb{R})$for all $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. Clearly, ∇ is linear and

$\nabla(0,f)(z-w) = (C_{\varphi_t} M_{\pi_t})\,(f(z) - f(w))$ for all $z, w \in \mathbb{R}$

$= f(z) - f(w)$ for all $z, w \in \mathbb{R}$.

Therefore $\nabla(0,f) = f$.

Also $\nabla(t+s, f) = \nabla(t, \nabla(s,(f)))$.

Therefore, In order to show that ,∇ is a dynamical system on $D(\mathbb{R})$, it suffices to show that ∇ is separately continuous[25]. Let $t_n \to t$ and let $(t_{n,}f_n)$ be a net in$\mathbb{R} \times D(\mathbb{R})$such that $(t_{n,}f_n) \to (t,f)$.

We shall show that

$\nabla\,(t_{n,}f_n) \to \nabla(t,f) \in D(\mathbb{R})$.

Then $\left\|\nabla(t_{n,}f_n) - \nabla(t,f)\right\| = \left\|(C_{\varphi_{t_n}}M_{\pi_{t_n}})(f_n(\mathrm{z}) - f_n(\mathrm{w})) - (C_{\varphi_t}M_{\pi_t})f(z) - f(w))\right\|$

for every z, w $\in \mathbb{R}$.

$= \sup\{\left\|((C_{\varphi_{t_n}}M_{\pi_{t_n}})f_n(\mathrm{z}) - (C_{\varphi_{t_n}}M_{\pi_{t_n}})f_n(\mathrm{w})) - ((C_{\varphi_t}M_{\pi_t})f(z) - (C_{\varphi_t}M_{\pi_t})f(w))\right\|\}$ for every z, w $\in \mathbb{R}$.

$$= \sup\{\left\|(f_n \circ \varphi_{t_n})(\mathrm{z}) - (f \circ \varphi_t)(\mathrm{z})\right\|\left\|e^{tf(z)}\right\| \textit{for every } \mathrm{z} \in \mathbb{R}\}$$

$+ \sup\|f(\varphi_t)(\mathrm{z})\|\left\|e^{t_n f_{n(z)}} - e^{tf(z)}\right\|$ *for every* z $\in \mathbb{R}\}$

$$= \sup\{\left\|f_n(\varphi_{t_n})(\mathrm{z}) - f(\varphi_t)(\mathrm{z})\right\|\left\|e^{tf(z)}\right\| \textit{for every } \mathrm{z} \in \mathbb{R}\}$$

$$+ \sup\|f(\varphi_t)(\mathrm{z})\|\left\|e^{t_n f_{n(z)}} - e^{tf(z)}\right\| \textit{ for every } \mathrm{z} \in \mathbb{R}\}$$

$\to 0$ as$t_n - t \to 0$ a

nd$f(\varphi_{t_n}) \to f(\varphi_t)$. Therefore $\nabla(t,f)$ is a

dynamical system on $D(\mathbb{R})$.

Theorem: 2.4. Let $C_\varphi: D \to D$ and $\pi_t : D \to \mathbb{C}$ be a continuous functions. Then $(M_{\pi_t}C_\varphi)(f)$ is bounded for every $\in \mathbb{R}$, $f \in D$ *and* $z, w \in D$.

Proof:

We shall show that $C_\varphi M_{\pi_t}$ *is* continuous at the origin. we claim that $(M_{\pi_t}C_\varphi)(B) \subseteq B$. We have,

$$\left|(M_{\pi_t}C_\varphi)(f(z) - f(w))\right|$$
$$\leq \left|(M_{\pi_t}C_\varphi)f(z) - (M_{\pi_t}C_\varphi)f(w)\right|$$
$$\leq |\pi_t(z)(f \circ \varphi)(z) - \pi_t(w)(f \circ \varphi)(w)|$$
$$\leq \|f(z) - f(w)\|$$
$$\leq \|f\||z - w|$$
$$\leq K\|f\|d_X(z,w) \text{for all } z, w \in D.$$

Therefore $f \in D$, $M_{\pi_t}C_\varphi \in D$ with, $\left\|M_{\pi_t}C_\varphi\right\| \leq K\|f\|$ is bounded with $\left\|M_{\pi_t} C_\varphi\right\| \leq K$; the proof is completed.

Theorem: 2.5.Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \to D(\mathbb{R})$ be a function defined by $\nabla(t,f) = (M_{\pi_t}C_{\varphi_t})\ (f)$ for every $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. Then ∇ is a linear dynamical system on $D(\mathbb{R})$.

Proof:

Since $M_{\pi_t}C_{\varphi_t}$ is a Toeplitz operator on $D(\mathbb{R})$for every $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$. It can be easily seen that $\nabla(0,f) = f$ and thus it follows that $\nabla(\mathrm{t},f) \in \mathrm{D}(\mathbb{R})$for all $t \in \mathbb{R}$ and $f \in D(\mathbb{R})$.

Clearly, ∇ is linear and $\nabla(0,f)(z - w) = ((M_{\pi_t}C_{\varphi_t})\ (f(z) - f(w))$ for all $z, w \in D$.

Therefore $\nabla(0,f) = f$. Also $\nabla(t + s, f)(z - w) = \nabla(t, \nabla(s, f))$. Therefore, In order to show that , ∇ is a dynamical system on $D(\mathbb{R})$, it suffices to show that ∇ is separately continuous[25]. Let $t_n \to t$ and let $(t_{n,} f_n)$ be a net in$\mathbb{R} \times D$ such that $(t_{n,} f_n) \to (t, f)$. We shall show that $\nabla\ (t_{n,} f_n) \to \nabla(t, f)$ is a dynamical system on $D(\mathbb{R})$ .since by theorem 2.3.

3. TENSOR PRODUCT OF COMPOSITION AND MULTIPLICATION OPERATOR ON ANALYTIC LIPSCHITZ SPACES

Let $\pi_t : D \to \mathbb{R}$ defined by $\pi_t(x) = e^{th(x)}$ for all $t \in \mathbb{R}$ and $f \in D$, where $h \in H_b(D, \mathbb{R})$ and $\|h\|_\infty$= sup { $\|h(x)\|$: $h \in D$ }. Also $\varphi_t : \mathbb{R} \to \mathbb{R}$ is defined by $\varphi_t(\omega) = t + \omega$ the self-map[7].

Theorem:3.1. Let $\varphi: D \to D$ and $\pi_t : D \to \mathbb{R}$ be a continuous functions. Then $(C_\varphi \otimes M_{\pi_t})(f \otimes g)$ is bounded for every $t \in \mathbb{R}$, $f \otimes g \in D \otimes D$and $z, w \in D$.

Proof:

We shall show that $C_\varphi \otimes M_{\pi_t}$ *is* continuous at the origin.We claim that $(C_\varphi \otimes M_{\pi_t})(B) \subseteq B$.

We have

$$\begin{aligned}
&\left|(C_\varphi f \otimes M_{\pi_t} g)(z) - (C_\varphi f \otimes M_{\pi_t} g)(w)\right| \\
&\qquad = \left|C_\varphi f(z) - C_\varphi f(w)\right|\left|M_{\pi_t} g(z) - M_{\pi_t} g(w)\right| \\
&\qquad = |(f \circ \varphi)(z) - (f \circ \varphi)(w)||\pi_t(z) g(z) - \pi_t(w) g(w)| \\
&\qquad \leq K\|(f \otimes g)\| d_X(z, w) \text{For all } z, w \in D.
\end{aligned}$$

Therefore $f \otimes g \in D \otimes D$, $C_\varphi f \otimes M_{\pi_t} g \in D \otimes D$ with, $\left\| C_\varphi f \otimes M_{\pi_t} g\right\| \leq K\|f \otimes g\|$ is bounded with $\left\| C_\varphi \otimes M_{\pi_t}\right\| \leq K$; the proof is completed.

Theorem:3.2. Let $B(E)$ be the Banach algebra of all bounded linear operators on E. Let $(D,B(E))$be the space of bounded continuous functions from $D$to $B(E)$. Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$ in $D^\infty(D, B(E))$ and let f_α be a sequence converging to f in (D, E). Then the product of $f_\alpha h_\alpha(\varphi_{t\alpha})$ converges to $fh(\varphi_{t)}$in $(D, E) \otimes\ (D, E)$.

Proof:

Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$in $D^\infty(D,B(E))$Then$\left\|f_n h_n(\varphi_{t_n}) - fh(\varphi_t)\right\|_E \to 0$ as$\left\|h_n(\varphi_{t_n}) - h(\varphi_{t)}\right\|_\infty \to 0$ and $\|f_n - f\|_\infty \to 0$.

Theorem: 3.3. Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \otimes D(\mathbb{R}) \to D(\mathbb{R}) \otimes D(\mathbb{R})$ be the function defined by $\nabla(t, f \otimes g) = (C_{\varphi_t} \otimes M_{\pi_t})\ (f \otimes g)$ for every $t \in \mathbb{R}$ and $f \otimes g \in D(\mathbb{R}) \otimes D(\mathbb{R})$. Then ∇ is a linear dynamical system on D($\mathbb{R}$) $\otimes D(\mathbb{R})$.

Proof:

Since $C_{\varphi_t} \otimes M_{\pi_t}$is a Tensor product of composition and multiplication operator on $D(\mathbb{R}) \otimes D(\mathbb{R})$for every $t \in \mathbb{R}$and $f \otimes g \in D(\mathbb{R}) \otimes D(\mathbb{R})$. It can be easily seen that $\nabla(0, f \otimes g) = f \otimes g$ and $\nabla(t + s, f \otimes g) = \nabla(t, \nabla(s, (f \otimes g))$.

Therefore, In order to show that , $\nabla(t, f \otimes g)$ is a dynamical system on $D(\mathbb{R}) \otimes D(\mathbb{R})$, it suffices to show that ∇ is continuous. Let $g_\alpha \to g$ and let $(t_{\alpha,} f_\alpha \otimes g_\alpha)$ be a net in $\mathbb{R} \times D(\mathbb{R}) \otimes D(\mathbb{R})$such that $(t_{\alpha,} f_\alpha \otimes g_\alpha) \to (t, f \otimes g)$. We shall show that $\nabla (t_{\alpha,} f_\alpha \otimes g_\alpha) \to \nabla(t, f \otimes g)$.

Then $\left\|\nabla(t_{\alpha,} f_\alpha \otimes g_\alpha) - \nabla(t, f \otimes g)\right\|$

$$= \left\|\left((C_{\varphi_{t_\alpha}} f_\alpha \otimes M_{\pi_{t_\alpha}} g_\alpha\right)(\mathrm{z}) - (C_{\varphi_{t_\alpha}} f_\alpha \otimes M_{\pi_{t_\alpha}} g_\alpha)(\mathrm{w})\,) - \quad ((C_{\varphi_t} f \otimes M_{\pi_t} g)(z) - (C_{\varphi_t} f \otimes M_{\pi_t} g)(z))\right\|$$

$$= sup\{\left\|((C_{\varphi_{t_\alpha}} f_\alpha(z) \otimes M_{\pi_{t_\alpha}} g_\alpha(z) - (C_{\varphi_{t_\alpha}} f_\alpha(w) \otimes M_{\pi_{t_\alpha}} g_\alpha(w)) - ((C_{\varphi_t} f(z) \otimes M_{\pi_t} g(z)) - ((C_{\varphi_t} f(w) \otimes M_{\pi_t} g(w))\right\|$$

$\to 0$ as$|g_\alpha - g| \to 0$ and $f(\varphi_{t_\alpha}) \to f(\varphi_t)$.

Therefore $\nabla(t, f \otimes g)$ is a dynamical system on $D(\mathbb{R})) \otimes D(\mathbb{R})$.

Theorem: 3.4. Let $C_\varphi : D \to D$ and $\pi_t : D \to \mathbb{C}$ be a continuous functions. Then $(M_{\pi_t} \otimes C_\varphi)(f \otimes g)$ is bounded for every $\in \mathbb{R}$, $f \otimes g \in D \otimes D$.

Proof:

We shall show that $M_{\pi_t} \otimes C_\varphi$ *is* continuous at the origin.We claim that $(M_{\pi_t} \otimes C_\varphi)(B) \subseteq B$.

We have

$$\left|(M_{\pi_t} f \otimes C_\varphi g)(z) - (M_{\pi_t} f \otimes C_\varphi g)(w)\right|$$
$$= \left|M_{\pi_t} f(z) - M_{\pi_t} f(w)\right|\left|C_\varphi g(z) - C_\varphi g(w)\right|$$
$$= |\pi_t(z) f(z) - \pi_t(w) f(w)||(g \circ \varphi)(z) - (g \circ \varphi)(w)|$$
$$\leq K\|(f \otimes g)\| d_X(z, w) \text{For all } z, w \in D.$$

Therefore $f \otimes g \in D \otimes D, M_{\pi_t} f \otimes C_\varphi g \in D \otimes D$ with, $\left\|M_{\pi_t} f \otimes C_\varphi g\right\| \leq K\|f \otimes g\|$ is bounded with $\left\|M_{\pi_t} \otimes C_\varphi\right\| \leq K$; the proof is completed.

Theorem: 3.5.Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \otimes D(\mathbb{R}) \to D(\mathbb{R}) \otimes D(\mathbb{R})$ be a function defined by $\nabla(t, f \otimes g) = (M_{\pi_t} \otimes C_{\varphi_t})(f \otimes g)$ for every $t \in \mathbb{R}$ and $f \otimes g \in D(\mathbb{R}) \otimes D(\mathbb{R})$. Then ∇ is a linear dynamical system on $D(\mathbb{R}) \otimes D(\mathbb{R})$.

Proof:

Since $M_{\pi_t} \otimes C_{\varphi_t}$ is a Tensor product of composition and multiplication operator on $D(\mathbb{R}) \otimes D(\mathbb{R})$ for every $t \in \mathbb{R}$ and $f \otimes g \in D(\mathbb{R}) \otimes D(\mathbb{R})$. It can be easily seen that $\nabla(0, f \otimes g) = f \otimes g$ and $\nabla(t + s, f \otimes g) = \nabla(t, \nabla(s, f \otimes g))$.

Therefore, In order to show that , $\nabla(t, f \otimes g)$ is a dynamical system on $D(\mathbb{R}) \otimes D(\mathbb{R})$, it suffices to show that ∇ is continuous. Let $(t_{\alpha,} f_\alpha \otimes g_\alpha)$ be a net in $\mathbb{R} \times D(\mathbb{R}) \otimes D(\mathbb{R})$ such that $(t_{\alpha,} f_\alpha \otimes g_\alpha) \to (t, f \otimes g)$. We shall show that $\nabla(t_{\alpha,} f_\alpha \otimes g_\alpha) \to \nabla(t, f \otimes g)$. Then $\nabla(\mathrm{t}, f \otimes g)$ is a dynamical system on $D(\mathbb{R}) \otimes D(\mathbb{R})$. Since by theorem 3.3.

4. TENSOR SUM OF COMPOSITION AND MULTIPLICATION OPERATOR ON ANALYTIC LIPSCHITZ SPACES

Let $\pi_t : D \to \mathbb{R}$ defined by $\pi_t(x) = e^{th(x)}$ for all $t \in \mathbb{R}$ and $f \in D$, where $h \in H_b(D, \mathbb{R})$ and $\|h\|_\infty$= sup { $\|h(x)\| : h \in D$ }. Also $\varphi_t : \mathbb{R} \to \mathbb{R}$ is defined by $\varphi_t(\omega) = t + \omega$ the self-map[7].

Theorem:4.1. Let $\varphi: D \to D$ and $\pi_t : D \to \mathbb{R}$ be a continuous functions. Then $(C_\varphi \boxplus M_{\pi_t})(f \boxplus g)$ is bounded for every $t \in \mathbb{R}$, $f \boxplus g \in D \boxplus D$and $z, w \in D$.

Proof:

We shall show that $C_\varphi \boxplus M_{\pi_t}$ *is* continuous at the origin.We claim that $(C_\varphi \boxplus M_{\pi_t})(B) \subseteq B$.

We have

$$\begin{aligned}
&\left|(C_\varphi f \boxplus M_{\pi_t} g)(z) - (C_\varphi f \boxplus M_{\pi_t} g)(w)\right| \\
&= \left|C_\varphi f(z) - C_\varphi f(w)\right| \left|M_{\pi_t} g(z) - M_{\pi_t} g(w)\right| \\
&= |(f \circ \varphi)(z) - (f \circ \varphi)(w)| |\pi_t(z) g(z) - \pi_t(w) g(w)| \\
&\leq K\|(f \boxplus g)\| d_X(z, w) \text{For all } z, w \in D.
\end{aligned}$$

Therefore $f \boxplus g \in D \boxplus D$, $C_\varphi f \boxplus M_{\pi_t} g \in D \boxplus D$ with, $\| C_\varphi f \boxplus M_{\pi_t} g\| \leq K\|f \boxplus g\|$ is bounded with $\| C_\varphi \boxplus M_{\pi_t}\| \leq K$; the proof is completed.

Theorem:4.2. Let $B(E)$ be the Banach algebra of all bounded linear operators on E. Let $(D, B(E))$be the space of bounded continuous functions from $D$to $B(E)$. Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$ in $D^\infty(D, B(E))$ and let f_α be a sequence converging to f in (D, E). Then the sumof $f_\alpha h_\alpha(\varphi_{t\alpha})$ converges to $f h(\varphi_{t)}$in $(D, E) \boxplus (D, E)$.

Proof:

Let $h_\alpha(\varphi_{t\alpha})$ converges to $h(\varphi_t)$in $D^\infty(D, B(E))$Then$\left\|f_n h_n(\varphi_{t_n}) - f h(\varphi_t)\right\|_E \to 0$

as$\left\|h_n(\varphi_{t_n}) - h(\varphi_{t)}\right\|_\infty \to 0$ and $\|f_n - f\|_\infty \to 0$.

Theorem: 4.3. Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \boxplus D(\mathbb{R}) \to D(\mathbb{R}) \boxplus D(\mathbb{R})$ be the function defined by $\nabla(t, f \boxplus g) = (C_{\varphi_t} \boxplus M_{\pi_t})\ (f \boxplus g)$ for every $t \in \mathbb{R}$ and $f \boxplus g \in D(\mathbb{R}) \boxplus D(\mathbb{R})$. Then ∇ is a linear dynamical system on $\mathrm{D}(\mathbb{R}) \boxplus D(\mathbb{R})$.

Proof:

Since $C_{\varphi_t} \boxplus M_{\pi_t}$is a Tensor sum of composition and multiplication operator on $D(\mathbb{R}) \boxplus D(\mathbb{R})$for every $t \in \mathbb{R}$and $f \boxplus g \in D(\mathbb{R}) \boxplus D(\mathbb{R})$. It can be easily seen that $\nabla(0, f \boxplus g) = f \otimes g$ and $\nabla(t + s, f \boxplus g) = \nabla(t, \nabla(s, (f \boxplus g))$.

Therefore, In order to show that , $\nabla(t, f \boxplus g)$ is a dynamical system on $D(\mathbb{R}) \boxplus D(\mathbb{R})$, it suffices to show that ∇ is continuous. Let $g_\alpha \to g$ and let $(t_{\alpha,} f_\alpha \boxplus g_\alpha)$ be a net in$\mathbb{R} \times D(\mathbb{R}) \boxplus D(\mathbb{R})$such that $(t_{\alpha,} f_\alpha \boxplus g_\alpha) \to (t, f \otimes g)$. We shall show that $\nabla (t_{\alpha,} f_\alpha \boxplus g_\alpha) \to \nabla(t, f \boxplus g)$.

Then $\left\|\nabla(t_{\alpha,} f_\alpha \boxplus g_\alpha) - \nabla(t, f \boxplus g)\right\|$

$$= \left\|\left((C_{\varphi_{t_\alpha}} f_\alpha \boxplus M_{\pi_{t_\alpha}} g_\alpha\right)(\mathrm{z}) - (C_{\varphi_{t_\alpha}} f_\alpha \boxplus M_{\pi_{t_\alpha}} g_\alpha)(\mathrm{w})\ ) - \qquad ((C_{\varphi_t} f \boxplus M_{\pi_t} g)(z) - (C_{\varphi_t} f \boxplus M_{\pi_t} g)(z))\right\|$$

$$= sup\{\left\|((C_{\varphi_{t_\alpha}} f_\alpha(z) \boxplus M_{\pi_{t_\alpha}} g_\alpha(z) - (C_{\varphi_{t_\alpha}} f_\alpha(w) \boxplus M_{\pi_{t_\alpha}} g_\alpha(w)) - ((C_{\varphi_t} f(z) \boxplus M_{\pi_t} g(z)) - ((C_{\varphi_t} f(w) \boxplus M_{\pi_t} g(w))\right\|$$

$\to$ 0 as$|g_\alpha - g| \to$ 0 and $f(\varphi_{t_\alpha}) \to f(\varphi_t)$.

Therefore $\nabla(t, f \boxplus g)$ is a dynamical system on $D(\mathbb{R})) \boxplus D(\mathbb{R})$.

Theorem: 4.4. Let $C_\varphi : D \to D$ and $\pi_t : D \to \mathbb{C}$ be a continuous functions. Then $(M_{\pi_t} \otimes C_\varphi)(f \boxplus g)$ is bounded for every $\in \mathbb{R}$, $f \boxplus g \in D \boxplus D$.

Proof:

We shall show that $M_{\pi_t} \boxplus C_\varphi$ *is* continuous at the origin.We claim that $(M_{\pi_t} \boxplus C_\varphi)(B) \subseteq B$.

We have

$$\left|(M_{\pi_t} f \boxplus C_\varphi\, g)(z) - (M_{\pi_t} f \boxplus C_\varphi\, g)(w)\right|$$

$$= \left|M_{\pi_t} f(z) - M_{\pi_t} f(w)\right|\left|C_\varphi g(z) - C_\varphi g(w)\right|$$

$$= |\pi_t(z) f(z) - \pi_t(w) f(w)||(g \circ \varphi)(z) - (g \circ \varphi)(w)|$$

$$\leq K\|(f \boxplus g)\| d_X(z, w) \text{For all } z, w \in D.$$

Therefore $f \boxplus g \in D \boxplus D, M_{\pi_t} f \otimes \ C_\varphi \ g \in D \boxplus D$ with, $\left\|M_{\pi_t} f \boxplus \ C_\varphi \ g\right\| \leq K\|f \boxplus g\|$ is bounded with $\left\|M_{\pi_t} \boxplus \ C_\varphi\right\| \leq K$; the proof is completed.

Theorem: 4.5.Let $\nabla : \mathbb{R} \times D(\mathbb{R}) \boxplus D(\mathbb{R}) \rightarrow D(\mathbb{R}) \boxplus D(\mathbb{R})$ be a function defined by $\nabla(t, f \boxplus g) = (M_{\pi_t} \boxplus C_{\varphi_t})\ (f \boxplus g)$ for every $t \in \mathbb{R}$ and $f \boxplus g \in D(\mathbb{R}) \boxplus D(\mathbb{R})$. Then ∇ is a linear dynamical system on $D(\mathbb{R}) \boxplus D(\mathbb{R})$.

Proof:

Since $M_{\pi_t} \boxplus C_{\varphi_t}$ is a Tensor sum of composition and multiplication operator on $D(\mathbb{R}) \boxplus D(\mathbb{R})$ for every $t \in \mathbb{R}$ and $f \boxplus g \in D(\mathbb{R}) \boxplus D(\mathbb{R})$. It can be easily seen that $\nabla(0, f \boxplus g) = f \boxplus g$ and $\nabla(t + s, f \boxplus g) = \nabla(t, \nabla(s, f \boxplus g))$.

Therefore, In order to show that $, \nabla(t, f \boxplus g)$ is a dynamical system on $D(\mathbb{R}) \boxplus D(\mathbb{R})$, it suffices to show that ∇ is continuous. Let $(t_{\alpha,} f_\alpha \boxplus g_\alpha)$ be a net in $\mathbb{R} \times D(\mathbb{R}) \boxplus D(\mathbb{R})$ such that $(t_{\alpha,} f_\alpha \boxplus g_\alpha) \rightarrow (t, f \boxplus g)$. We shall show that $\nabla(t_{\alpha,} f_\alpha \boxplus g_\alpha) \rightarrow \nabla(t, f \boxplus g)$. Then$\nabla$(t, $f \boxplus g$) is a dynamical system on $D(\mathbb{R}) \boxplus D(\mathbb{R})$. Since by theorem 4.3.

REFERENCES:

1. *J. S. Manhas, Weighted composition operators on weighted locally convex spaces of analytic functions, Southeast Asian Ball. Math.29(2005), no.1, 127-140.*
2. *D. Senthilkumar and P. Chandra Kala, Operator and Dynamical system in measurable function spaces, Global journal of mathematical sciences: theory of practical, Vol 4, no. 2(2012), pp 107-104.*
3. *D. Senthilkumar and P.Chandra Kala, Tensor Sum and Dynamical systems, Acta MathematicaScientia. Vol 4,No. 2(2012),pp 107-114.*
4. *D. Senthilkumar, P. Chandra Kala and T. Prasad, Tensor product operators induce Dynamical systems on weighted locally convex space, Jordan Journal of Mathematics and statistics(JJMs) 6, 2013, pp.169-181*
5. *J.H. Shpiro, composition operator and classical function theory, Universal: Tracts in Mathematics, Springer, New York, Ny, USA, 1993 hpiro, composition operator and classical*
6. *R. K. Singh and J. S. Manhas, composition operator on function spaces, North Holland Mathematics Studies, 1993.*
7. *R. K. Singh, and J. S. Manhas, Multiplication operators and dynamical systems.Journal of Austral.Math.Soc., (Series A), 53, 92-102, 1992.*
8. *R. K. Singh, and J. S. Manhas, Composition operators onweighted spaces of vector-valued continuous functions, Acta. Sci. Math, 1988.*
9. *R. K. Singh, and R. David Chandra Kumar, Weighted composition operators on functional spaces. Bull. Aust. Math. Soc., vol 31, 117-126, 1985.*

10. *R. K. Singh, and J. S. Manhas, Multiplication operators on weighted spaces of vector-valued continuous functions. J. Aust. Math. Soc., (Series A) vol 50, 98-107, 1991.*
11. *R. K. Singh, and J. S. Manhas, weighted composition operators and weighted space of cross- section. Actascientiarummathematicarum, Vol 59, no-3-4, 1994.*
12. *R. K. Singh, and J. S. Manhas, Composition operators and Multiplication operators on weighted spaces of analytic functions, research article, international journal of mathematics and mathematical sciences, vol 2007.*
13. *R. K. Singh, and J. S. Manhas, Composition operators on function spaces, North Holland publishing Co., Amsterdom, London, N.Y.1993*
14. *R. K. Singh, and W.H. Summers: Composition operator on weighted spaces of continuous functions, J. Austral. Math Soc. Series A 45 (1988), 303-319.*
15. *R. K. Singh, V. Sahai and V. K. Pandey: Weighted substitution operator on* $L^p(\mathbb{R})$ *and wavelet theory, mathematics Newsletter (Ramanujam Math Soc), Vol.21, June 2011, 6-16.*
16. *G.S. Vigneswari, P.Chandra Kala And V.Anusuya Dynamical system induced by tensor sum tensor product and toeplitz operator on weighted continuous and locally convex space, Advances and Application in Mathematical Sciences, vol 21, issue 3, January 2022, pages 1139-1156.*
17. *G.S. Vigneswari, P. Chandra Kala And V. AnusuyaDynamical system induced by weighted tensor sum and tensor product operator with Lie group representations Page 1 Malaya Journal of Matematik, Vol. 9, No. 1, 409-414, 2021*
18. *G.S. Vigneswari, P. Chandra Kala And V. AnusuyaOperator and Dynamical System In Analytic Lipschitz Space, Journal of Physics: Conference Series, Volume 1947*

DOI No. : 10.5958/2394-9309.2022.00039.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

MATHEMATICAL METHOD OF FORECASTING BASED ON FUZZY TIME SERIES

V. Anithakumari, S. Imalin, V.M. Arul Flower Mary
Assistant Professor, Department of Mathematics, Muslim Arts College, Thiruvithancode, 629174
*Research Scholar , Department of Mathematics, Holy Cross College, Nagercoil. 629004
Associate Professor, Department of Mathematics, Holy Cross College, Nagercoil. 629004
Affiliated to Manonmaniam Sundaranar University Abishekapatti,Tirunelveli-627012
E-mail : anithakumari_v@yahoo.co.in, imalinjas25@gmail.com, arulflowermary@gmail.com

ABSTRACT
Forecasting accuracy is one of the most favorable critical issues in Fuzzy Time Series models. This study compares the application of three forecasting methods on the Hydro power Generation of in India. The Fuzzy time series including the Chen rules, Yule rule, and Doane rule. The Yule's rule and Doane rule show smaller than predicted error and closer. Fuzzy models resulted in more accurate forecast of the Hydro power Generation in India. However, this impact reduces with time and production of hydro power generation in the time series analysis. The Yule's rule and Doane rule can be utilized to predicted hydro power generation production value accurately.

Keywords: Fuzzy Time series; Sturges rule; Yule's pravidlo rule, Chen rule, Linguistic Variable and Hydropower Generation.

1. INTRODUCTION

In the last decade, fuzzy time series have received more concentration to deal with the vagueness and incompleteness inherent in time series data. Different types of models have been developed either to get better forecasting accuracy or reduce computation overhead. However, the issues of controlling uncertainty in forecasting, effectively partitioning intervals, and consistently achieving forecasting accuracy with different interval lengths have been rarely investigated. In the literature survey most of the model is of first order fuzzy time series model. In the past decade many forecasting models based on the concepts of fuzzy time series have been proposed. It has been applied to predict enrollments, temperature, crop production and stock index, etc. Time series forecasting studied the relations on the sequential set of past data measured over time to forecast the future values. Forecasting technique are frequently conducted by using statistical tools like regression analysis, moving averages, integrated moving average and autoregressive integrated moving average. The historical data defined as linguistic terms are not being considered in forecasting technique which is one of the major limitations of these methods. Fuzzy set theory and fuzzy logic was first introduced by Zadeh (1965) which provides a general method for usage uncertainty and vagueness in information available in linguistic terms. Song and Chissom (1993) used the fuzzy set theory given by Zadeh to develop models for fuzzy time series forecasting and considered the problem of forecasting enrollments on the time series data of University of Alabama.

2. FUZZY SET AND FORECASTING

The fuzzy sets theory can be defined as a mathematical formalization that enables us to eliminate indefiniteness and deal with incomplete, inaccurate information of both qualitative and quantitative by nature. The fuzzy sets theory, advanced one of the well-known representatives of modern applied mathematics, by excluding any definite report of the task offers such a solution scheme of the problem that a personal reasoning and evaluation plays a principal role in evaluating indefinite, unclear fact. Thus anyone, encountering indefinite, incomplete information and data, can form some conclusion, if even in a rough way, by passing through his/her reasoning all these realities. The use of fuzzy verbal notions in every-day speech (much, more, little, small, many, a Number of etc.) enables us to give a qualitative report of the problem which must be tackled and take account of its indefinite nature as well as attain the report of the factors that can't be described qualitatively.

The advent of fuzzy logic made it possible to tackle a great many problems with fuzzy input data. One of them was a forecasting problem. Many of the structural elements of the latter (input data and interdependence between its components, interval evaluation of indicators and their interdependence, expert evaluations and judgments etc.) are either of a fuzzy nature or, by being in fuzzy relationships, condition the fuzzy description of the problem.

The application of fuzzy logic to the usage of forecasting problems was undertaken by the researches in which the mathematical models of fuzzy time series were described in a fuzzy form for conduct the problem with fuzzy input data This approach was developed later by other scientists dealing with the solution of similar problems To tackle the task, the authors proposed a model of fuzzy time series and tried to reduce the average forecasting error by making adequate alterations in the model.

Thus, the major purpose of the proposed approach is methodological: 1) putting forth an evaluation method based on fuzzy time series for estimating model parameters; 2) testing the extent to which the model is adequate to reflect the real process, that is to say, computing the method error; 3) conducting the comparative analysis of computation results; 4) revealing the practical and theoretical importance of the model.

3. FUZZY TIME SERIES

Fuzzy Time series is assumed to be a fuzzy variable along with associated membership function Song and Chissom [1993] have proposed a procedure for solving fuzzy time series model described in the following steps.Time series represents a consecutive series of observation that is conducted by equal time intervals and lies at the root of exploring real processes in economics, meteorology and natural sciences etc.

Universe of Discourse:

The universe of discourse, where $U = \{u_1, u_2 \ldots u_n\}$ of the given historical data the minimum data is denoted D_{min} and the maximum data is denoted D_{max}. The universe of discourse.

$U = [D_{min} - D_1, D_{max} - D_2]$

Where, D_1,D_2 are two positive real number, A Fuzzy set A_i of U is defined by$A_i = f_{Ai}(u_1)/u_1 +$
$f_{Ai}(u_2)/u_2 + \ldots \ f_{Ai}(u_n)/u_n$

Where f_{Ai} is the membership function of the fuzzy set.

$A_i, f_{Ai} : U \rightarrow [0,1]$

A. *Fuzzy set:*

A fuzzy set is totally characterized by a membership function(MF). Fuzzy set introduced by lotfi A.Zadeh and Dieter klaua in 1965. Fuzzy set can be applied in more then field.

If there exists a fuzzy relationship R(t-1,t), such that $F(t) = F(t-1) \circ R(t-1,t)$, Where, $\circ$is an operator, then F(t) is said to be caused by F(t-1) . The relationship between F(t) and F(t- 1) can be denoted by $F(t-1) \rightarrow F(t)$

Membership function

Let X be a nonempty set. A fuzzy set A in X is characterized by its membershipfunction A: X [0, 1]. The mapping A is also called the membership function of fuzzy set A.

$\mu_F(x) : U \rightarrow [0,1]$

Fuzzy Relationship:

Fuzzy logical relationship can be further grouped together into fuzzy logicalrelationship groups according to the equal sides of the fuzzy logical relationships

For example, there are fuzzy logical relationships with the same left hand sides Ai

$A_i \rightarrow A_{j1}$

$A_i \rightarrow A_{j2}$

…

$A_i \rightarrow A_{j1}, A_{j2}, \ldots$

Song and Chissom applied time invariant and time variant models to forecast the enrollment at the University of Alabama. The time variant and invariant models include the following steps:

- Define universe of discourse and the intervals
- Divider the intervals

- Classify the fuzzy sets
- Fuzzify the data
- Create the fuzzy relationships
- Defuzzify the forecasting results
- Forecasting

CHEN RULE

A rule for decisive the desirable number of groups into which a distribution of observation should be classified the number of groups or classes is 1+3.322log n, where n is the number of observation.

Sturges rule is used to decide the number of classes that will best show the outline of a distribution given a set of information.

Fuzzy time series using Sturges rule:

A best forecasting accuracy using time variant and time invariant fuzzy time series and it is emphasized that the forecast uses only historical data. The significance of thesis is to reduce the mean square error when compared with the existing forecast approaches.

The appropriate number of intervals is computed as

$$I = 1 + 3.322 \log_{10} n$$

Where, I is the number of classes in the data. n is the number of observation in the data and log is common lag value to n.

To compute the length of intervals as

$$L = \max - \min / I$$

Yule's Rule

Yule rules (1929) recommended the following formula to arrive at approximate number ofclass.

$$I = 2.5 * (\sqrt[4]{n})$$

Where, I is the number of classes in the data. n is the number of observation in the data

Doane Rule

Doane (1976) proposed the method for calculating the number of classes according to

$$I = 3.3 * \log_{10} n$$

4 . FORECASTING HYDRO POWER GENERATION

To implement of the above rule for forecasting hydro power generation production in India. Production is based on 27 year (1995 to 2021) time series Indian hydro power generation production data.

Sturges Rule

Step 1: A best forecasting accuracy using time invariant fuzzy time series and it is emphasized that the forecast uses only historical data. The significance of reduce the mean square error when compared with the existing forecast approaches.

We are discussing our proposed method Sturges rule. The historical data and proposed method are Table1. The length of Intervals six and the universe of discourse U = [2000, 6500] , the length of intervals is 750.

Step2:Partition the universe of discourse into six equal length intervels $u_1, u_2, \ldots, u_6,$ where u_1 = [2000 , 2750] ,u_2 = [2750, 3500] , u_3 = [3500 , 4250] ,u_4 = [4250 , 5000], u_5 = [5000 ,

5750] ,u_6 = [5750 , 6500].

Step3: Fuzzy logical relationship and fuzzy logical relationship groups. From A_i in table 1, the fuzzy logical relationship. The fuzzy logical relationship can be rearranged into fuzzy logical relationship groups as in table 2.

Table 1: production Data sets

Intervels	Fuzzy set
u1= [2000 2750]	A1
u2= [2750 3500]	A2
u3= [3500 4250]	A3
u4= [4250 5000]	A4
u5= [5000 5750]	A5
u6= [5750 6500]	A6

Table 2: Group of Fuzzy logical relationship groups

Fuzzy logical relationship
A1 → A1, A2
A2 → A2, A3
A3 → A3,A4
A4 → A4,A5
A5 → A5,A6
A6 → A6

Step4: Fuzzy set A_i , In this case the linguistic variable is "hydro power production in India "Each fuzzy set A_i is assigned to a linguistic term: A_1 = (Poor production),A_2 = (Average production),A_3 = (Good production),A_4 = (Very good production),A_5 = (Excellent production),A_6 = (Bumper production). Each A_i is defined by the intervals $u_1, u_2, \ldots, u_6$.

$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6$ $A_2 = 0.5/u_1 + 1/u_2$
$+0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6$ $A_3 = 0.5/u_1 + 0.5/u_2 + 1/u_3 + 0/u_4$
$+0/u_5 + 0/u_6$

$A_4 = 0.5/u_1 + 0.5/ u_2 +0.5/u_3 + 1/ u_4 +0/u_5 + 0/ u_6$ $A_5 = 0.5/u_1 + 0.5/ u_2 +0.5/u_3 + 0.5/ u_4 +1/u_5 + 0/ u_6$ $A_6 = 0.5/u_1 + 0/ u_2 +0/u_3 + 0/ u_4 +0.5/u_5 + 1/ u_6$

Table 1 list the natural rubber production data from 1995 to 2021 and thecorresponding fuzzy production forecasting A_i.

Step 5: If the trend in hydro power value leads to an increase, then the fuzzy sets in the height part are all selected. Otherwise, If the trend in power production value leads to a decrease, then the fuzzy sets in the low part are all selected. Otherwise, if the trend in power generation production value leads to no change, in table 2 the basis of fuzzy set is selected.

Yule's Rule

Yule rules (1929) recommended the following formula to arrive at approximate numberof class.

$$I=2.5*(4\sqrt{n})I=5$$

Where, I is the number of classes in the data. n is the number of observation in the data

Doane Rule

Doane (1976) proposed the method for calculating the number of classes according to

$$I = 3.3*\log_{10} n$$

4 . FORECASTING HYDRO POWER GENERATION

To implement of the above rule for forecasting hydro power generation production in India. Production is based on 27 year (1995 to 2021) time series Indian hydro power generation production data.

STURGES RULE

Step 1: A best forecasting accuracy using time invariant fuzzy time series and it is emphasized that the forecast uses only historical data. The significance of reduce the mean square error when compared with the existing forecast approaches.

We are discussing our proposed method Sturges rule. The historical data and proposed method are Table1. The length of Intervals six and the universe of discourse U = [2000, 6500] , the length of intervals is 750.

Step2: Partition the universe of discourse into six equal length intervels $u_1, u_2, \ldots, u_6,$ where u_1 = [2000 , 2750] ,u_2 = [2750, 3500] , u_3 = [3500 , 4250] ,u_4 = [4250 , 5000], u_5 = [5000 , 5750] ,u_6 = [5750 , 6500].

Table 3: production Data sets

Intervels	Fuzzy set
u_1 = [2000 2900]	A1
u_2 = [2900 3800]	A2
u_3 = [3800 4700]	A3
u_4 = [4700 5600]	A4
u_5 = [5600 6500]	A5

Table 4: Group of Fuzzy logical relationship groups

Fuzzy logical Relation ship
A1 → A1, A2
A2 → A2, A3
A3 → A3,A4
A4 → A4,A5
A5 → A5,A6

Step3: Fuzzy logical relationship and fuzzy logical relationship groups. From A_i in table 3, the fuzzy logical relationship. The fuzzy logical relationship can be rearranged into fuzzy logical relationship groups as in table 4.

Step4: Fuzzy set A_i , In this case the linguistic variable is "Hydro power generation production in India" Each fuzzy set A_i is assigned to a linguistic term: A_1 = (Poor production),A_2 = (Average production),A_3 = (Good production),A_4 = (Very good production),A_5 = (Excellent production). Each A_i is defined by the intervals $u_1, u_2, \ldots, u_5.$

$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5$ $A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5$

. . .

. . .

. . .

$A_5 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5$

Table 3 list the Hydro power generation production data from 1995 to 2021 and the corresponding fuzzy production forecasting A_i.

Step 5: If the trend in power generation value leads to an increase, then the fuzzy sets in the height part are all selected. Otherwise, if the trend in hydro power generation production value leads to a decrease, then the fuzzy sets in the low part are all selected. Otherwise, if the trend in natural rubber production value leads to no change, in table 4 the origin fuzzy set is selected.

DOANE RULE

Doane rule is used to divide the class and intervals

$$I = 3.3 * \log_{10} n$$

Where, I is the number of classes in the data. n is the number of observation in the data.

Step 1: We are discussing our proposed method Doane Rule. The historical data and proposed method are Table3. The total number of observation is 27 .The length of Intervals five and the universe of discourse U = [2000, 6500] , the length of intervals is 900.

Table 5: production Data sets

Intervels	Fuzzy set
u_1= [2000 2900]	A1
u_2= [2900 3800]	A2
u_3= [3800 4700]	A3
u_4= [4700 5600]	A4
u_5= [5600 6500]	A5

Table 6: Group of Fuzzy logical relationship groups

Fuzzylogical Relationship
A1 → A1, A2
A2 → A2, A3
A3 → A3,A4
A4 → A4,A5
A5→ A5,A6

Step2:Partition the universe of discourse into five equal length intervels u1,u2,...,u5, where u1 = [2000 2900], u2 = [2900 3800],u3 = [3800 4700],u4 = [4700 5600],u5 = [5600

6500] .

Step3: Fuzzy logical relationship and fuzzy logical relationship groups. From A_i in table 5, the fuzzy logical relationship. The fuzzy logical relationship can be rearranged into fuzzy logical relationship groups as in table 6.

Step4: Fuzzy set A_i , In this case the linguistic variable is "Hydro power generation production in India" Each fuzzy set A_i is assigned to a linguistic term: A_1 = (Poor production),A_2 = (Average production),A_3 = (Good production),A_4 = (Very good production),A_5 = (Excellent production). Each A_i is defined by the intervals $u_1, u_2, \ldots, u_5$.

$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5$ $A_2 = 0.5/u_1$
$+ 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5$

. . .

. . .

. . .

$A_5 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5$

Table 3 list the Hydro power generation production data from 1995 to 2021 and the corresponding fuzzy production forecasting $A_{i.}$

Step 5: If the trend in power generation value leads to an increase, then the fuzzy sets in the height part are all selected. Otherwise, if the trend in hydro power generation production value leads to a decrease, then the fuzzy sets in the low part are all selected. Otherwise, if the trend in natural rubber production value leads to no change, in table 4 the origin fuzzy set is selected.

5. A COMPARISON OF THE RESULT OF THREE RULES

We compare the forecasting results of the three rules :

A Comparison of the mean square errors (MSE) of different rules is shown in Table 7,where the mean square error (MSE) is defined as follows:

$$MSE = \frac{\sum_{i=1}^{n} (Actual\ production\ value of\ i - Forecast\ production\ value\ of\ i)^2}{n}$$

Table 7: A Comparison of the MSE of Three Rules

	Chen Rule	Yule's Rule	Doane Rule
MSE	3.192	2.12	2.12

Table 8: Forecasting Value of Hydropower Generation in Yule and Doane Method

Year	Hydropower	Forecasting	Year	Hydropower	Forecasting
1995	3963.77		2009	5307.294	5489.891
1996	3957.082	4213.857	2010	5090.157	5489.891
1997	4870.091	3620.446	2011	2125.82	5638.541
1998	5181.323	5489.891	2012	3915.120	4359.217
1999	4502.29	4359.219	2013	4111.131	4259.219
2000	5367.981	5489.891	2014	5121.272	5479.891
2001	4603.738	3620.446	2015	3077.125	1078.686
2002	2737.409	837.686	2016	6469.425	6279.958
2003	2334.426	5638.545	2017	5489.891	5579.891
2004	3916.55	4359.219	2018	5489.891	5579.891
2005	5807.139	7484.852	2019	6363.337	6279.958
2006	6469.919	6279.958	2020	5307.294	5479.891
2007	6340.202	5489.891	2021	5489.891	5579.891
2008	5625.227	5489.891			

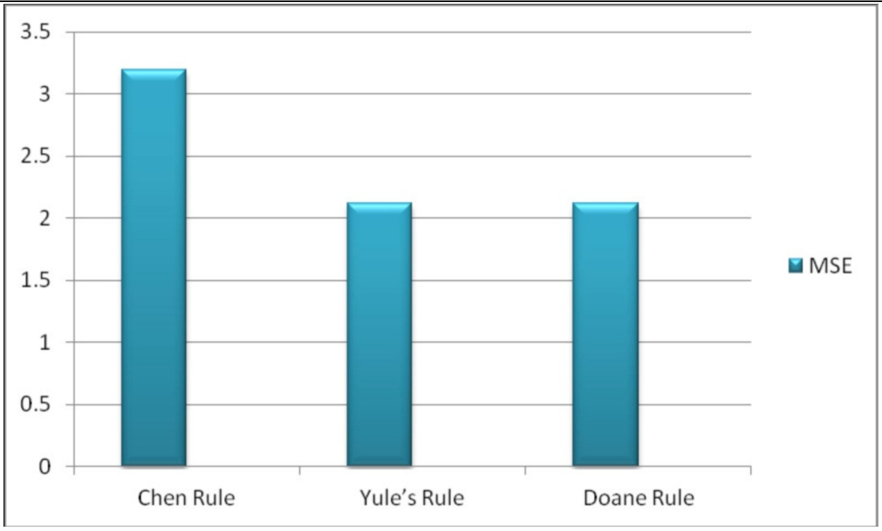


Figure 1: The Mean Square Error result of the Three Rule in Hydro power Generation production in India

CONCLUSION

In this paper, we presented a time invariant fuzzy time series method for forecasting Hydro Power generation production in India. In case of hydro power production all the three rule provide similar forecast .The forecasted value from Yule and Doane rule are in close agreement with each other, where as the Chen rule exhibits some variation with the Yule and Doane rule and which can be visualized in Fig 1.The testing of these three rules shows that all the three rules under study provide forecast with an mean Square Error (MSE) for Chen rule, Yule's rule and Doane rules are Table 7 respectively.

REFERENCES

1. *Chi-chen wang (2019), A comparison study between fuzzy time series model and ARIMA model for Forecasting Taiwan export, Expert system with Application, 38, 2019.9296-9304.*
2. *Yun-sheng Hsu et.al (2021), A Comparison of ARIMA forecasting and heuristic modeling, Applied Financial Economics, 2021, 21, 1095-1102.*
3. *P.Yolum et.al (2005) , A New Trend Heuristic Time variant fuzzy time series method for forecasting enrollment, Springer-verlag Berlin Heidelberg 2005,ISCIS2005,LNCS 3733,pp 553-564,2005.*
4. *R.K.Srivastava (2018), Fuzzy – Neural Techniques for Short Term Forecast of Food Grain Production, International journal of Information Technology and Knowledge Management, 2018, volume 4, No.2, Pp.385-390.*
5. *Narendarakumar et.al (2010), Fuzzy time series forecasting of Wheat production, International on Computer Science and Engineering l02, No.03, 2010, 65-640.*
6. *Shyi-ming Chen (2002),"Forecasting enrollment Based on high order fuzzy time series," Cybernetics and system, An International Journal, 33, 1-16, 2002 copyright, Taylor and Francis.*
7. *Chung-ming own pao-Ta-Yu (2005), Forecasting fuzzy time series on a heuristic high order model,Cybernetics and system, An International Journal, 36,705-717,copyright 2005, Taylor& Francis Inc.*
8. *Kunhuang Huarng (2001), Heuristic model of fuzzy time series for forecasting, Fuzzy set and system 123, 2001, 369-386, Elsevier.*
9. *Hemant kumar Pathak et.al (2011), A New Bandwidth Interval Based Forecasting Method for Enrollment Using Fuzzy Time Series, Scientific Research, Applied Mathematics, 2011, 2,504 -507.*
10. *Tahseen A.Jilani, et.al (2007), Multivariate High Order Fuzzy Time Series Forecasting for Car Road Accidents. World Academy of Science, Engineering andTechnology 25, 2007.*

DOI No. : 10.5958/2394-9309.2022.00039.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

PREDICTION, ACCURACY, CLASSIFICATION AND VISUALIZATION OF RAINFALL DATABASE USING MACHINE LEARNING ALGORITHMS IN NORTH EAST SATES OF INDIA

P. Arumugam*, A. Poom Pavai, A. Poongothai***, Manimannan G******
and R. Lakshmi Priya*****
*Assistant Professor, Department of Statistics, Annamalai University, Chidambaram.
,** Assistant Professor, Department of Statistics, Apollo Arts and Science College Chennai, Guduvanchery, Chennai
*** Assistant Professor, Department of Statistics, Muthayammal College of Arts and Science, Namakkal.
***** Assistant Professor, Department od Statistics, Dr. Ambedkar Govt. Arts College, Vyasarpadi, Chennai

ABSTRACT

This research paper attemptsto identify the rainfall database prediction, accuracy, classification and visualization in the North East States (NES) namely Arunachal Pradesh, Assam, Meghalaya, Manipur, Tripura, Mizoram and Nagaland (Sister of Seven States) in India. The secondary sources of data were collected from Indian Metrological Department from 1901 to 2017. The database was available in two categories,monthly and seasonal. This research paper analyze seasonal database of NES using machine learning algorithms. Indian Climates is split into four categories and are named as winter (December to February), summer (March to May), monsoon (June to September) and post monsoon (October to November).Initially the seasonal rainfall database is normalized and visualized. Five machine learning algorithms achieved 99 percent classification accuracy and remaining one percent is misclassified due to large volume of database or outliers. Kappa Statistics of Naïve Bayes and Multilayer Perceptron (MLP) are 0.91 and 0.96 and rest of the machine learning algorithms J48, Random Forest and Instance Based Learner (IBk) results shows 1. The Kappa Statistics >0.80 is almost perfect classification. The summary statistics of various machine algorithms were discussed. Finally the four seasons of NES rainfall database were classified as Very High Rainfall, High Rainfall, Moderate Rainfall and Low rainfall.

Keywords: NES, Rainfall, Seasons, Machine Learning Algorithms, Classification, Prediction, Accuracy and Visualization.

1. INTRODUCTION

India's most of the economy is based on agriculture. The one of the primary source for its success is water. There are many water reservoirs such as Bore wells, canals, rivers which can be used for agricultural purposes. But the main resource for these all reservoirs is rain water. Due to geographical and economic conditions every farmer is unable to afford bore well and other conventional systems. Hence, rainfall is a greatest asset and is most essential factor in agriculture. To predict the rainfall is vital task and its accuracy will provide benefits to various fields such as agriculture, pesticides, tourism, event management, water conservation, flood and drought prediction, etc. Here we are going to use five most popular machine learning techniques to predict the rainfall. Those techniques are Naïve Bayes, Multilayer Perceptron, J48, Random Forest and Instance Based Learner (IBL). This study will analyze and comparing their results with Kappa Statistics of various ML Algorithms.

2. REVIEW OF LITERATURE

Mohini Prasad Mishra study carries historical weather data collected locally at Faisalabad city, Pakistan that was analyzed for useful knowledge by applying data mining techniques. Data includes ten years' period [2007-2016] and it had been tried to extract useful practical knowledge of weather data on monthly based historical analysis. Analysis and investigation was done using data mining techniques by examining changing patterns of weather parameters which includes maximum temperature, minimum temperature, wind speed and rainfall. After preprocessing of data and outlier analysis, K-means clustering algorithm and Decision Tree algorithm were applied. Two clusters were generated by using K-means Clustering algorithm with lowest and highest of mean parameters. Whereas in decision tree algorithm, a model was developed for modeling meteorological data and it was used to train an algorithm known as the classifier. 10-fold cross validation used to generate trees. The result obtained with smallest error (33%) was selected on test data set. While for the number of rules generated of the given tree was selected with minimum error of 25%. The results showed that for the given enough set data, these techniques can be used for weather analysis and climate change studies.

Nabila WardahZamani and SitiShalizaMohdKhairi paper, studies on data mining techniques used to predict rainfall using meteorological data of Subang Weather Station collected from January 2009 to December 2016. The data preparation process involves five weather factors which are maximum temperature, minimum temperature, evaporation, wind speed and cloud with 2922 observations. Predictive Decision Tree model, Artificial Neural Network model and Naïve Bayes model are developed for rainfall prediction and comparison. As a results show that the performance of Decision Tree model is better as compared to the other predictive models with the misclassification rate of 0.15 and RMSE=0.35. Given enough set of data, rainfall can be predicted using the data mining techniques.

3. DATA SOURCE

The data for the study are secondary sources of data and were collected from Indian Metrorological Department (IMD) during the Period of 1901 to 2017 with the parameter of twelve months.

4. METHODOLOGY AND MACHINE LEARNING ALGORITHM

4.1 Naïve Bayes

The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:Naïve: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Bayes: It is called Bayes because it depends on the principle of Bayes' Theorem. Bayes' theorem is also known as Bayes' Rule or Bayes' law, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.The formula for Bayes' theorem is given as:

$$P(\frac{A}{B}) = \frac{P\left(\frac{B}{A}\right)P(A)}{P(B)}$$

Where,

$P(\frac{A}{B})$is Posterior probability: Probability of hypothesis A on the observed event B.

$P\left(\frac{B}{A}\right)$is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

$P(A)$**is Prior Probability**: Probability of hypothesis before observing the evidence.

$P(B)$**is Marginal Probability**: Probability of Evidence.

4.2 Multilayer Perceptron

A Perceptron network with one or more hidden layers is called a Multilayer perceptron network. A multi perceptron network is also a feed-forward network. It consists of a single input layer, one or more hidden layers and a single output layer.The pictorial representation of multi-layer perceptron learning is as shown below

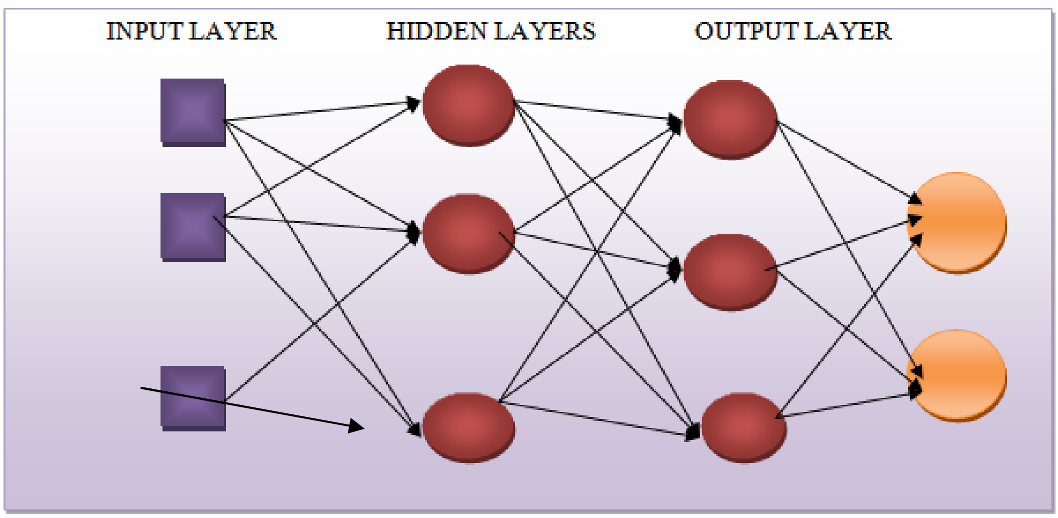


Figure 1. Multilayer Perceptron

MLP networks are used for supervised learning format. A typical learning algorithm for MLP networks is also called back propagation's algorithm.

4.3 Decision Tree

A decision tree is a flowchart-like tree structure, where each internal node represents a test happening an attribute, each branch represents an ending of the test, and class label is represented by each leaf node or terminal node. Given each tuple the attribute value of the tuple are tested next to the decision tree. A path is traced beginning the root to a leaf node which holds the class prediction used for the tuple. It is simple to convert decision trees into classification rules. Decision tree learning uses a decision tree because a predictive model which maps observations on an item to conclusions about the item's object value. It is single of the predictive modeling approaches utilize in statistics, data mining and machine learning. Tree models where the object variable can take a finite set of value are called classification trees, inside this tree structure, leaves correspond to class labels and branches represent conjunction of features that lead to individual's class labels. Decision tree can be constructed moderately quick compare to other methods of classification.

4.4 Random Forest

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of *combining multiple classifiers to solve a complex problem and to improve the performance of the model.* **Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. **The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.** The below diagram explains the working of the Random Forest algorithm:

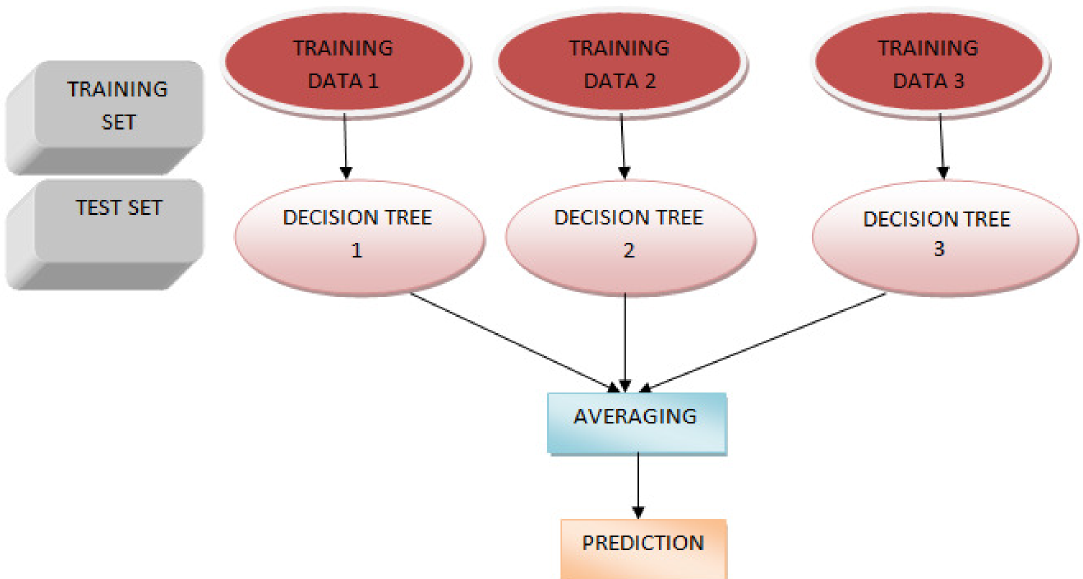


Figure 2. Random Forest

4.5 Instance Based Learner

Instance-based learning is the systems that learn the training examples by heart and then generalizes to new instances based on some similarity measure. It is called instance-based because it builds the hypotheses from the training instances. It is also known as **memory-based learning** or **lazy-learning.** The time complexity of this algorithm depends upon the size of training data. The worst-case time complexity of this algorithm is **O (n)**, where n is the number of training instances.

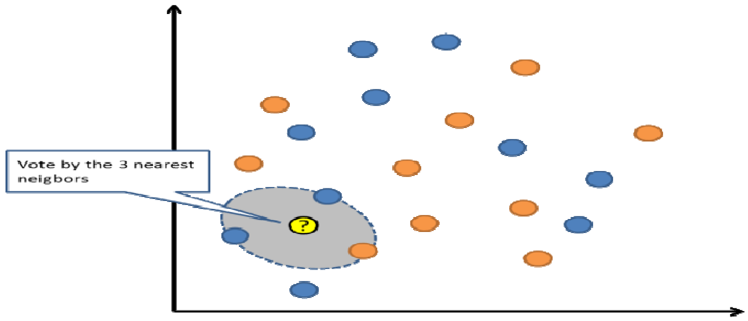


Figure 3. Instant Based Learner

5. RESULT AND DISCUSSION

In the following figure 4, Visualized seasonal rainfall database and prediction using various Methods of ML Algorithms of NORTH EAST STATES (NES) in India.

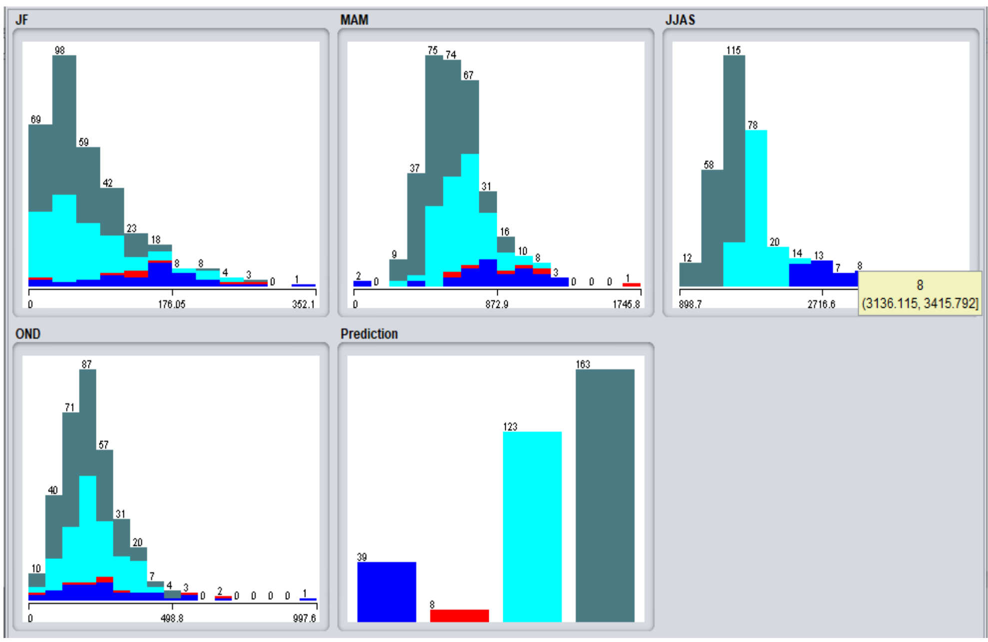


Figure 4.Seasonal Variationof Years

MODELS	CORRECTLY CLASSIFIED	INCORRECTLY CLASSIFIED	KAPPA STATISTIC
NAÏVE BAYES	95%	5.%	0.916
MULTILAYER PERCEPTION	98%	2.0%	0.965
DECISION TREE	100%	0%	1
RANDOM FOREST	100%	0%	1
INSTANCE BASED LEARNER	100%	0%	1

Table 1.Accuracy and Kappa statistic of Various Models

The above table1 show that classification accuracy and Kappa Statistic based different ML Algorithm models. For Naïve Bayes' and Multilayer Perception, the Kappa Statistic is 0.916 and 0.9654 and rest of the ML Algorithms shows the result as 1.The Kappa Statistic >0.80 is considered as perfect models.The Model of ML algorithms is visualized in the following figure 5. The Kappa Statistic and correctly classified instances are very close 95 percent and above and incorrectly classified instance less than 5 percent.

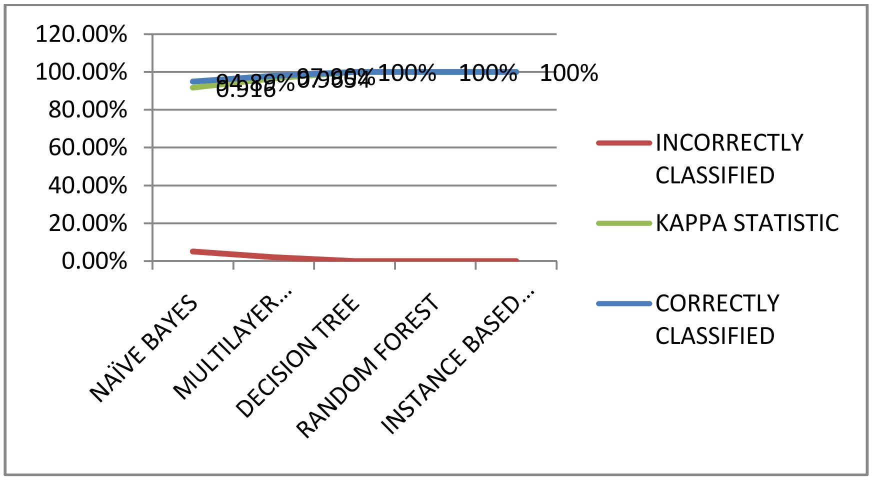


Figure 5.Visualization of Accuracy and Kappa Statistic

Models	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Root Relative Squared Error
Naïve Bayes	0.058	0.1427	18.9777%	36.5405%
Multilayer Perception	0.0158	0.0769	5.1748%	19.6972%
Decision Tree	0.000	0.000	0.000%	0.000%
Random Forest	0.0096	0.0371	3.1487%	9.5152%
Instance Based Learner	0.0045	0.0051	1.4562%	1.3165%

Table 2.Summary Statistics of Various Models

The Summary statistics of Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Relative Absolute Error (RAE) and Root Relative Squared Error (RRSE) other statistics of various models were discussed in this table2.This result is shows that best model for ML Algorithm.

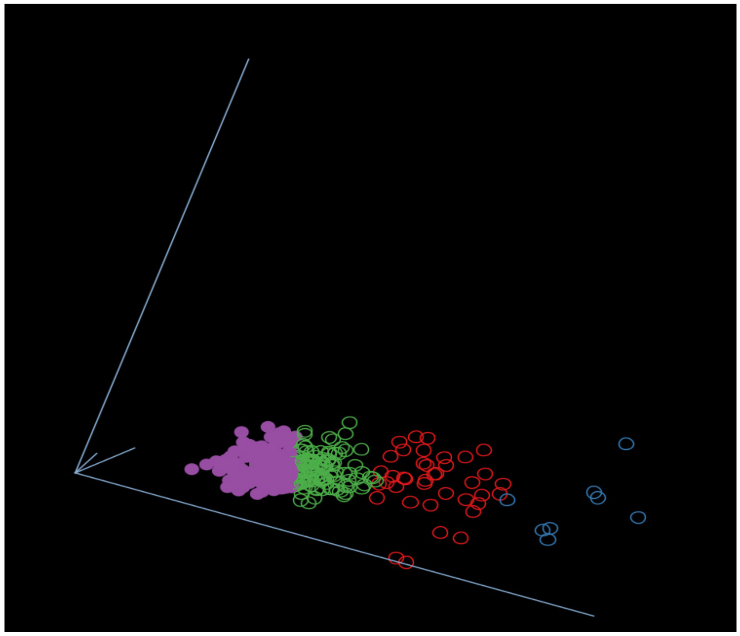


Figure 6. Four Seasons of NES Rainfall

The four seasons of NES rainfall database were classified and labeled as Very High Rainfall, High Rainfall, Moderate Rainfall and Low Rainfall. And this various seasons were visualized using Waikato Environment for Knowledge Analysis, (WEKA) in figure 6.

6. CONCLUSION

This research paper attempts to identify the rainfall database prediction, accuracy, classification and visualization in the North East States (NES) namely Arunachal Pradesh, Assam, Meghalaya, Manipur, Tripura, Mizoram and Nagaland (Sister of Seven States) in India. This research paper analyze seasonal database of NES using machine learning algorithms. Indian Climates is split into four categories and are named as winter (January to February), summer (March to May), monsoon (June to September) and post monsoon (October to December). Five machine learning algorithms achieved 95 percent and above classification accuracy and remaining less than 5 percent is misclassified due to large volume of database or outliers. Kappa Statistics of Naïve Bayes and Multilayer Perceptron (MLP) is 0.91 and 0.96 and rest of the machine learning algorithms J48, Random Forest and Instance Based Learner (IBk) results shows 1. The Kappa Statistics >0.80 is almost perfect classification. The summary statistics of various machine algorithms were discussed and their results showed that best model of ML algorithms. Finally the four seasons of NES rainfall database were classified and labeled as Very High Rainfall, High Rainfall, Moderate Rainfall and Low rainfall.

REFERENCES

1. Nikhilkumar B Shardoor, Dr. MandapatiVenkateshwarRao, Yashwant Mane, PushpaChettiar, VeenaKumari, KamleshArya (2019)” Journal of Emerging Technologies and Innovative Research (JETIR) www.jetir.org 72 Rainfall Prediction and Forecasting using Time Series Analysis” Journal of Emerging Technologies and Innovative Research (JETIR) Volume 6, Issue 4 www.jetir.org (ISSN-2349-5162).
2. Sara khan , MohdMuqeem , NashraJaved, (2016) “A Critical Review of Data Mining Techniques in Weather Forecasting” International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 4,
3. Md. MostafizurRahman and M. SayedurRahman, (2019), “Performance of different data mining methods for predicting rainfall of Rajshahi district, Bangladesh” 7th Int. Conf. on Data Science & SDGs EM – 041, pp 665-669.
4. Dr. MohdSadiq Ali Khan, (2021)” Indian Summer Monsoon Pattern Of Recoveries By Data Mining” International Journal Of Advance Scientific Research & ET, Volume 4, Issue 12, January 2021, ISSN (Online) 2456-0774.
5. Testud J, Chong M. Threedimensional wind field analysis from dual-Doppler radar data. Part I: Filtering, interpolating and differentiating the raw data. Journal of Applied Meteorology and Climatology. 1983 Jul 1; 22(7):1204-1215.
6. Nabila Wardah Zamani and Siti Shaliza Mohd Khairi, (2018) “A Comparative Study of Data Mining techniques for Rainfall Prediction in Subang” AIP Conference Proceedings, Volume 2013, Issue1, https://doi.org/10.1063/1.5054241.

Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
DOI No. : 10.5958/2394-9309.2022.00041.5
ISSN (Print) : 0975-7139
ISSN (Online) : 2394-9309

DESIGNING OF SINGLE SAMPLING PLANS FOR THE SUPPLY CHAIN CONTRACTS - A SIMULATION STUDY WITH QUALITY COSTS OPTIMIZATION

S. Ravisankar* & J. Sinthiya**

*Department of Statistics, Government Arts College, Coimbatore, India
** Department of Mathematics, Sri Ramakrishna College of Arts & Science, Coimbatore, India
E-mail id: sinthiya@srcas.ac.in

ABSTRACT

Supply Chain Contract (SCC) is a relatively new method which introduces a new mechanism for managing uncertainties in projects. In this paper, an attempt is made to design asimulation model to determine the optimal single sampling plan for supply chain contract that minimizes the producer's total cost and the consumer's total cost by satisfying both the producer's and the consumer's quality requirements. Numerical examples are also given so as to select optimal single sampling plan in order to obtain optimum lot size.

Key Words : Supply Chain Contract, Acceptance Sampling, Contracts, Supply Chain Management

1. INTRODUCTION

Acceptance sampling has become important field of Statistical Quality Control and it was popularized by Dodge and Romig. Acceptance sampling inspection can be carried outat the beginning of the production process while receiving the raw material from a supplier, the component orthe parts orthe products at the end of the production. It is usually used when testing is destructive; the cost of 100% inspection is very high; and 100% inspection takes too long.

According to Dodge (1969), the general areas of acceptance sampling are: (i) Lot-by-lot sampling by the method of attributes (ii) Lot-by-lot sampling by the method of variables (iii) Continuous sampling of a flow of units by the methods of attributes (iv)Special purpose plans.

In designing a sampling plan, one has to accomplish a number of different purposes. According to Hamaker (1960), the important factors are

- To strike a proper balance between the consumer's requirements, the producer's capabilities and inspection capacity,
- To separate bad lots from good ones,
- Simplicity of procedures and administration,
- Economy in number of observations,
- To reduce the risk of wrong decisions with increasing lot size,
- To use accumulated sample data as a valuable source of information,

- To exert pressure on the producer or supplier when the quality of the lots received is unreliable or not up to the standard and
- To reduce sampling when the quality is reliable and satisfactory.

1.1 Supply Chain Contract

A Supply chain contract isthe agreement between the buyer and the supplier on the issues like pricing and discounts,minimum and maximum purchasing quantities, etc.

The following three types of contracts are mainly used in supply chain. They are

- Fixed Price Contract
- Cost-Plus Contract
- Time and Material Contract

Fixed Price Contract (FPC)

A fixed price contract is also known as Firm Fixed Price Contract. The fixed price contract is the mostly used by all the suppliers. The buyer also accepts it because once the price is fixed. They know that there will be no change in the price and no discounts. In this type of contract, price, material, time and discountswere fixed. Thebuyer need not to pay extra price or time. The FPC is the most favored contract type for buying organizations to buy when there are no discounts and to pay actual cost.

Cost-Plus Contract

A Cost-Plus Contract is an agreement to reimburse from a company for expenses incurred plus a specific amount of profit, usually stated as a percentage of the contract's full price. These type of contracts are primarily used in construction where the buyer assumes some of the risks but also provides the degree of flexibility to the contractor. As this contract covers all the expenses related to the project, there are no surprises. The contractors must provide proof of all related expenses, including direct and indirect costs. The main advantage of cost plus contract is higher quality of material, since, the contractor has incentive to use the best labour, material and the less chance of having the project overbid.

Time and Material Contract

A Time and Material (T&M) contract involves both seller and buyer agreeing on pre-determined unit rates for labour and material.This type of contract is used when it is impossible to get an accurate estimate of the total project cost. The schedule cannot be defined or changes are likely to be made during construction. The main advantage of this contract is flexibility to adjust requirements, shift directions and replace features. This contract is to avoid fixed-price bidding process which helps to save time and also to motivate the work more efficiently.

The notations used in this study are:

c : Acceptance Number

C : Total Cost

n : Sample Size

N : Optimum Lot Size

$P(q)$: Production Quantity

SD(q) : Standard deviation of Quantity
T(q): Target Quantity

2. REVIEW OF LITERATURE

The various research articles related to Acceptance Sampling and Supply Chain (SC) in which simulation techniques have been reviewed.

Radhakrishnan and Sekkizhar (2007) have applied intervened random effect Poisson distribution in a business process control and constructed sampling plans. Radhakrishnan and Ravisankar (2009) presented a procedure for constructing single sampling plans for three class attribute plans using AQL as the quality standard. These plans were also compared with two class attribute plans.

Jamkhaneh, E. B., Sadeghpour-Gildeh, B., &Yari, G. (2010) proposed a method for designing acceptance single sampling plans with fuzzy quality characteristic with using fuzzy Poisson distribution. These plans are well defined since if the fraction of defective items is crisp they reduce to classical plans.

Starbird (2001) investigated the effect of rewards, penalties and inspection policies on the behaviour of the expected cost minimizing to the supplier. Further pre-discussed about the reward and/or penalty that motivates a supplier to deliver the buyer's target quality depending upon the inspection policy. Starbird (2003) also developed a mathematical model for a simplified supply chain in which conformance quality is one of the supplier's decision variables. Also both the supplier and the customer are trying to minimize the expected annual cost. The author also found that the buyers who are using coordinated replenishment may be trading higher quality products for lower cost.

Lie-Fern Hsu andJia-Tzer Hsu (2012) have developed an economic model to determine the optimal sampling plan in a two-stage supply chain that minimizes the producer's and the customer's quality and risk requirements. They found the products inspection, internal failure and externalfailure costs which are having an effect on the optimal sampling plan. Cao and Zhang (2011) achieved greater collaborative advantages with their supply chain partners in the past few decades. These supply chain collaborative advantages have a bottom-line influence on firm performance.

Foster (2008) proved that the increasing importance given to supply chain management which has been developed for operations management. The author also defined the Supply Chain Quality Management (SCQM) as a system-based approach to the performance improvement that leverages opportunities created by upstream and downstream linkages with the supplierand the customer.

3. OBJECTIVE OF THE STUDY

The main objective of this study is to design single sampling plans in supply chain contracts that minimizesthe producer's total cost and consumer's total cost by satisfying both producer's and consumer's quality requirements.

4. RESEARCH METHODOLOGY

In this study, a simulation software (GoldSim14.0) is used to design the simulation model for single sampling plan in supply chain contract in order to find out the economic order quantity. The inventory cost, the shortage cost, the rewards and the penalties are taken into account to minimize the total cost using the simulation model. The simulation is an important tool that provides a way in which alternative designs, plans and policies can be evaluated without having to experiment in a real time system, which may be costly, time-consuming or simply impractical to do.

GoldSim(14.0) is a dynamic, probabilistic simulation software developed by GoldSim technology. It is a general purpose softwarefor simulation framework. It is not specialized to a particular type of problem. The GoldSimsoftware has the powerful discrete event simulation capabilities. Itwas designed primarily to model systems exhibiting both continuous and discrete changes.

5. SINGLE SAMPLING PLAN (SSP) AND ITS BASIC MEASURES

The single sampling plan is a decision rule to accept or reject a lot based on the results of one random sample from the lot.

5.1 Procedure for Single Sampling Plan

The operatingprocedure for implementing SSP to arrive at a decision about the lot is described in the following steps.

i) Draw a random sample of size (n) from the lot of size (N) received from the supplier.

ii) Count the number of defective units (d) in the sample.

iii) If $d \leq c$, the acceptance number, the lot is accepted.

iv) If $d > c$, the lot is rejected and go for 100% inspection.

5.2Economic Design of Acceptance Sampling Plan

The probability of acceptance for the SSP is given by

$$P_a(p) = \sum_{k=0}^{c} \frac{e^{-np}(np)^k}{k!} \qquad \ldots\ldots (5.1)$$

where ‘p’ is the proportion defective.

The script used in the GoldSimsoftware to calculate probability of acceptance $P_a(p)$ for various ‘p’ values is given in figure-1.

```
Script
Statement List
Define: prob = 0.0
FOR (k = 0; ~k <= c1; k = ~k + 1)
  Define: fact = 1
  FOR (i = 1; ~i <= ~k; i = ~i + 1)
    fact = ~fact*~i
  END FOR
  prob = (n*p)^~k/~fact+~prob
END FOR
Result = ~prob*exp(-n*p)
```

Figure-1: Script used in GoldSim

5.3Average Outgoing Quality (AOQ)

The AOQ is the expected proportion of defectives that the plan will allow to pass. It is assumed that all the defective items in the lot will be replaced with good items.

The AOQ of the SSP is given by

$$AOQ = \frac{p(N-n)P_a(\text{p})}{N} \qquad ....(5.2)$$

6. DESIGNING A SIMULATION MODEL FOR INTEGRATING SSP WITH SCC

In this study, the simulation model presented in Figure-2 for supply chain contractsis developed using theGoldSim Software.

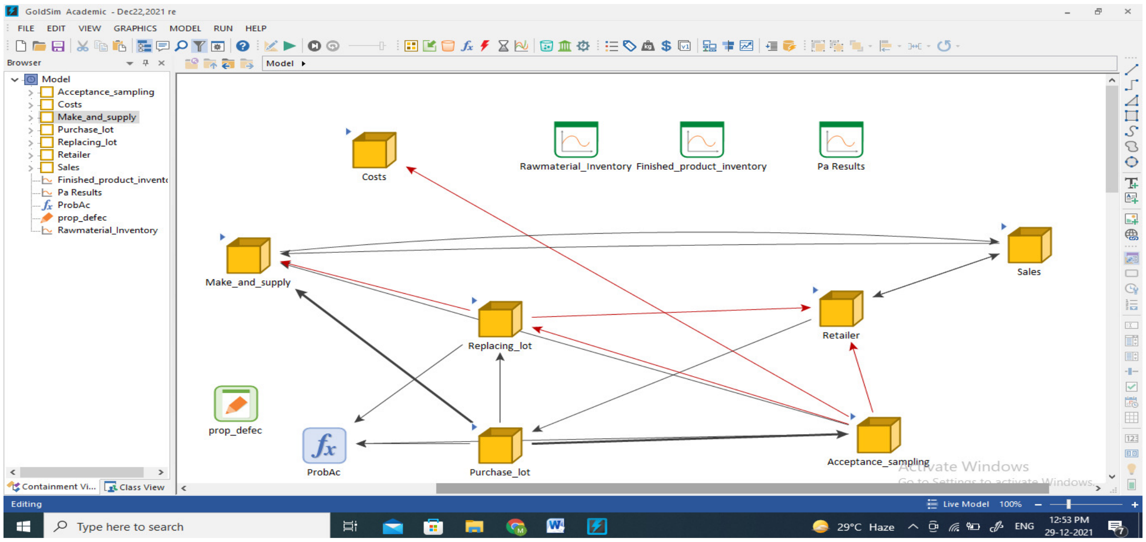


Figure 2.Simulation Model for Integrating SSP with SCC

The process of supply chain contractsis shown inthe figure 2. A lot of the raw material is being sent to the sampling inspection after which the lot is sent to the production process.

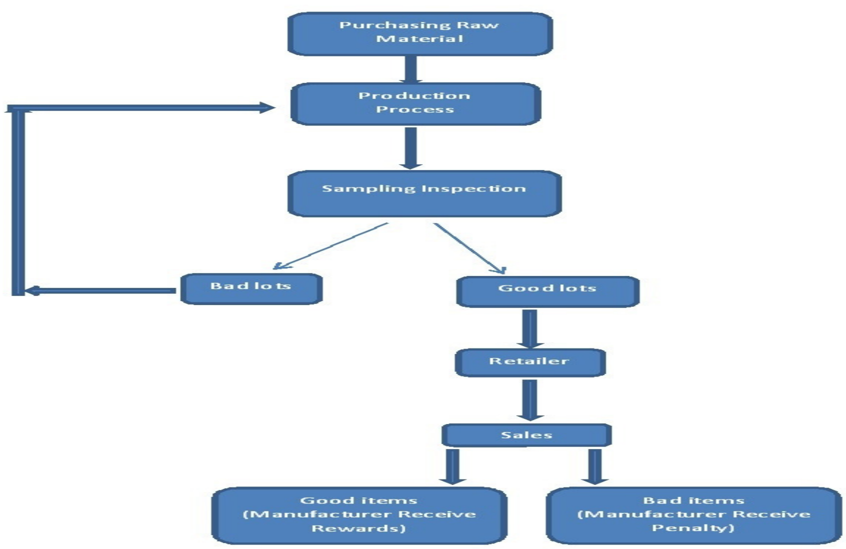


Figure- 3: Integrating SSP with SCC

6. CONSTRUCTION OF SSP (N,C) AND CALCULATING (NC)

In this section, for various combination of T(q), SD(q) & P(q) and fixed reward Rs.0.5, penalty Rs.250, the SSP (n,c) with optimum lot size (N) and the total cost (C) has been constructed for various fixed demand values 200, 500 &700 and presented in Table-1, Table-2 and Table-3 respectively.

Table-1: SSP (n,c) with (N, C) for Demand=200, Reward = Rs. 0.5 andPenalty =Rs. 250

T(q)	SD(q)	P(q)	n	c	N	C
100	30	120	121	8	858	81,43,980
100	30	150	97	4	300	96,41,720
120	30	150	91	3	858	81,44,310
120	20	150	146	3	857	81,35,790
120	40	150	72	6	858	81,48,550
120	30	120	87	7	856	81,15,600
120	20	120	47	7	856	81,16,710
130	20	150	101	3	856	81,16,090
130	30	150	120	5	856	81,15,620
130	40	150	101	4	856	81,15,610

Example-1:

As the manufacturer has to minimize the total cost,he receives both the reward for better quality and the penalty for the failure.It is assumed that the reward and penalty are Rs.0.5/unit and as Rs. 250/unit respectively. Suppose the manufacturer requires the target quantityT(q)=100, standard deviation of quantitySD(q) =30 and production quantityP(q) =120, then from the Table-1;

For the lot size (N) 858, the manufacturer will obtainthe SSP(121,8) and Rs. 81,43,980 as the total cost(C).

Table 2: SSP (n,c) with N and C for Demand=500, Reward = Rs. 0.5 andPenalty =Rs. 250

T(q)	SD(q)	P(q)	n	c	N	C
100	30	120	151	2	2166	1,99,55,800
100	30	130	176	5	2000	1,90,63,900
120	30	150	141	5	1999	1,89,57,700
120	20	150	184	5	1999	1,92,28,000
120	40	150	129	5	2000	1,90,63,900
120	30	120	33	2	2096	1,99,88,100
120	20	120	98	9	1999	1,89,57,100
130	20	150	124	6	1999	1,89,69,200
130	30	150	42	2	2000	1,91,61,500
130	40	150	184	4	2096	1,99,91,400

Example-2:

In a contract, the demand of an item is 500 units per day, the production quantity= 120, target quantity =100 and standard deviation=30 and then from the Table 2, the manufacturer get the SSP(151,2) along with the optimal lot size 2166 units and the total cost Rs. 1,99,55,800.

Table 3: SSP (n,c) with N & C for Demand=700, Reward = Rs. 0.5 andPenalty =Rs. 250

T(q)	SD(q)	P(q)	n	c	N	C
100	40	120	130	2	2799	2,91,63,900
100	30	130	178	7	2799	2,91,64,200
120	30	150	136	1	3033	3,05,54,800
120	20	150	163	2	3078	3,09,38,900
120	40	150	147	8	2935	3,06,47,800
120	30	120	108	3	3033	3,05,50,600
120	20	120	135	7	2799	2,91,63,700
130	20	150	151	4	2800	2,93,70,200
130	30	150	154	3	3033	3,05,54,700
130	40	150	138	3	2799	2,91,63,900

Example-3:

Suppose the manufacturer requires the target quantityT(q)=100, standard deviation of quantity =40, production quantity=120, then from the Table 1, he can get the SSP(130,2) along with lot size (N) =2799 units and total cost (C) = Rs. 2,91,63,900.

CONCLUSION

In this study, SSP is used to monitor the quality of the raw material and the components delivered by suppliers or the finished product. It is concluded that the risk of rejection motivates the supplier's to improve the quality of the production. Customers have an economic justification for the use of acceptance sampling. Numerical examples are provided to show the effectiveness of producers and consumers justification.The supplier and buyer coordinate replenishment may be focused on further study. The study may also be extended for other sampling inspection plans.

REFERENCES:

1. *Dodge, H. F. (1969). Notes on the evolution of Acceptance Sampling Plans, Part I. Journal of Quality Technology, 1(sup1), 1-12.*
2. *Hamaker, H. C. (1960). Attribute Sampling in Operation. Bulletin of the International Statistical Institute, 37(2), 265-281.*

3. *Radhakrishnan, R., &Sekkizhar, J. (2007). Construction of Sampling Plans using Intervened Random Effect Poisson Distribution. International Journal, 2(1-2), 88-96.*
4. *Jamkhaneh, E. ,Sadeghpour-Gildeh, B., &Yari, G. (2010). Acceptance single sampling plan by using of Poisson distribution. Journal of Mathematics and Computer Science, 1(1), 6-13.*
5. *Starbird, S. A. (2001). Penalties, rewards, and inspection: provisions for quality in supply chain contracts. Journal of the Operational Research Society, 52(1), 109-115.*
6. *Starbird, S. A. (2003). Effect of coordinated replenishment policies on quality. Journal of the Operational Research Society, 54(1), 32-39.*
7. *Hsu, Lie-Fern, and Jia-Tzer Hsu. (2012). Economic design of acceptance sampling plans in a two-stage supply chain. Advances in Decision Sciences, 2012.*
8. *Cao, M., & Zhang, Q. (2011). Supply chain collaboration: Impact on collaborative advantage and firm performance. Journal of operations management, 29(3), 163-180.*
9. *Jamkhaneh, E. B., Sadeghpour-Gildeh, B., &Yari, G. (2010). Acceptance single sampling plan by using of Poisson distribution. Journal of Mathematics and Computer Science, 1(1), 6-13.*
10. *Dodge, H. F. (1965). Evaluation of a sampling inspection system having rules for switching between normal and tightened inspection. Technical Report 14.C.*
11. *https://www.goldsim.com/Web/Customers/Downloads/*

DOI No. : 10.5958/2394-9309.2022.00042.7
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

K-LEHMER THREE MEAN LABELING OF SUBDIVISION OF SOME GRAPHS

M J Abisha* ,K. Rubin Mary**

*Research Scholar, Reg No:19113232092003,Department of Mathematics, St Jude's College, Thoothoor-629176,TamilNadu,India.

**Assistant professor,Department of Mathematics, St Jude's College, Thoothoor-629176, Tamil Nadu, India.

Affliated to ManonmaniamSundaranarUniversity,Abishekapatti, Tirunelveli-627012,TamilNadu,India.

E-mail : abisharejeesh@gmail.com, rubinjude@yahoo.com

ABSTRACT

A graph $G = (V, E)$ with r vertices and s edges is called a k- Lehmer-3 mean graph if it is possible to label the vertices $v \in V$ with distinct labels $h(x)$ from $k, k+1, k+2, \ldots, k+s$ in such a way that when each edge $e = xy$ is labeled with $h(e) = \left\lceil \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rceil$ (or) $\left\lfloor \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rfloor$ then the edge labels are distinct.In this case "h" is called k-Lehmer-3 mean labeling of G. In this paper we investigate k-Lehmer three mean labeling of some disconnected graphs.

Keyword : Lehmer three mean labeling, k- Lehmer three mean labeling, Subdivision of graph.

1. INTRODUCTION:

Graphs described here is simple,undirected and connected graphs.Let V(G) and E(G) be stated as the vertex and edge set of graph G.We refer Gallian for more comprehensive survey [1].We follow Harrary for some standard words,expressions and symbols[2],The concept and notation of mean labeling was first introduced by S somasundaram and R Ponraj[3].S Somasundaram,S S Sandya and T Pavithra introduced the concept of Lehmer three mean graphs [4]. Here we investigating some standard graphs in K-Lehmer three mean labeling of Subdivision of some graphs.

Definition1.1

Let G be a(r, s) graph.A function$h$is called Lehmer three mean labeling of graph G, if it is possible to label the vertices $v \in V$ with distinct labels $h(x)$ from $1,2,3, \ldots \ldots \ldots, s+1$ in such a way that when each edge $e = xy$ is labeled with $h(e) = \left\lceil \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rceil$ (or) $\left\lfloor \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rfloor$ then the edge labels are distinct.In this case "h" is called Lehmer-3 mean labeling of G.

Definition1.2

Let G be a (r, s) graph.A function h is called k-Lehmer three mean labeling of graph G,if it is possible to label the vertices $v \in V$ with distinct labels $h(x)$ from $k, k+1, k+2, \ldots, k+s$in such a way that each edge $e = xy$ is labeled with$h(e) = \left\lceil \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rceil$ (or) $\left\lfloor \frac{h(x)^3+h(y)^3}{h(x)^2+h(y)^2} \right\rfloor$ then the edge labels are distinct.A graph which admits a k-Lehmer three mean labeling is called k-Lehmer three mean graph.

2. MAIN RESULTS

Theorem 2.1

The graph $S(P_m \odot k_1)$ is a k-Lehmer three mean graphs.

Proof:

We subdivide $P_m \odot k_1$ in three different cases.

Case (i)

Subdividing the vertices u_j and u_{j+1} by a new vertex named w_j.

Let $V(G)\} = \{u_j, v_j \;\; 1 \le j \le m \; \& w_j; 1 \le j \le m-1\}$ and

$E\{G)\} = \{u_j v_j;\; 1 \le j \le m\} \cup \{u_j w_j;\; w_j u_{j+1}; 1 \le j \le m-1\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = k + (3j-3) \qquad 1 \le j \le m$$
$$h(v_j) = k + (3j-2) \qquad 1 \le j \le m$$
$$h(w_j) = k + (3j-1) \qquad 1 \le j \le m$$

Then the induced edge labels are

$$h^*(u_j v_j) = k + (3j-3) \qquad 1 \le j \le m$$
$$h^*(u_j w_j) = k + (3j-2) \qquad 1 \le j \le m-1$$
$$h^*(w_j u_{j+1}) = k + (3j-1) \qquad 1 \le j \le m-1$$

Case (ii)

Subdividing the vertices u_j and v_j by a new vertex named w_j.

Let $V(G)\} = \{u_j, v_j, w_j \; 1 \le j \le m, \}$ and

$E\{G)\} = \{u_j w_j; w_j v_j;\; 1 \le j \le m\} \cup \{u_j u_{j+1}; 1 \le j \le m-1\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = k + (3j-3) \qquad 1 \le j \le m$$
$$h(v_j) = k + (3j-2) \qquad 1 \le j \le m$$
$$h(w_j) = k + (3j-1) \qquad 1 \le j \le m$$

Then the induced edge labels are

$$h^*(u_j w_j) = k + (3j-3) \qquad 1 \le j \le m$$
$$h^*(w_j v_j) = k + (3j-2) \qquad 1 \le j \le m-1$$
$$h^*(u_j u_{j+1}) = k + (3j-1) \qquad 1 \le j \le m-1$$

Case (iii)

Subdividing each edge by a new vertex.

Let $V(G)\} = \{u_j, v_j, w_j \; 1 \le j \le m \;\; \& x_j \; 1 \le j \le m-1, \}$

$E\{G)\} = \{u_j x_j; x_j u_{j+1};\; 1 \le j \le m-1\} \cup \{u_j w_j; w_j v_j 1 \le j \le m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = k + (4j-4) \qquad 1 \le j \le m$$
$$h(v_j) = k + (4j-2) \qquad 1 \le j \le m$$
$$h(w_j) = k + (4j-3) \qquad 1 \le j \le m$$

$$h(x_j) = k + (4j - 1) \qquad 1 \le j \le m - 1$$

Then the induced edge labels are

$$h^*(u_j w_j) = k + (4j - 4) \qquad 1 \le j \le m$$
$$h^*(w_j v_j) = k + (4j - 3) \qquad 1 \le j \le m$$
$$h^*(u_j x_j) = k + (4j - 2) \qquad 1 \le j \le m - 1$$
$$h^*(x_j u_{j+1}) = k + (4j - 1) \qquad 1 \le j \le m - 1$$

Thus S($P_m \odot k_1$) is a k-Lehmer three mean graphs.

Theorem 2.2

The Subdivision of a triangular snake $T_m$is a k-Lehmer three mean graphs for K$\ge$ 6.

Proof:

We subdivide T_m in three different cases.

Case (i)

Subdividing the vertices u_j and u_{j+1} by a new vertex named w_j.

Let $V(G)\} = \{u_j;\ 1 \le j \le m - 1\ \&\ v_j w_j\ 1 \le j \le m\}$ and

$E\{G)\} = \{u_j w_j;\ w_j u_{j+1}, u_j v_j, v_j u_{j+1};\ 1 \le j \le m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s-1\}$ by

$$h(u_1) = k$$
$$h(u_j) = k + (4j - 5) \qquad 2 \le j \le m$$
$$h(v_j) = k + (4j - 3) \qquad 1 \le j \le m\text{-}1$$
$$h(w_j) = k + (4j - 2) \qquad 1 \le j \le m\text{-}1$$

Then the induced edge labels are

$$h^*(u_j v_j) = k + (4j - 4) \qquad 1 \le j \le m - 1$$
$$h^*(u_j w_j) = k + (4j - 3) \qquad 1 \le j \le m - 1$$
$$h^*(w_j u_{j+1}) = k + (4j - 1) \qquad 1 \le j \le m - 1$$
$$h^*(v_j u_{j+1}) = k + (4j - 2) \qquad 1 \le j \le m - 1$$

Case(ii)

Subdividing two sides of triangle.

Let $V(G)\} = \{u_j; 1 \le j \le m + 1\ \& v_j, w_j, x_j 1 \le j \le m,\}$ and

$E\{G)\} = \{u_j u_{j+1}; u_j w_j; w_j v_j; v_j x_{j+1}; x_j u_{j+1}\ 1 \le j \le m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = k + (5j - 5) \qquad 1 \le j \le m$$
$$h(v_j) = k + (5j - 3) \qquad 1 \le j \le m - 1$$
$$h(w_j) = k + (5j - 4) \qquad 1 \le j \le m - 1$$
$$h(x_j) = k + (5j - 2) \qquad 1 \le j \le m - 1$$

Then the induced edge labels are

$$h^*(u_j w_j) = k + (5j - 3) \qquad 1 \le j \le m - 1$$

$$h^*(w_j v_j) = k + (5j-4) \qquad 1 \le j \le m-1$$
$$h^*(u_j u_{j+1}) = k + (5j-2) \qquad 1 \le j \le m-1$$
$$h^*(v_j x_j) = k + (5j-3) \qquad 1 \le j \le m-1$$
$$h^*(x_j u_{j+1}) = k + (5j-1) \qquad 1 \le j \le m-1$$

Case (iii)

Subdividing each edge by a new vertex.

Let $V(G)\} = \{u_j; 1 \le j \le m+1 \,\&v_j, w_j, x_j, y_j 1 \le j \le m,\}$ and

$E\{G)\} = \{y_j u_{j+1}; u_j y_j; u_j w_j; w_j v_j; v_j x_j; x_j u_{j+1}\ 1 \le j \le m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s-1\}$ by

$$h(u_1) = k$$
$$h(u_j) = k + (6j-7) \qquad 2 \le j \le m$$
$$h(v_j) = k + (6j-4) \qquad 1 \le j \le m-1$$
$$h(w_j) = k + (6j-5) \qquad 1 \le j \le m-1$$
$$h(x_j) = k + (6j-2) \qquad 1 \le j \le m\text{-}1$$
$$h(y_j) = k + (6j-3) \qquad 1 \le j \le m\text{-}1$$

Then the induced edge labels are

$$h^*(u_j y_j) = k + (6j-4) \qquad 1 \le j \le m-1$$
$$h^*(u_j w_j) = k + (6j-6) \qquad 1 \le j \le m-1$$
$$h^*(w_j v_j) = k + (6j-5) \qquad 1 \le j \le m-1$$
$$h^*(u_j u_{j+1}) = k + (5j-2) \qquad 1 \le j \le m-1$$
$$h^*(v_j x_j) = k + (6j-3) \qquad 1 \le j \le m-1$$
$$h^*(x_j u_{j+1}) = k + (6j-1) \qquad 1 \le j \le m-1$$

Thus S(T_m) is a k-Lehmer three mean graphs.

Theorem 2.3

The graph S($P_m \odot k_{1,2}$) is a k-Lehmer three mean graphs

Proof:

We subdivide $P_m \odot k_{1,2}$ in three different cases.

Case (i)

Subdividing the vertices u_j and u_{j+1} by a new vertex named x_j.

Let $V(G)\} = \{x_j;\ 1 \le j \le m-1 \,\&u_j, v_j, w_j\ 1 \le j \le m\}$ and

$E\{G)\} = \{u_j x_j;\ x_j u_{j+1};\ 1 \le j \le m-1\} \cup \{u_j v_j;\ u_j w_j;\ 1 \le j \le m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = k + (4j-4) \qquad 1 \le j \le m$$
$$h(v_j) = k + (4j-3) \qquad 1 \le j \le m$$

$$h(w_j) = k + (4j - 2) \qquad 1 \leq j \leq m$$
$$h(x_j) = k + (4j - 1) \qquad 1 \leq j \leq m$$

Then the induced edge labels are

$$h^*(u_j x_j) = k + (4j - 2) \qquad 1 \leq j \leq m - 1$$
$$h^*(u_j w_j) = k + (4j - 3) \qquad 1 \leq j \leq m$$
$$h^*(x_j u_{j+1}) = k + (4j - 1) \qquad 1 \leq j \leq m - 1$$
$$h^*(u_j v_j) = k + (4j - 4) \qquad 1 \leq j \leq m$$

Case(ii)

Subdividing the vertices u_j and u_{j+1} by a new vertex named w_j and Subdividing the vertices u_j and y_j by a new vertex named w_j.

Let $V(G)\} = \{u_j, v_j, w_j, x_j, y_j\ 1 \leq j \leq m\}$ and

$E\{G)\} = \{u_j u_{j+1};\ 1 \leq j \leq m - 1\} \cup \{u_j w_j, w_j v_j, u_j x_j, x_j y_j;\ 1 \leq j \leq m\}$

Case (a) When m is odd

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s+1\}$ by

$$h(u_j) = \begin{cases} k + (5j - 5) & for\ i = 1,3,5, \dots, m \\ k + (5j - 6) & for\ i = 2,4,6, \dots, m - 1 \end{cases}$$
$$h(v_j) = k + (5j - 3) \qquad 1 \leq j \leq m$$
$$h(w_j) = k + (5j - 4) \qquad 1 \leq j \leq m$$
$$h(x_j) = k + (5j - 2) \qquad 1 \leq j \leq m$$
$$h(y_j) = \begin{cases} k + 5j & for\ i = 1,3,5, \dots, m \\ k + (5j - 1) & for\ i = 2,4,6, \dots, m - 1 \end{cases}$$

Case (b) When m is even

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s\}$ by

$$h(u_j) = \begin{cases} k + (5j - 5) & for\ i = 1,3,5, \dots, m - 1 \\ k + (5j - 6) & for\ i = 2,4,6, \dots, m \end{cases}$$
$$h(v_j) = k + (5j - 3) \qquad 1 \leq j \leq m$$
$$h(w_j) = k + (5j - 4) \qquad 1 \leq j \leq m$$
$$h(x_j) = k + (5j - 2) \qquad 1 \leq j \leq m$$
$$h(y_j) = \begin{cases} k + 5j & for\ i = 1,3,5, \dots, m - 1 \\ k + (5j - 1) & for\ i = 2,4,6, \dots, m \end{cases}$$

Then the induced edge labels are

$$h^*(u_j u_{j+1}) = k + (5j - 2) \qquad 1 \leq j \leq m - 1$$
$$h^*(u_j x_j) = k + (5j - 3) \qquad 1 \leq j \leq m$$
$$h^*(u_j w_j) = k + (5j - 5) \qquad 1 \leq j \leq m$$
$$h^*(w_j v_j) = k + (5j - 4) \qquad 1 \leq j \leq m$$
$$h^*(x_j y_j) = k + (5j - 2) \qquad 1 \leq j \leq m$$
$$h^*(u_j v_j) = k + (4j - 4) \qquad 1 \leq j \leq m$$

Case(iii)Subdividing each edge by a new vertex.

Let $V(G)\} = \{u_j, v_j, w_j, x_j, y_j 1 \leq j \leq m \ \& t_j; 1 \leq j \leq m-1\}$ and

$E\{G)\} = \{u_j t_j; t_j u_{j+1}; 1 \leq j \leq m-1\} \cup \{u_j w_j; w_j v_j; u_j x_j; x_j y_j \ 1 \leq j \leq m\}$

Define a function $h: V(G) \to \{k, k+1, k+2, \dots, k+s+1\}$ by

$$h(u_j) = k + (6j-6) \qquad 1 \leq j \leq m$$
$$h(v_j) = k + (6j-4) \qquad 1 \leq j \leq m$$
$$h(w_j) = k + (6j-5) \qquad 1 \leq j \leq m$$
$$h(x_j) = k + (6j-3) \qquad 1 \leq j \leq m$$
$$h(y_j) = k + (6j-1) \qquad 1 \leq j \leq m$$
$$h(t_j) = k + (6j-2) \qquad 1 \leq j \leq m$$

Then the induced edge labels are

$$h^*(u_j t_j) = k + (6j-3) \qquad 1 \leq j \leq m-1$$
$$h^*(t_j u_{j+1}) = k + (6j-1) \qquad 1 \leq j \leq m-1$$
$$h^*(w_j v_j) = k + (6j-5) \qquad 1 \leq j \leq m$$
$$h^*(u_j w_j) = k + (6j-6) \qquad 1 \leq j \leq m$$
$$h^*(u_j x_j) = k + (6j-4) \qquad 1 \leq j \leq m$$
$$h^*(x_j y_j) = k + (6j-2) \qquad 1 \leq j \leq m$$

Thus S($P_m \odot k_{1,2}$) is a k-Lehmer three mean graphs.

REFERENCE

1. *J.A. Gallian, "A Dynamic Survey of Graph Labeling", The Electronic Journal of Combinatories,(2019),DS6.*
2. *F.Harary , "Graph theory" , Narosa publication House reading ,New Delhi (1988).*
3. *S. Somasundaram and R .Ponraj, "Mean labeling of Graphs", National Academy of Science Letters,26(2003),210-213.*
4. *S. Somasundaram, S.S. Sandhya and T.S.Pavithra," Lehmer-3 mean Labeling of graphs" "International Journal of Mathematical Forum,12(17)(2017),819-825.*
5. *M.J.Abisha and K.Rubin Mary, "K-Lehmer three mean labeling of some graphs", Malaya Journal of Mathematik, Vol 8,No.3,1219-1221,2020.*

DOI No. : 10.5958/2394-9309.2022.00043.9
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

PERCEPTION OF ACCOMMODATION PROVIDERS IN KANCHIPURAM DISTRICT

Dr. A. Mahendrakumar*, Mr. P.Murugaian, Dr N. Kannan*****
Assistant Professor, Department of Commerce, St Thomas College of Arts and Science, Koyambedu -107.
Ph.D Scholar(Part time), Department of Statistics Annamalai University, Chidambaram.
Professor, Department of Statistics, Annamalai University, Chidambaram, Tamilnadu, India.

ABSTRACT

Accommodation is a sordid of tourism industry as it is a energetic and fundamental part of tourism supply. Tourists in their travel need location where they can rest and revive during their travel. As a result, commercial accommodations are in existence. Hotels are getting closely correlated and are an integral part of tourism. Accommodation providers respond to the tourists need and offer services to standard criteria and desire of tourists.Recent years, hotels in Kanchipuram district have encountered difficult economic times due to increasing customer demands and strong internal industry development competition. Domestic travelers stay in the hotels for more number of days compared to the foreign tourist. Since most of the foreign travelers are on a continuous movement and they have to cover other destinations as per their travel plan. In many of the hotels, there is lack of infrastructures like laundry, dry cleaning, security, swimming pools etc. The hospitality industry's main concern globally is to serve its customers' needs and desires, most of which are addressed through personal services. Hence, the hotel businesses that are able to provide quality services to its ever-demanding customers in a warm and efficient manner are those businesses which will be more likely to obtain a long-term competitive advantage over their rivals.in this research paper the researcher analyzed that perception of accommodation providers in Kanchipuram district.

Keyword s: Tourist, Hotels, Accommodation, Mean, Standard deviation, Chi-Square testetc.

1. INTRODUCTION

Hotels are an important component of the tourism industry. It contributes in the overall tourism experience through the standards of facilities and services offered by them. With the aim of providing contemporary standards of facilities and services available in the hotels, the Ministry of Tourism has formulated a voluntary scheme for classification of operational hotels which will be applicable to the following categories: Star Category Hotels: 5 Star Deluxe, 5 Star, 4 Star, 3 Star, 2 Star & 1 Star Heritage Category Hotels: Heritage Grand, Heritage Classic & Heritage Basic.

In the Temple State of Tamil Nadu, hospitality is a tradition that never fails to make the guests feel like being the most important persons on earth. They enjoy the architectural wonders of Dravidian culture, music and dances, wildlife and beaches, and the serenity of hill resorts in Tamil Nadu. All the major hotel chains of India have their hotels-up and running in the major cities especially Chennai".

Accommodation is the key element in the tourism product and an essential component of tourism. Accommodation is a core area of the tourist industry and plays a distinctive role in the development of this ever- expanding industry. The United Nations conference on International Travel and Tourism held in Rome in 1963 considered, in particular, the issues relating to means of accommodation.

HOTELS

Although the earliest hotels date from the eighteenth century, their growth occurred only in the following century when the railways created sufficiently large markets. Hotels provide accommodation, meals and refreshments for irregular periods of time for those who may reserve their accommodation either in advance or on the premises.

Hotels provide facilities to meet the needs of the modern traveller. "A place where all who conduct themselves properly, and who, being able and ready to pay for their entertainment are received, if there be accommodation for them, and who without any stipulated engagement as to the duration of their stay or as to the rate of compensation, are, while there, supplied at a reasonable cost with their meals, lodging, and such services and attention as are necessarily incident to the use of the house as a temporary home".

TYPES OF HOTELS

Over the years the concept and the format of hotels have changed a great deal. There are various types of hotels catering to the increasing demands of tourists. The size, the façade, architectural features and the facilities and amenities provided differ from one establishment to another. The following are the main types of hotels:

- International Hotels
- Commercial Hotels
- Residential Hotels
- Heritage Hotels

SUPPLEMENTARY ACCOMMODATION

Supplementary accommodation consists of various types of accommodation other than the conventional hotels type. There are a series of other installations that are able to offer tourists lodging, Food and corresponding services. This is popularly known as supplementary accommodation and some of the principal forms of supplementary accommodation are:

- Motel
- Youth Hostels
- Camping sites
- Bed and breakfast establishments
- Tourist holiday villages and
- Time share and resorts condominiums.

This present chapter deals with the perception of accommodation providers in Kanchipuram district. In fact, the hotels have created a positive impact on the development of tourism industry. Overall the tourist are very much satisfied regarding the location of the hotels, seasonal variations, price fluctuations and the facilities offered for them. The researcher has made an attempt to provide the bird's view of the private hotels in the tourist places of Kanchipuram district.

OBJECTIVES OF THE STUDY

- To study the background of the accommodation providers.
- To find out the occupancy rate in the accommodation providers in Kanchipuram district.

HYPOTHESIS OF THE STUDY

- There is no direct relationship between the medical facilities and the business experience of the accommodation providers.
- There is no association between the medical facilities and location of the accommodation providers.

LIMITATIONS OF THE STUDY

Due to time constraints the researcher collected the 120 accommodation providers in Kanchipuram district. The accommodation providers are not given the fullest information to interviewer for the fear. Some of the accommodation providers were only given the general information.

RESEARCH METHODOLOGY

The researcher collected the various data for the research in primary and secondary data. Primary data was collected through Interview schedule method secondary data was collected from various books, journal, research articles, websites, Unpublished records and working papers etc. the researcher using the exploratory research. The researcher also used the convenience sampling technique. The researcher has collected the information for 120 accommodation providers. The statistical in the research tools has been used for mean, standard deviation, for the occupancy rate per week and chi square test.

REVIEW OF LITERATURE

To achieve the research objectives, exploratory research of descriptive nature was used. According to Rodrigues (2007), exploratory study aims to provide greater familiarity with the problem. Thus, the study intends to expand possible discoveries on the topic of OTAs in Brazil, more specifically online consumers perception about these companies, whose studies are still in an incipient stage.

It also has a descriptive character. In the view of Rodrigues (2006), descriptive research aims to observe, register, classify, and interpret, making use of standard techniques of data collection, such as questionnaires and systematic observation, thus showing the interrelationships among factors that influence the formation and development of the research and contribute to reach the objectives.

The present study has as preponderant characteristic, the quantitative approach, since its objective is to quantify results and show the relationships between them, having as target population the consumers over 18 years who have already booked accommodation using OTAs. It was used a purposely, non-probability sample, with the intention of contemplating at least 130 respondents, due to the selected analysis technique.

According to Dencker (2000), before the administration of the questionnaire, a pre-test must be conducted to assess any problems in the writing of questions, imprecisions, exhaustion, among others. From this perspective, the pilot test evaluated the data collection process and had the purpose of adjusting the questions and verifying their relationship with the research objectives. Twenty questionnaires were used with visitors from Centro de Lançamentos da Barreira do Inferno (CLBI), in Parnamirim, RN, which is a base of the Brazilian Air Force, where Cultural and Tourist Information Center (CCEIT) is located, focusing on tourists who are on their way to the Potiguar coast. The questionnaires were applied from 26 October to 2 November 2016. It was verified that the data collection instrument achieved the proposed research objective, with only minor changes required to the questionnaire, which was important to improve the instrument for the final data collection.

Therefore, the final questionnaire was adapted and designed considering previous literature, specifically the research carried out by Law (2009) and Del Chiappa (2013), which investigated the perception of different online consumer groups of hotels with different online shopping experiences. Data collection took place from December 2016 to March 2017 (during high season), so hotels had high occupancy rates, which facilitates the approach. Structured questionnaires were applied to guests staying in two three-star hotels, located in Ponta Negra neighborhood, in Natal, RN. The choice for three-star hotels was due to the ease of access of the researcher, since the most luxurious hotels are located in places of difficult access in the city. A total of 131 questionnaires were collected from the guests. In one of the hotels the approach was made by a receptionist, who was trained to administer the questionnaires. In the other hotel, the researcher herself administered the questionnaires to the guests, after authorization from the management.

ANALYSIS AND DISCUSSION

Table – 1- General Information About Accommodation Providers

Sl No	Source	General information	No of hotels	Per centage
1	Age (in years)	Less than 5years	32	26.67
		5-10 years	41	34.16
		Above 10 years	47	39.17
2	**Investments (Rs. In crores)**	Less than Rs.2 crore	48	40.00
		Rs.1 crore – Rs. 3 crores	36	30.00
		Rs.3 crore – Rs. 5 Crore	21	17.50
		Above Rs.5 crores	15	12.50
3.	**Forms of proprietor**	Proprietor	26	21.67
		Partnership	66	55.00
		Private	28	23.33
4.	**Types of accommodation**	Ordinary	60	50.00
		Five star	18	15.00
		Three star	21	17.50
		A/C hotels	21	17.50
5	**Mode of acquisition**	Newly started	57	47.50
		Lease	38	31.67
		Inheritance	25	20.83
6	**Location of the hotels**	Near Bus stand	64	53.33
		Outskirts of the town	38	31.67
		Near railway station	18	15.00

Source: Primary Data

It is inferred that 39.17 per cent (47) of the accommodation providers are experienced in the hotel industry tune-up their service 10 years and above. It is found that 40 per cent (48) of the hoteliers' investment of fund is less then Rs. 25, 00,000 for miniature type of hotels. The present study shows that 55 per cent (66) of the hotels were formed in partnership type. i.e. individual cannot invest huge amount into hotels. So they form partnership firm in Kanchipuram district. It is observed that 50 per cent (60) of the hotels are ordinary level. It requires minimum capital. It is easy to get financial assistance from financial institution for running ordinary type of hotels. It is learned from the study that, 47.50 per cent (57) of the hoteliers are newly entered to the hotel industry. The preponderance 29.16 per cent (35) of the respondents started this business as they possess required finance to start business. Finance is the pre- requisite to start any business .

Table 2- Activities Outsourced In Private Accommodations

Sl.no	Activities	No.of Accommodation providers	% of Total
1	Frequent cleaning of rooms	5	4.17
2	Better laundry facilities	24	20.00
3	Round the clock services	40	33.33
4	Well safety and security	21	17.50
5	Swimming pool maintenance	16	13.33
6	Hygiene food	14	11.67
	Total	**120**	**100.00**

Source: Primary Data

Out of the sample hoteliers 4.17 per cent (5) of them have provided frequent cleaning the room; 20.00 per cent (24) of them have provided better laundry facilities; 33.33 per cent (40) of them have provided round the clock service, 17.50 per cent (21) of them have provided well safety and security; 13.33 per cent (16) of them have provided swimming pool facility and the rest 11.67 per cent (14) of them have provided hygiene food to their guests.

Stay of tourists per day in a week in hotels may vary from Monday to Sunday. Hence occupancy per day in a week is analyzed and shown in table 3.

Table 7.17 Occupancy Per Day In A Week

Occupancy per cent for a week	Minimum	Maximum	Sum	Mean	Standard deviation
Monday	30	72	6296	57.76	8.512
Tuesday	50	82	7311	67.07	6.287
Wednesday	45	78	7063	64.80	6.719
Thursday	54	76	6883	63.15	5.810
Friday	-	76	6414	58.84	9.379
Saturday	44	75	6480	59.45	7.641
Sunday	55	82	7356	67.49	6.699

Source: Computed from Primary Data

It is lucid from Table 7.17 the maximum stay of guests in hotels is on Tuesday and Wednesday followed by Sunday. Thursday and Friday occupies the third position. However, it is an accepted fact that no hotels can expect 100 per cent occupancy on all days of the week. The average occupancy will be around 75 per cent. Festival time and holidays time the occupancy is 100 per cent occupied in the study area.

Table 4 - Medical Facilities

Sl. No	Opinion	No.of Accommodation providers	% of Total
1	Yes	115	95.83
2	No	05	04.17
Total		**120**	**100**

Source: Primary Data

Out of 120 sample hotels, 95.83 per cent (115) of the hotels provide medical facilities in their hotels and 4.17 per cent (5) of the hotels are not providing any kind of medical facilities in their premises.

TESTING OF HYPOTHESIS

In order to prove the fact, if there is any significant relationship between medical facilities and the business experience of the hotel owners. A null hypothesis is formulated chi-square test is applied accordingly.

A null hypothesis is framed that H_{01}**"There is no direct relationship between the medical facilities and the business experience of the accommodation providers".**

In statistic is useful for comparison of observed frequencies with theoretical frequencies and to draw decision whether there is any significant difference between these two sets Chi- Square test is a non-parametric test.

Table 5 Medical Facilities And Business Experience

Business experience	Medical facilities		Total
	Yes	No	
Less than 5 years	32 (27.83%)	0(-)	32 (26.67%)
5 - 10 years	40 (34.78%)	1(20%)	41(34.17%)
10 years and above	43 (37.39%)	4 (80%)	47 (39.17%)
Total	**115 (100%)**	**05 (100%)**	**120(100%)**

The calculated Chi-Square value is 3.6468- significant (at 2 degrees of freedom table value is 5.991 at 5% level of significance). The computed chi-square value is 3.6468 is less than the table value 5.991. So formulation of null hypothesis is accepted. Hence it is concluded that, H_{01}**"There is no direct relationship between the medical facilities and the business experience of the accommodation providers".**

TESTING OF HYPOTHESIS

H_{02}**"There is no association between the medical facilities and location of the accommodation providers".**

Table 7.24 Medical facilities and location of the accommodation providers

Location of the accommodation providers	Medical facilities		Total (% of Total)
	Yes (% of Total)	No (% of Total)	
Near bus stand	62 (53.91%)	02 (40%)	64 (53.33%)
Outskirts of the town	38 (33.04%)	00 (-)	38(31.67%)
Near railway station	15 (13.04%)	03(60%)	18 (15.00%)
Total	**115**	**05**	**120**

Calculated chi- square value is 1.9053- significant (at 2 degrees of freedom the table value is 5.991 at 5 % level of significance).The computed chi-square value is 1.9053 is less than the table value 5.991. So formulation of null hypothesis accepted.Hence it is concluded that, H_{052}**"There is no association between the medical facilities and location of the accommodation providers".**

FINDINGS OF THE STUDY

It is inferred that 39.17 per cent (47) of the accommodation providers are experienced in the hotel industry tune-up their service 10 years and above. It is found that 40 per cent (48) of the hoteliers' investment of fund is less then Rs. 2 crores for miniature type of hotels.The present study shows that 55 per cent (66) of the hotels were formed in partnership type. i.e. individual cannot invest huge amount into hotels. So they form partnership firm in Kanchipuram district.It is observed that 50 per cent (60) of the hotels are ordinary level. It requires minimum capital. It is easy to get financial assistance from financial institution for running ordinary type of hotels.

It is learned from the study that, 47.50 per cent (57) of the hoteliers are newly entered to the hotel industry.The preponderance 29.16 per cent (35) of the respondents started this business as they possess required finance to start business. Finance is the pre- requisite to start any business.It is a logically analyzed that 45.00 per cent (54) share of accommodation providers said that the tourists are coming to their hotel one times (Table 7.7). By providing the good quality of service, customer service, more facilities the hotel can attract tourists to come repeatedly the same hotel.The study reveals that around 53.33 per cent (64) of the hotels are located near bus stand for the convenient of the tourists mostly using the road transport.

It is lucid from table 7.17 the maximum stay of guests in hotels is on Tuesday and Wednesday followed by Sunday. Thursday and Friday occupy the third position. However, it is an accepted fact that no hotels can expect 100 per cent occupancy on all days of the week. The average occupancy will be around 75 per cent. Festival time and holidays time the occupancy is 100 per cent occupied in the study area.Almost 95.83 per cent (115) of the accommodation providers in the Kanchipuram district are providing medical facilities to their guest whenever needed.

RECOMMENDATIONS

- The Government has to attract more foreign tourists by the way of online advertisements with highlighting tourist places in Kanchipuram district.
- Government should take step to reduce the parking charges and to fix the nominal cost for all the vehicles to two wheelers, four wheelers, auto, etc.
- The study reveals that the majority of the accommodation providers prefer payments only in cash. Accommodation providers must introduce all the easy way of new innovations in banking sector for collecting the payment which will attract more number of customers.
- Tourist prefers non A/C deluxe and double room. So the accommodation providers should construct more number of double rooms and non A/C deluxe rooms to the requirement of the tourist.
- The accommodation providers should try to provide a quality and hygiene healthy food to the tourist in the study area. And the tourist expects more recreational activities like dances, expo etc. while visiting the Kanchipuram District.
- Medical facilities must be kept geared up at all times at the hotel premises in the tourist place to attract the tourist.
- Any type of hotel authorities may go out of their way in serving their customers with the best food, accommodation services and bringing about flexibility in their services to make their stay a pleasurable experience.

CONCLUSION

To conclude, all these improvements and facilities are done and twisted, in order to bring a sea change in the tourism diligence in general and specifically for Kanchipuram District in particular. The economic significance of tourism in terms of employment, income, foreign exchange earnings and regional development is a major driving force to place tourism appropriately in development. Medical facility is more important so the accommodation providers are gives more services offered by all hotels. In emergency first aid will be given for tourists. Many hotels possess pharmacy, round the clock medical services and doctors are available at any time in the hotels.

REFERENCE

1. Batra, G.S, Dangwal, R.C: “Tourism Promotion and Development” Deep and Deep Publication, New Delhi, 2000.
2. Chauhan, Ravee: “Eco Tourisms-Trends and Challenges”, Vista International Publishing House, New Delhi, 2006.
3. Ghosh, Bishwnath: “Tourism and Travel Management”, Vikas Publishing House pvt. Ltd., New Delhi, 2001.
4. Biju.M.R Hospitality industry in Kerala: Performances and promises SAJOSPS Vol.2No.2 June 2002.
5. Biju.M.R., Facets of Tourism: Global and Indian Dimensions, SAJOSPS, Vol.8No.1 December 2007.
6. Ghimire Him Lal: “New Strategies for Tourism Marketing in Nepal”, The Journal of Business Studies, Vol.2, No.1, 2004.
7. Connell, J: “Medical tourism: Sea, Sun, Sand and … Surgery”, Tourism Management, Vol. 27, No.6, 2006.
8. Ministry of Tourism, Annual Report 2013-14
9. World Tourism Organization # Bureau of immigration of India 2019-2020.
10. State Tourism / UT Departments 2018-2019.
11. Tourist statistics, Department of tourism, Tamil Nadu Government of Tamil Nadu2011-2021.

DOI No. : 10.5958/2394-9309.2022.00044.0
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

CONNECTED HUB SETS AND CONNECTED HUB POLYNOMIALS OF SQUARE OF PATHS

T. Angelinshiny* And T. Anitha Baby**
*Research Scholar (Reg. No. 20213282092009), Department of Mathematics, Women's Christian College, Nagercoil, Tamil Nadu, India
**Assistant Professor, Department of Mathematics, Women's Christian College, Nagercoil, Tamil Nadu, India
Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli - 627 012.
E-mail :angelinshinyt@gmail.com, anithasteve@gmail.com,

ABSTRACT

Let P_p^2 be the square of path with p vertices. Let H_c (P_P^2,k) denotes the family of connected hub sets ofP_p^2 of cardinality k. Then, the polynomial,

$$H_c(P_p^2, x) = \sum_{k=\hbar_c(P_p^2)}^{|V(P_p^2)|} h_c(P_p^2, k)\, y^k,$$

is called the connected hub polynomial of $P_p^2$where $h_c(P_p^2, k)$ denotes the number of connected hub sets of P_p^2 of cardinality k, and $\hbar_c(P_p^2)$ is the connected hub number of P_p^2.
In this paper we derived a recursive formula for $h_c(P_p^2, k)$. Using this recursive formula we construct the connected hub polynomial of the square of path as,

$$H_c(P_p^2, x) = \sum_{k=\left\lceil \frac{p-2}{2} \right\rceil}^{p} h_c(P_p^2, k)\, y^k$$

and some of the properties of this polynomial also have been studied.
Mathematics Subject Classification Code 05C38, 05C31, 05C99
Keywords : Square of path, Connected hub sets, Connected hub number, Connected hub polynomials.

1. INTRODUCTION.

Let $G = (V, E)$ be a simple graph. The number of vertices in G is called the order of G and the number of edges in G is called the size of G.

The square of a path P_p is the graph obtained by joining every pair of vertices of distance two in the path and is denoted by P_p^2.

As usual we use $\lfloor p \rfloor$ for the largest integer less than or equal to p and $\lceil p \rceil$ for the smallest integer greater than or equal to p. Also, we denote the set $\{1,2,3,\cdots,p\}$ by $[p]$, throughout this paper.

2. CONNECTED HUB SETS OF SQUARE OF PATHS.

In this section, we give the connected hub number of the square of paths and some of the properties of connected hub sets.

Definition 2.1.

Let $G = (V, E)$ be a connected graph. A subset H of V is called a hub set of G if for any two distinct vertices $u, v \in V - H$, there exists a $u - v$ path P in G, such that all the internal vertices of P are in H.The minimum cardinality of a hub set of G is called the hub number of G and is denoted by $\hbar(G)$.

Definition 2.2.

A hub set H of G is called a connected hubset if the induced subgraph $< H >$ is connected. The minimum cardinality of a connected hub set of G is called connected hub number of G and is denoted by $\hbar_c(G)$.

Lemma 2.3.

Let P_p^2 be the square of path with p vertices. Then its connected hub number is, $\hbar_c\left(P_p^2\right) = \left\lfloor \frac{p-2}{2} \right\rfloor$.

Lemma 2.4.

Let P_p^2, $p \geq 4$ be the square of path with order p. Then $h_c\left(P_p^2, k\right) = 0$ if $k < \left\lfloor \frac{p-2}{2} \right\rfloor$ or $k > p$ and $h_c\left(P_p^2, k\right) > 0$ if $\left\lfloor \frac{p-2}{2} \right\rfloor \leq k \leq p$.

Proof:

There is no connected hub set of cardinality k, when $k < \left\lfloor \frac{p-2}{2} \right\rfloor$ or $k > p$.

Therefore, $h_c\left(P_p^2, k\right) = 0$ if $k < \left\lfloor \frac{p-2}{2} \right\rfloor$ or $k > p$.

By Lemma 2.3, the cardinality of the minimum connected hub set of the square of path P_p^2 is $\left\lfloor \frac{p-2}{2} \right\rfloor$ and the cardinality of the maximum connected hub set of the square of path P_p^2 is p.

Therefore, $h_c\left(P_p^2, k\right) > 0$ if $k \geq \left\lfloor \frac{p-2}{2} \right\rfloor$ or $k \leq p$.

Hence, we have,

$h_c\left(P_p^2, k\right) = 0$ if $k < \left\lfloor \frac{p-2}{2} \right\rfloor$ or $k > p$ and $h_c\left(P_p^2, k\right) > 0$ if $\left\lfloor \frac{p-2}{2} \right\rfloor \leq k \leq p$.

Lemma 2.5.

Let P_p^2, $p \geq 4$ be the square of path with $\left|V\left(P_p^2\right)\right| = p$. Then,

1. If $H_c\left(P_{p-1}^2, k-1\right) = \phi$ and $H_c\left(P_{p-3}^2, k-1\right) = \phi$ then $H_c\left(P_{p-2}^2, k-1\right) = \phi$.
2. If $H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-3}^2, k-1\right) \neq \phi$ then $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$.
3. If $H_c\left(P_{p-1}^2, k-1\right) = \phi$ and $H_c\left(P_{p-2}^2, k-1\right) = \phi$ then $H_c\left(P_p^2, k\right) = \phi$.
4. If $H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$ then $H_c\left(P_p^2, k\right) \neq \phi$.

Proof:

1. Since, $H_c\left(P_{p-1}^2, k-1\right) = \phi$ and $H_c\left(P_{p-3}^2, k-1\right) = \phi$, by Lemma 2.4,

$k-1 > p-1$ or $k-1 < \left\lfloor\frac{p-3}{2}\right\rfloor$ and

$$k-1 > p-3 \text{ or } k-1 < \left\lfloor\frac{p-5}{2}\right\rfloor$$

Therefore, $k-1 > p-1$ or $k-1 < \left\lfloor\frac{p-5}{2}\right\rfloor$

Therefore, $k-1 > p-2$ or $k-1 < \left\lfloor\frac{p-4}{2}\right\rfloor$ holds.

Hence, $H_c(P_{p-2}^2, k-1) = \phi$.

2. Assume that $H_c(P_{p-1}^2, k-1) \neq \phi$ and $H_c(P_{p-3}^2, k-1) \neq \phi$

Suppose, $H_c(P_{p-2}^2, k-1) = \phi$

Then by Lemma 2.4,

$k-1 > p-2$ or $k-1 < \left\lfloor\frac{p-4}{2}\right\rfloor$.

If $k-1 > p-2$, then $k-1 > p-3$.

Therefore, $H_c(P_{p-3}^2, k-1) = \phi$, a contradiction.

If $k-1 < \left\lfloor\frac{p-4}{2}\right\rfloor$, then $k-1 < \left\lfloor\frac{p-3}{2}\right\rfloor$.

Therefore, $H_c(P_{p-1}^2, k-1) = \phi$, a contradiction.

Hence, $H_c(P_{p-2}^2, k-1) \neq \phi$.

3. Since, $H_c(P_{p-1}^2, k-1) = \phi$ and $H_c(P_{p-2}^2, k-1) = \phi$,

By Lemma 2.4,

$k-1 > p-1$ or $k-1 < \left\lfloor\frac{p-3}{2}\right\rfloor$ and

$k-1 > p-2$ or $k-1 < \left\lfloor\frac{p-4}{2}\right\rfloor$.

Therefore, $k-1 > p-1$ or $k-1 < \left\lfloor\frac{p-4}{2}\right\rfloor$.

Therefore, $k > p$ or $k < \left\lfloor\frac{p-2}{2}\right\rfloor$.

Hence, $H_c(P_p^2, k) = \phi$.

4. Assume that $H_c(P_{p-1}^2, k-1) \neq \phi$ and $H_c(P_{p-2}^2, k-1) \neq \phi$

Then by Lemma 2.4, we have,

$\left\lfloor\frac{p-3}{2}\right\rfloor \leq k-1 \leq p-1$ and $\left\lfloor\frac{p-4}{2}\right\rfloor \leq k-1 \leq p-2$.

Therefore, $\left\lfloor\frac{p-4}{2}\right\rfloor \leq k-1 \leq p-1$.

Therefore, $\left\lfloor\frac{p-2}{2}\right\rfloor \leq k \leq p$

Hence, $H_c(P_p^2, k) \neq \phi$.

Lemma 2.6.

Let P_p^2, $p \geq 4$ be the square of path with $|V(P_p^2)| = p$. Suppose that $H_c(P_p^2, k) \neq \phi$, then,

1. If $H_c(P_{p-1}^2, k-1) = \phi$ and $H_c(P_{p-2}^2, k-1) \neq \phi$ iff $p = 2j+3$ and $k = j$.
2. If $H_c(P_{p-1}^2, k-1) \neq \phi$ and $H_c(P_{p-2}^2, k-1) = \phi$ iff $k = p$.
3. If $H_c(P_{p-1}^2, k-1) \neq \phi$ and $H_c(P_{p-2}^2, k-1) \neq \phi$ iff $\left\lfloor\frac{p-3}{2}\right\rfloor + 1 \leq k \leq p-1$.

Proof:

Assume that $H_c(P_p^2, k) \neq \phi$.

1. Since, $H_c(P_{p-1}^2, k-1) = \phi$, by Lemma 2.4,

$$k-1 > p-1 \text{ or } k-1 < \left\lfloor \frac{p-3}{2} \right\rfloor$$

If $k-1 > p-1$, then $k > p$.

Therefore, by Lemma 2.4, we have,

$H_c(P_p^2, k) = \phi$, a contradiction.

Therefore, $k-1 < \left\lfloor \frac{p-3}{2} \right\rfloor$

That is $k < \left\lfloor \frac{p-1}{2} \right\rfloor$ (1)

Also, since, $H_c(P_{p-2}^2, k-1) \neq \phi$, by Lemma 2.4,

$\left\lfloor \frac{p-4}{2} \right\rfloor \leq k-1 \leq p-2$.

Therefore, $\left\lfloor \frac{p-2}{2} \right\rfloor \leq k$ (2)

From (1) and (2), we get,

$\left\lfloor \frac{p-2}{2} \right\rfloor \leq k < \left\lfloor \frac{p-1}{2} \right\rfloor$, and this inequality is true only when $p = 2j+3$ and $k = j$.

Conversely, if $p = 2j+3$ and $k = j$, then by Lemma 2.4,

we have,

$H_c(P_{p-1}^2, k-1) = H_c(P_{2j+2}, j-1) = \phi$ and $H_c(P_{p-2}^2, k-1) = H_c(P_{2j+1}, j-1) \neq \phi$.

2. Since, $H_c(P_{p-2}^2, k-1) = \phi$,by Lemma 2.4,

$k-1 > p-2$ or $k-1 < \left\lfloor \frac{p-4}{2} \right\rfloor$.

If $k-1 > p-2$, then $k < \left\lfloor \frac{p-2}{2} \right\rfloor$.

Therefore, $H_c(P_p^2, k) = \phi$, a contradiction.

Therefore,$k-1 > p-2$,

Therefore, $k > p-1$,

Hence, $k \geq p$ (3)

Also, since $H_c(P_{p-1}^2, k-1) \neq \phi$, by Lemma 2.4,

$\left\lfloor \frac{p-3}{2} \right\rfloor \leq k-1 \leq p-1$.

Therefore $k \leq p$ (4)

From (3) and (4) we get,

$k = p$.

Conversely, if $k = p$, then by Lemma 2.4,

$H_c(P_{p-1}^2, k-1) = H_c(P_{p-1}, p-1) \neq \phi$ and $H_c(P_{p-2}^2, k-1) = H_c(P_{p-2}, p-1) = \phi$.

3. Assume that, $H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$,

Then by Lemma 2.4,
$\left\lfloor\frac{p-3}{2}\right\rfloor \leq k-1 \leq p-1$ and $\left\lfloor\frac{p-4}{2}\right\rfloor \leq k-1 \leq p-2$.
Therefore, $\left\lfloor\frac{p-3}{2}\right\rfloor \leq k-1 \leq p-2$.
Therefore, $\left\lfloor\frac{p-3}{2}\right\rfloor + 1 \leq k \leq p-1$.

Conversely, suppose,
$\left\lfloor\frac{p-3}{2}\right\rfloor + 1 \leq k \leq p-1$, then by Lemma 2.4,
$H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$

Lemma 2.7.
For every $p \geq 4$, $p \geq \left\lfloor\frac{p-2}{2}\right\rfloor$.

1. If $H_c\left(P_{p-1}^2, k-1\right) = \phi$ and $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$ then $H_c\left(P_p^2, k\right) = \left\{X \cup \{p-2\}\right\}$, where $X \in H_c\left(P_{p-2}^2, k-1\right)$.
2. If $H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-2}^2, k-1\right) = \phi$ then $H_c\left(P_p^2, k\right) = \left\{X \cup \{p\}\right\}$, where $X \in H_c\left(P_{p-1}^2, k-1\right)$.
3. If $H_c\left(P_{p-1}^2, k-1\right) \neq \phi$ and $H_c\left(P_{p-2}^2, k-1\right) \neq \phi$ then

(i) When $k \neq p-4$,

$$H_c\left(P_p^2, k\right) = \begin{cases} X \cup \{p-3\}\ if\ X\ ends\ with\ \{p-4\} \\ \cup\ X \cup \{p-2\}\ if\ X\ ends\ with\ \{p-3\} \\ \cup\ X \cup \{p-1\}\ if\ X\ ends\ with\ \{p-2\} \\ \cup\ Y \cup \{p-2\}\ if\ Y\ ends\ with\ \{p-4\} \\ \cup\ Y \cup \{p-1\}\ if\ Y\ ends\ with\ \{p-3\} \\ \cup\ Y \cup \{p\}\ if\ Y\ ends\ with\ \{p-2\}\ or\ \{p-1\} \end{cases}$$

where $X \in H_c\left(P_{p-2}^2, k-1\right)$ and $Y \in H_c\left(P_{p-1}^2, k-1\right)$.

(ii) When $k = p-4$,

$$H_c\left(P_p^2, k\right) = \begin{cases} X \cup \{p-2\}\ if\ X\ ends\ with\ \{p-4\}\ or\ \{p-3\} \\ \cup\ X \cup \{p-1\}\ if\ X\ ends\ with\ \{p-2\} \\ \cup\ Y \cup \{p-1\}\ if\ Y\ ends\ with\ \{p-3\} \\ \cup\ Y \cup \{p\}\ if\ Y\ ends\ with\ \{p-2\} \end{cases}$$

where $X \in H_c\left(P_{p-2}^2, k-1\right)$ and $X \neq \{1,2,\cdots,p-5\}, \{4,5,\cdots,p-2\}$,
$Y \in H_c\left(P_{p-1}^2, k-1\right)$

Proof:

1. In this case, $H_c\left(P_{p-2}^2, k-1\right)$ has only one connected hub set and that connected hub set end with $\{p-4\}$.

Let X be the only connected hub set of P_p^2 of cardinality $k-1$. Adjoin $\{p-2\}$to X. Hence the connected hub set X of $H_c\left(P_{p-2}^2, k-1\right)$ belongs to $H_c\left(P_p^2, k\right)$ by adjoining $\{p-2\}$only. Therefore, $H_c\left(P_p^2, k\right) = \left\{X \cup \{p-2\}\right\}$, where $X \in H_c\left(P_{p-2}^2, k-1\right)$.

2. In this case, $H_c\left(P_{p-1}^2, k-1\right)$ has only one connected hub set and that connected hub set end with$\{p-1\}$.

Let X be the only connected hub set of P_{p-1}^2 of cardinality $k-1$. Adjoin $\{p\}$to X. Hence the connected hub set X of $H_c\left(P_{p-1}^2, k-1\right)$ belongs to $H_c\left(P_p^2, k\right)$ by adjoining $\{p\}$ only. Therefore, $H_c\left(P_p^2, k\right) = \{X \cup \{p\}\}$, where $X \in H_c\left(P_{p-1}^2, k-1\right)$.

3. (i) The construction of $H_c\left(P_p^2, k\right)$ from $H_c\left(P_{p-2}^2, k-1\right)$ and $H_c\left(P_{p-1}^2, k-1\right)$ when $k \neq p-4$.

Let X be a connected hub set of P_{p-2}^2 of cardinality $k-1$.

Then all the elements of X end with $\{p-4\}$ or $\{p-3\}$ or $\{p-2\}$

If X ends with $\{p-4\}$, adjoin $\{p-3\}$ to X.

If X ends with $\{p-3\}$, adjoin $\{p-2\}$ to X.

If X ends with $\{p-2\}$, adjoin $\{p-1\}$ to X.

In this way, every X of $H_c\left(P_{p-2}^2, k-1\right)$ belongs to $H_c\left(P_p^2, k\right)$.

Now, let Y be a connected hub set of P_{p-1}^2 of cardinality $k-1$.

Then all the elements of Y end with $\{p-4\}$ or $\{p-3\}$ or $\{p-2\}$ or $\{p-1\}$.

If Y ends with $\{p-4\}$, adjoin $\{p-2\}$ to Y.

If Y ends with $\{p-3\}$, adjoin $\{p-1\}$ to Y.

If Y ends with $\{p-2\}$ or $\{p-1\}$, adjoin $\{p\}$ to Y.

In this way every Y of $H_c\left(P_{p-1}^2, k-1\right)$ belongs to $H_c\left(P_p^2, k\right)$.

Therefore,

$$H_c\left(P_p^2, k\right) = \begin{cases} X \cup \{p-3\} \; if \; X \; ends \; with \; \{p-4\} \\ \cup \, X \cup \{p-2\} \; if \; X \; ends \; with \; \{p-3\} \\ \cup \, X \cup \{p-1\} \; if \; X \; ends \; with \; \{p-2\} \\ \cup \, Y \cup \{p-2\} \; if \; Y \; ends \; with \; \{p-4\} \\ \cup \, Y \cup \{p-1\} \; if \; Y \; ends \; with \; \{p-3\} \\ \cup \, Y \cup \{p\} \; if \; Y \; ends \; with \; \{p-2\} \; or \; \{p-1\} \end{cases}$$

where $X \in H_c\left(P_{p-2}^2, k-1\right)$ and $Y \in H_c\left(P_{p-1}^2, k-1\right)$.

(ii) The construction of $H_c\left(P_p^2, k\right)$ from $H_c\left(P_{p-2}^2, k-1\right)$ and $H_c\left(P_{p-1}^2, k-1\right)$ when $k = p-4$.

Let X be a connected hub set of P_{p-2}^2 of cardinality $k-1$.

Remove the sets $\{1,2,3,\cdots,p-5\}$ and $\{4,5,6,\cdots p-2\}$ from $H_c\left(P_{p-2}^2, k-1\right)$. The remaining sets end with $\{p-4\}$ or $\{p-3\}$ or $\{p-2\}$.

If X ends with $\{p-4\}$ or $\{p-3\}$, adjoin $\{p-2\}$ to X.

If X ends with $\{p-2\}$, adjoin $\{p-1\}$ to X.

In this way, every X of $H_c\left(P_{p-2}^2, k-1\right)$ belongs to $H_c\left(P_p^2, k\right)$ except two sets.

Now, let Y be a connected hub set of P_{p-1}^2 of cardinality $k-1$.

Then all the elements of Y end with $\{p-3\}$ or $\{p-2\}$.

If Y ends with $\{p-3\}$, adjoin $\{p-1\}$ to Y.

If Y ends with $\{p-2\}$, adjoin $\{p\}$ to Y.

In this way every Y of $H_c(P_{p-1}^2, k-1)$ belongs to $H_c(P_p^2, k)$.

Therefore,

$$H_c(P_p^2, k) = \begin{cases} X \cup \{p-2\}\ if\ X\ ends\ with\ \{p-4\}\ or\ \{p-3\} \\ \cup\ X \cup \{p-1\}\ if\ X\ ends\ with\ \{p-2\} \\ \cup\ Y \cup \{p-1\}\ if\ Y\ ends\ with\ \{p-3\} \\ \cup\ Y \cup \{p\}\ if\ Y\ ends\ with\ \{p-2\} \end{cases}$$

Where $X \in H_c(P_{p-2}^2, k-1)$ and$X \neq \{1,2,\cdots,p-5\},\{4,5,\cdots,p-2\}$,

$Y \in H_c(P_{p-1}^2, k-1)$

Theorem 2.8.

Let $H_c(P_p^2, k)$ be the family of connected hub sets of P_p^2 with cardinality k, where $k \geq \left\lfloor\frac{p-2}{2}\right\rfloor$. Then

(i) $h_c(P_p^2, k) = h_c(P_{p-1}^2, k-1) + h_c(P_{p-2}^2, k-1)$, if $\left\lfloor\frac{p-2}{2}\right\rfloor \leq k \leq p$ and $k \neq p-4$.

(ii) $h_c(P_p^2, k) = h_c(P_{p-1}^2, k-1) + h_c(P_{p-2}^2, k-1) - 2$, if $k = p-4$.

Proof:

We consider the four cases given in Theorem 2.7.

Case (i)

By Theorem 2.7 (1), we have

$H_c(P_p^2, k) = \{X \cup \{p-2\}\}$, where $X \in H_c(P_{p-2}^2, k-1)$.

Since $H_c(P_{p-1}^2, k-1) = \phi$, we have

$h_c(P_{p-1}^2, k-1) = 0$

Therefore, $h_c(P_p^2, k) = h_c(P_{p-1}^2, k-1)$.

Case (ii)

By Theorem 2.7 (2), we have

$H_c(P_p^2, k) = \{X \cup \{p\}\}$, where $X \in H_c(P_{p-1}^2, k-1)$.

Since $H_c(P_{p-2}^2, k-1) = \phi$, we have

$h_c(P_{p-2}^2, k-1) = 0$

Therefore, $h_c(P_p^2, k) = h_c(P_{p-2}^2, k-1)$.

Case (iii)

When $k \neq p-4$

By Theorem 2.7 (3(i)), we have,

$$H_c\left(P_p^2,k\right)=\begin{cases} X\cup\{p-3\}\ if\ X\ ends\ with\ \{p-4\} \\ \cup X\cup\{p-2\}\ if\ X\ ends\ with\ \{p-3\} \\ \cup X\cup\{p-1\}\ if\ X\ ends\ with\ \{p-2\} \\ \cup Y\cup\{p-2\}\ if\ Y\ ends\ with\ \{p-4\} \\ \cup Y\cup\{p-1\}\ if\ Y\ ends\ with\ \{p-3\} \\ \cup Y\cup\{p\}\ if\ Y\ ends\ with\ \{p-2\}\ or\ \{p-1\} \end{cases}$$

where $X\in H_c\left(P_{p-2}^2,k-1\right)$ and $Y\in H_c\left(P_{p-1}^2,k-1\right)$.

Therefore, $h_c\left(P_p^2,k\right)=h_c\left(P_{p-1}^2,k-1\right)+h_c\left(P_{p-2}^2,k-1\right)$ if $k\neq p-4$.

Case (iv)

When $k=p-4$

By Theorem 2.7 (3(ii)), we have,

$$H_c\left(P_p^2,k\right)=\begin{cases} X\cup\{p-2\}\ if\ X\ ends\ with\ \{p-4\}\ or\ \{p-3\} \\ \cup X\cup\{p-1\}\ if\ X\ ends\ with\ \{p-2\} \\ \cup Y\cup\{p-1\}\ if\ Y\ ends\ with\ \{p-3\} \\ \cup Y\cup\{p\}\ if\ Y\ ends\ with\ \{p-2\} \end{cases}$$

where $X\in H_c\left(P_{p-2}^2,k-1\right)$ and$X\neq\{1,2,\cdots,p-5\},\{4,5,\cdots,p-2\}$,

$Y\in H_c\left(P_{p-1}^2,k-1\right)$

Therefore, $h_c\left(P_p^2,k\right)=h_c\left(P_{p-1}^2,k-1\right)+h_c\left(P_{p-2}^2,k-1\right)-2$, if $k=p-4$.

3. CONNECTED HUB POLYNOMIALS OF SQUARE OF PATHS.

In this section, we define and find the connected hub polynomial of the square of path P_p^2. Also, we give the properties hold by the coefficients of the polynomial.

Definition 3.1.

Let $H_c\left(P_p^2,k\right)$ denotes the family of connected hub sets of P_p^2 with cardinality k. Then the connected hub polynomial $H_c\left(P_p^2,y\right)$ of P_p^2 is defined as

$$H_c\left(P_p^2,y\right)=\sum_{k=\hbar_c\left(P_p^2\right)}^{\left|V\left(P_p^2\right)\right|} h_c\left(P_p^2,k\right)y^k,$$

where, $h_c\left(P_p^2,k\right)$ denotes the number of connected hub sets of P_p^2 of cardinality k, and $\hbar_c\left(P_p^2\right)$ is connected hub number of P_p^2.

Theorem 3.2.

For every $p\geq 4$, $H_c\left(P_p^2,y\right)=y\left[H_c\left(P_{p-1}^2,y\right)+H_c\left(P_{p-2}^2,y\right)\right]-2y^{p-4}$ with initial values,

$$H_c(P_4^2,y)=4y+5y^2+4y^3+y^4$$

$$H_c(P_5^2,y)=y+7y^2+8y^3+5y^4+y^5$$

Proof:

From the definition of connected hub polynomial, we have,

$$H_c(P_p^2, y) = \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} h_c(P_p^2, k) y^k$$

$$= \sum_{\substack{k=\left\lfloor \frac{p-2}{2} \right\rfloor \\ k \neq p-4}}^{p} \left[h_c(P_{p-1}^2, k-1) + h_c(P_{p-2}^2, k-1)\right] y^k$$

$$+\left[h_c(P_{p-1}^2, k-1) + h_c(P_{p-2}^2, k-1) - 2\right] y^{p-4}$$

$$= \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} \left[h_c(P_{p-1}^2, k-1) + h_c(P_{p-2}^2, k-1)\right] y^k$$

$$-2y^{p-4}$$

$$= \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} h_c(P_{p-1}^2, k-1) y^k$$

$$+ \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} h_c(P_{p-2}^2, k-1) y^k - 2y^{p-4}$$

$$= y \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} h_c(1, k-1) y^{k-1}$$

$$+ y \sum_{k=\left\lfloor \frac{p-2}{2} \right\rfloor}^{p} h_c(P_{p-2}^2, k-1) y^{k-1} - 2y^{p-4}$$

$$= y\left[H_c(P_{p-1}^2, y) + H_c(P_{p-2}^2, y)\right] - 2y^{p-4}$$

Hence, $H_c(P_p^2, y) = y\left[H_c(P_{p-1}^2, y) + H_c(P_{p-2}^2, y)\right] - 2y^{p-4}$ for every $p \geq 4$ with initial values,

$$H_c(P_4^2, y) = 4y + 5y^2 + 4y^3 + y^4$$

$$H_c(P_5^2, y) = y + 7y^2 + 8y^3 + 5y^4 + y^5$$

Table 1: $\boldsymbol{h_c(P_p^2, k)}$**, the number of connected hub sets of** $\boldsymbol{P_p^2}$**of cardinality** $\boldsymbol{k}$**.**

p \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	3	3	1												
4	4	5	4	1											
5	1	7	8	5	1										
6	0	3	12	12	6	1									
7	0	1	8	20	17	7	1								
8	0	0	4	18	32	23	8	1							
9	0	0	1	12	36	49	30	9	1						
10	0	0	0	5	30	66	72	38	10	1					
11	0	0	0	1	17	66	113	102	47	11	1				
12	0	0	0	0	6	47	132	183	140	57	12	1			
13	0	0	0	0	1	23	113	245	283	187	68	13	1		
14	0	0	0	0	0	7	70	245	428	421	244	80	14	1	
15	0	0	0	0	0	1	30	183	490	11	606	312	93	15	1

Theorem 3.2.

The coefficients of the $H_c(P_p^2, y)$ satisfy the following properties.

1. $h_c(P_p^2, p) = 1$, for every $p \geq 4$.
2. $h_c(P_p^2, p-1) = p$, for every $p \geq 4$.

3. $h_c\left(P_p^2, p-2\right) = \frac{1}{2}(n^2 - 3n + 6)$, for every $p \geq 4$.
4. $h_c\left(P_p^2, p-3\right) = \frac{1}{6}(n^3 - 9n^2 + 38n - 48)$, for every $p \geq 4$.
5. $h_c\left(P_{2p+3}^2, p\right) = 1$, for every $p \geq 1$.
6. $h_c\left(P_{2p+2}^2, p\right) = p + 1$, for every $p \geq 2$.

REFERENCES:

1. *Walsh, Matthew. "The hub number of a graph." Int. J. Math. Comput. Sci 1, no. 1 (2006):117-124.*
2. *Veettil, Ragi Puthan, and T. V. Ramakrishnan. "Introduction to hub polynomial ofgraphs." Malaya Journal of Matematik (MJM) 8, no. 4, 2020 (2020): 1592-1596.*
3. *S. Alikhani and Y.-H. Peng. Dominating sets and domination polynomials of paths.International journal of Mathematics and mathematical sciences, 2009, 2009.*
4. *S. Alikhani and Y.-h. Peng. Introduction to domination polynomial of a graph. arXivpreprintar Xiv:0905.2251, 2009.*
5. *S. Alikhani and Y.-H. Peng. Dominating sets and domination polynomials of certain graphs,ii. Opuscula Mathematica, 30, 01 2010.*
6. *A. Vijayan and T. A. Baby. Connected total dominating sets and connected total dominationpolynomials of square of paths. International Journal of Mathematics Trends and Technology,11(1), 2014.*

DOI No. : 10.5958/2394-9309.2022.00045.2
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

THE UPPER EDGE MONOPHONIC GLOBAL DOMINATION NUMBER OF A GRAPH

V. Selvi, V*. Sujin Flower**
*Register Number. 20123042092008, Research Scholar, Department of Mathematics,
Holy Cross College (Autonomous), Nagercoil - 629 004, India
Assistant Professor, Department of Mathematics Holy Cross College (Autonomous), Nagercoil - 629 004, India
Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli - 627 012
E-mail: selvi.maths1983@gmail.com,sujinflower@gmail.com

ABSTRACT

An edge monophonic global dominating set of G is called a minimal edge monophonic global dominating set of G if no proper subset of S is an edge geodetic global dominating set of G. The upper edge monophonic global domination number $\underline{\gamma}_{me}^{+}(G)$ is the maximum cardinality of a minimal edge monophonic global dominating set of G. Some general properties satisfied by this concept are studied. The upper edge monophonic global domination number of some standard graphs are determined. Connected graphs of order $p \geq 2$, the upper edge monophonic global domination number 2 or p are characterized. It is shown that for every pair of integers a and b with $3 \leq a \leq b$, there exists a connected graph G such that $\underline{\gamma}_{me}(G) = a$ and $\underline{\gamma}_{me}^{+}(G) = b$.

Keywords: upper edge monophonic global domination number, edge monophonic global domination number, global domination number, monophonic number.
AMS Subject Classification: 05C12, 05C69.

1. INTRODUCTION

Let $G = (V, E)$ be a graph with a vertex set $V(G)$ and edge set $E(G)$ (or simply V and E, respectively, Furthermore, we say that a graph G has order $p = |V(G)|$ and size $m = |E(G)|$. For basic graph theoretic terminology, we refer to [2]. A vertex v is adjacent to another vertex u if and only if there exists an edge $e = uv \in E(G)$. If $uv \in E(G)$, we say that u is a *neighbor* of v and denote by $N_G(v)$, the set of neighbors of v. The degree of a vertex $v \in V$ is $(v) = |N_G(v)|$. A vertex v is said to be *universal vertex* if $(v) = p - 1$. A vertex v is called an *extreme vertex* if the sub- graph induced by v is complete.

The *length* of a path is the number of its edges. Let u and v be vertices of a connected graph G. A shortest u-v path is also called a u-v *geodesic*. The (shortest path) *distance* is defined as the length of a u-v geodesic in G and is denoted by $d_G(u, v)$ *or* $d(u, v)$ for short if the graph is clear from the context. A *chord* of a path P is an edge which connects two non-adjacent vertices of P. A u-v path is called a monophonic path if it is a chordless path. For two vertices x and y, the closed interval $J_e[x, y]$ consists of all edges lying in a $x - y$ monophonic. If x and y are adjacent, then $J_e[x, y] = \{xy\}$. For a set M of vertices, let $J_e[M] = \cup_{x,y \in M} J_e[x, y]$. Then certainly $M \subseteq J_e[M]$. A set $M \subseteq V(G)$ is called a *monophonic se*t of G if $J_e[M] = E(G)$. The edge monophonic number $m_e(G)$ of G is the minimum order of its monophonic sets and any edge monophonic set of order $m_e(G)$ is called a m_e-set of G. The edge monophonic number of a graph was studied in [1,5-9].

A subset $D \subseteq V(G)$ is called a *dominating set* if every vertex in $V \setminus D$ is adjacent to at least one vertex of D. The *domination number*, $\gamma(G)$, of a graph G denotes the minimum cardinality of such dominating sets of G. A minimum dominating set of a graph G is hence often called as a γ-set of G. The domination concept was studied in [3,4]. A subset $D \subseteq V$ is called a *global dominating set* in G if D is a dominating set in G and $\underline{G}$. The *global domination number* $\underline{\gamma}(G)$ is the minimum cardinality of a minimum global dominating set in G. The concept of global domination in graph was introduced in [10]. A set $M \subseteq V$ is said to be a *monophonic global dominating se*t of G if M is both a monophonic set and a global dominating set of G. The minimum cardinality of a monophonic global dominating set of G is the monophonic global domination number of G and is denoted by $\underline{\gamma_m}(G)$. A monophonic global dominating set of cardinality $\underline{\gamma_m}(G)$ is called a $\underline{\gamma_m}$-set of G. The concept of monophonic global domination in graph was introduced in [11]. An edge monophonic set S of a connected graph is said to be an edge monophonic global dominating set of G if S is both an edge monophonic and a global dominating set of G. The minimum cardinality of an edge monophonic global dominating set is the edge monophonic global domination number of G and is denoted by $\underline{\gamma_{me}}(G)$. A edge monophonic global dominating set of cardinality $\underline{\gamma_{me}}(G)$ is called a $\underline{\gamma_{me}}$-set of G. The concept of edge monophonic global domination in graphs was introduced in [12].

The following theorem is used in sequel.

Theorem 1.1. [11] Each extreme vertex of a connected graph G belongs to every monophonic global dominating set of G.

2. THE UPPER EDGE MONOPHONIC GLOBAL DOMINATION NUMBER OF A GRAPH

Definition 2.1. An edge monophonic global dominating set S of G is called a *minimal edge monophonic global dominating set* of G if no proper subset of S is an edge geodetic global dominating set of G. *The* *upper edge monophonic global domination number* $\underline{\gamma}^{+}_{me}(G)$ is the maximum cardinality of a minimal edge monophonic global dominating set of G.

Example 2.2. For the graph G given in Figure 2.1, $S_1 = \{v_1, v_4\}$, $S_2 = \{v_2, v_4, v_6, v_7\}$ and $S_3 = \{v_1, v_3, v_5, v_8\}$ are the three minimal edge monophonic global dominating sets of G so that $\underline{\gamma}^{+}_{me}(G) \geq 4$. It is easily verified that there is no minimal edge monophonic global dominating set of cardinality more than four. Therefore $\underline{\gamma}_{me}(G) = 2$ and $\underline{\gamma}^{+}_{me}(G) = 4$.

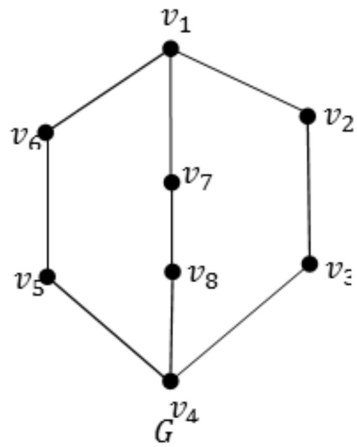


Figure 2.1

Remark 2.3. Every minimum edge monophonic global dominating set of G is a minimal edge monophonic global dominating set of G. But the converse need not be true. For the graph G given in Figure 2.1, $S_3 = \{v_1, v_3, v_5, v_8\}$ is a minimal edge monophonic global dominating set of G. But not a minimum edge monophonic global dominating set of G.

Theorem 2.4. For a connected graph G of order $p \geq 2, 2 \leq \underline{\gamma}_{me}(G) \leq \underline{\gamma}^{+}_{me}(G) \leq p$.

Proof: Any edge monophonic global dominating set needs at least two vertices and so $\underline{\gamma}_{me}(G) \geq 2$. The inequality $\underline{\gamma}_{me}(G) \leq \underline{\gamma}^{+}_{me}(G)$ follows from Remark 2.3. Also since $V(G)$ is an edge monophonic global dominating set and therefore $\underline{\gamma}^{+}_{me}(G) \leq p$. Thus $2 \leq \underline{\gamma}_{me}(G) \leq \underline{\gamma}^{+}_{me}(G) \leq p$.

Remark 2.5. The bounds in Theorem 2.4 are sharp. For the path graph P_4, $\underline{\gamma}_{me}(G) = 2$. For $G = C_4$, $\underline{\gamma}_{me}(G) = \underline{\gamma}^{+}_{me}(G) = 3$. All the inequalities in Theorem 2.4 can be strict. For the graph $G = P_7$ with vertex set $V(G) = \{v_1, v_2, \ldots, v_7\}$, $S_1 = \{v_1, v_4, v_7\}$ and $S_2 = \{v_1, v_3, v_5, v_7\}$ are the only two minimal edge monophonic global dominating sets of G so that $\underline{\gamma}_{me}(G) = 3, \underline{\gamma}^{+}_{me}(G) = 4$ and $p = 7$. Thus $2 < \underline{\gamma}_{me}(G) < \underline{\gamma}^{+}_{me}(G) < p$.

Corollary 2.6. Let G be a connected graph of order $p \geq 2$. If $\underline{\gamma}^{+}_{me}(G) = 2$, then $\underline{\gamma}_{me}(G) = 2$.

Proof: This follows from Theorem 2.4. ■

Remark 2.7. The converse of the Corollary 2.6 need not be true. For the graph G given in Figure 2.1, $\underline{\gamma}_{me}(G) = 2$. But $\underline{\gamma}^{+}_{me}(G) = 4$.

Corollary 2.8. Let G be a connected graph of order $p \geq 2$. If $\underline{\gamma}_{me}(G) = p$. Then $\underline{\gamma}^{+}_{me}(G) = p$.

Proof: This follows from Theorem 2.4. ■

Theorem 2.9. Let G be a connected graph of order $p \geq 2$. Then $\underline{\gamma}^{+}_{me}(G) = 2$ if and only if G is either K_2 or P_4.

Proof. Let $\underline{\gamma}^{+}_{me}(G) = 2$. Then by Corollary 2.6, $\underline{\gamma}_{me}(G) = 2$. Let $S = \{x, y\}$ be a $\underline{\gamma}_{me}$- set of G. If $G = K_2$, then we have done. If $G \neq K_2$, then by Theorem 2.4, $d(x, y) = 3$. Since every edge of G lies on x-y monophonic path, it follows that every vertex of G lies on x-y monophonic path. Let u be a vertex on x-y monophonic path such that $ux \in E(G)$. We prove that x is an end vertex of G. On the contrary suppose that x is not an end vertex of G. Then u is not a cut vertex of G. Then there exists at least two x-y monophonic path. Let $P_1, P_2, \ldots, P_r$ $(r \geq 2)$ be the x-y monophonic path of G. Let $u_i \in V(P_i)$ such that $u_i x \in E(G)$ $(1 \leq i \leq r)$ (u may be any one of $u_1, u_2, \ldots, u_r$). Since $d(u_i, y) = d\left(u_i, u_j\right) = 2$ $(i \neq j)$ $(1 \leq i, j \leq r)$, $S_1 = \{u_1, u_2, \ldots, u_r, y\}$ is an edge monophonic global dominating set of G. Also, since no proper subset of S_1 is an edge monophonic global dominating set of G, S_1 is a minimal edge monophonic global dominating set of G and so $\underline{\gamma}^{+}_{me}(G) \geq r + 1 \geq 3$. which is a contradiction. Therefore x is an end vertex of G. Similarly we can prove that y is an end vertex of G. Since $d(x, y) = 3$, x and y are end vertices of G. Since the x - y monophonic path contains only two internal vertices say u, v with $deg\ deg\ (u)$ $= deg\ deg\ (v)$, we have $G = P_4$. The converse is clear.

Theorem 2.10. Let G be a connected graph of order $p \geq 2$, Then $\underline{\gamma}_{me}(G) = p$ if and only if $\underline{\gamma}^{+}_{me}(G) = p$.

Proof: Let $\underline{\gamma}_{me}(G) = p$. Then by Corollary 2.8, $\underline{\gamma}^{+}_{me}(G) = p$. Conversely assume that $\underline{\gamma}^{+}_{me}(G) = p$. Then the set of all vertices of G is the unique minimal edge monophonic global dominating set of G. It follows that G contains no proper edge monophonic global dominating sets and so the set of all vertices is the minimum edge monophonic global dominating set of G. Hence $\underline{\gamma}_{me}(G) = p$. ■

Corollary 2.11. For the complete graph $G = K_p$ $(p \geq 2)$, $\underline{\gamma}_{me}(G) = \underline{\gamma}^{+}_{me}(G) = p$.

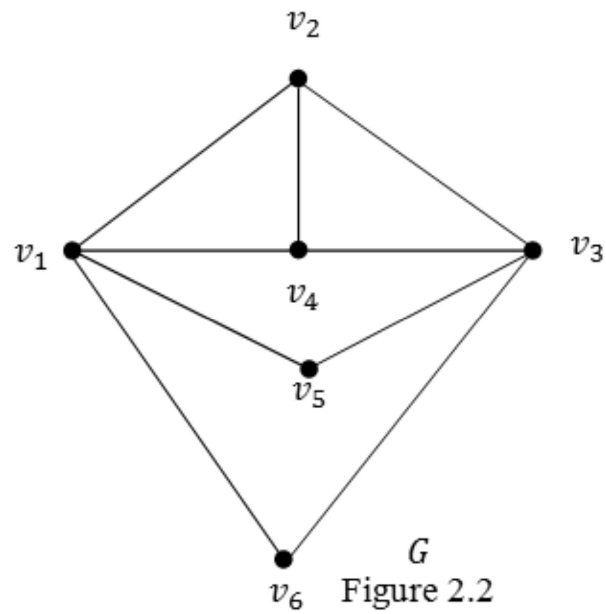
Proof: This follows from Theorem 2.4. ■

Theorem 2.12. Let G be a connected graph of order $p \geq 2$. If $\underline{\gamma}_{me}(G) = p - 1$ then $\underline{\gamma}^{+}_{me}(G) = p - 1$.

Proof: Let $\underline{\gamma}_{me}(G) = p - 1$. Then by Theorem 2.4, either $\underline{\gamma}_{me}^{+}(G) = p$ or $p - 1$. If $\underline{\gamma}_{me}^{+}(G) = p$, then by Theorem 2.10, $\underline{\gamma}_{me}(G) = p$, which is a contradiction to our assumption. Therefore $\underline{\gamma}_{me}^{+}(G) = p - 1$.

■

Remark 2.13. The converse of Theorem 2.12 need not be true. For the graph G given in Figure 2.2, $S_1 = \{v_1, v_2, v_3, v_4\}$, $S_2 = \{v_1, v_2, v_4, v_5, v_6\}$ and $S_3 = \{v_2, v_3, v_4, v_5, v_6\}$ are the only three minimal edge monophonic global dominating set of G so that $\underline{\gamma}_{me}(G) = 4 = p - 2$ and $\underline{\gamma}_{me}^{+}(G) = 5 = p - 1$.



G
Figure 2.2

Theorem 2.14. For the complete bipartite graph $G = K_{m,n}$ $(2 \leq m \leq n)$, $_{me}^{+}(G) = \{m, n\} + 1$.

Proof. Let $X = \{x_1, x_2, \ldots, x_m\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ be the two bipartition sets of G. Let $m < n$. Then $S = X$ and $S_1 = Y$ are the are the minimal edge monophonic dominating sets of G. Since S and S_1 are not global dominating sets of G. S and S_1 are not a minimal edge monophonic global dominating sets of G and so $\underline{\gamma}_{me}(G) \geq n + 1$. Let $S_2 = S_1 \cup \{x_1\}$ and $S_3 = S \cup \{y_1\}$. Then S_2 and S_3 are minimal edge monophonic global dominating sets of G, $\underline{\gamma}_{me}(G) \geq n + 1$. We prove that $\underline{\gamma}_{me}^{+}(G) = n + 1$. On the contrary suppose that $\underline{\gamma}_{me}^{+}(G) \geq n + 2$. Then there exists a minimal edge monophonic global dominating set S' such that $|S'| \geq n + 2$. Hence it follows that $S' \subset X \cup Y$. Let $S' = S_1' \cup S_2'$ where $S_1' \subset X$ and $S_2' \subset Y$. Let $x, y \in V \backslash S_1'$ such that $x \in X \backslash S_1'$ and $y \in Y \backslash S_2'$. Then $xy \notin I_e[S']$, which is a contradiction to S' an edge monophonic global dominating set of G. Therefore $\underline{\gamma}_{me}^{+}(G) = n + 1$. ■

Theorem 2.15. For every pair of integers a and b with $3 \leq a \leq b$, there exists a connected graph G such that $\underline{\gamma}_{me}(G) = a$ and $\underline{\gamma}_{me}^{+}(G) = b$.

Proof: Let $H = K_{3,b-a+2}$ with bipartite sets $U = \{u_1, u_2, u_3\}$, $V = \{v_1, v_2, \dots, v_{b-a+2}\}$. Let G be a graph obtained from H by adding new vertices $z_1, z_2, \dots, z_{a-2}$ and introducing the edge $u_3 z_i (1 \leq i \leq a - 2)$. The graph G is shown in Figure 2.3.

First we prove that $\underline{\gamma}_{me}(G) = a$. Let $Z = \{z_1, z_2, \dots, z_{a-2}\}$ be the set of all end vertices of G. Then by Theorem 1.1 (i), Z is a subset of every edge monophonic global dominating set of G and so $\underline{\gamma}_{me}(G) \geq a - 2$. Since $I_e[Z] \neq E(G)$, Z is not an edge monophonic global dominating set of G and so $\underline{\gamma}_{me}(G) \geq a$. It is easily verified that $Z \cup \{x\}$, where $x \notin Z$ is not an edge monophonic global dominating set of G and so $\underline{\gamma}_{me}(G) \geq a$. Let $S = Z \cup \{u_1, u_2\}$. Then S is an edge monophonic global dominating set of G so that $\underline{\gamma}_{me}(G) = a$.

Next we prove that $\underline{\gamma}^{+}_{me}(G) = b$. Let $S' = Z \cup \{v_1, v_2, \dots, v_{b-a+2}\}$. Then S' is an edge monophonic global dominating set of G. We prove that S' is a minimal edge monophonic global dominating set of G. On the contrary, suppose that S' is not a minimal edge monophonic global dominating set of G. Then there exists an edge monophonic global dominating set M such that $M \subset S'$. By Theorem 1.1 (i), $Z \subset M$. Let $x \in S'$ such that $x \notin M$. Then $x = v_i$ for some $i\ (1 \leq i \leq b - a + 2)$. Hence it follows that $xu_i \notin I_e[M]$ for $1 \leq i \leq 3$. Therefore M is not an edge monophonic global dominating set of G, which is a contradiction. Therefore S' is a minimal edge monophonic global dominating set of G and so $\underline{\gamma}^{+}_{me}(G) \geq b$. We prove that $\underline{\gamma}^{+}_{me}(G) = b$. On the contrary, suppose that $\underline{\gamma}^{+}_{me}(G) \geq b + 1$.

Then there exists a minimal edge monophonic global dominating set M' of G such that Y. $|M'| \geq b + 1$. Then $M' \subset X \cup Y \cup Z$. By Theorem 1.1 (i), $Z \subset M'$. Let $u \in X$ and $v \in$Then $uv \notin I_e[M']$. By the similar way if $u = u_2$, then $uv \notin I_e[M']$, which is a contradiction. If $u \in Y \backslash S'$ and $v \in X$, then $uv \notin I_e[M']$, which is a contradiction. Therefore $\underline{\gamma}^{+}_{me}(G) = b$.

■

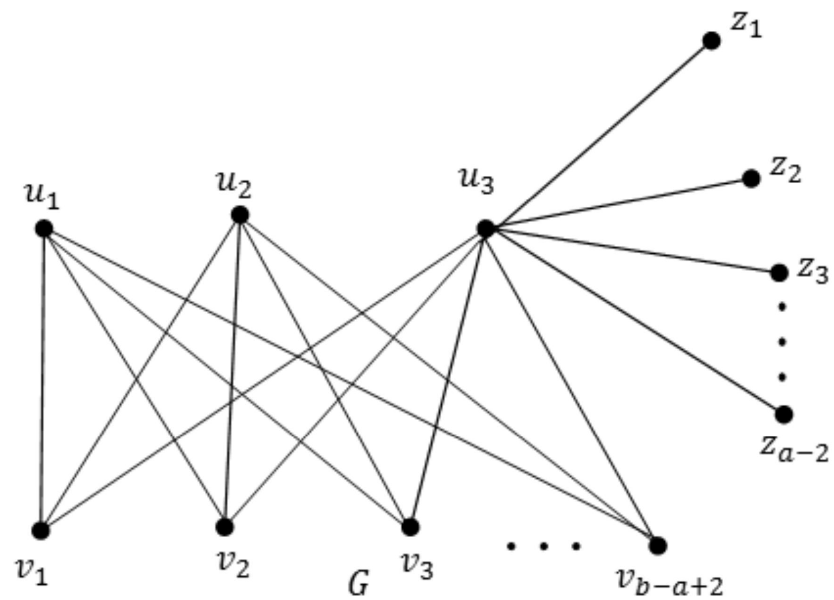


Figure 2.3

CONCLUSION

In this paper, we introduced the upper edge monophonic global domination number of a graph and determined it for some standard graphs. Connected graphs of order p?2 with the upper edge monophonic global domination number 2 or ?? are characterized. We will introduce new parameters connecting the edge monophonic global domination number of a graph with some conditions in future work.

REFERENCES

1. P.Arul Paul Sudhahar and C.Bency, The forcing restrained monophonic number of a graph, International Journal of Applied Engineering Research, 13(16), (2018), 12449-12452.
2. F.Buckley and F.Harary, Distance in Graph, Addition-Wesly-Redwood city, CA, (1990).
3. G.Chartrand, H.Gavlas, R. C. Vandell and F. Harary, The forcing domination number of a graph, Journal of Combinatorial Mathematics and Combinatorial Computing, 25 (1997), 161- 174.
4. Haynes.T.W, Hedetniemi.S.T and Slater, P.J, Fundamentals of Domination in Graphs, Marcel Dekker, New York, (1998).
5. J.John and P.A.P.Sudhahar, The forcing edge monophonic number of a graph, SCIENTIA Series A: Mathematical Sciences, 23, (2012), 87-98.
6. J.John and S.Panchali, The forcing monophonic number of a graph, International Journal of Mathematical Archives, (3), (2012), 935 - 938.

7. J.John and V.Mary Gleeta, The forcing monophonic hull number of a graph, International Journal of Mathematics Trends and Technology, 3(2), (2012), 43- 46.
8. J.John, P. Arul Paul Sudhahar and D.Stalin, On the (M,D) number of a graph, Proyecciones Journal of Mathematics,38,2, (2019), 255-266.
9. J.John, The forcing monophonic and forcing geodetic numbers of a graph, Indonesian Journal of Combinatorics, 4 (2) (2020),114-125.
10. E.Sampathkumar, The global domination number of a graph, J. Math. Phys. Sci., 23, (1989), 377-385.
11. V.Selvi and V.Sujin Flower, The Monophonic Global Domination Number of a Graph, J.Math. Comput.Sci, 11(5)(2021),6007-6017.
12. V.Selvi and V.Sujin Flower, The Edge Monophonic Global Domination Number of a Graph, Conference Proceeding of Recent Trends in Modern Mathematics, RTMM-(Volume 2) (2021), 43-48.

DOI No. : 10.5958/2394-9309.2022.00046.4
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ODD AVERAGE HARMONIOUS LABELING OF DISJOINT UNION OF GRAPHS

G. S. Ganeshwari*, G. Sudhana**
*Research Scholar (Reg No: 20123112092025), Department of Mathematics, Nesamony Memorial Christian College, Marthandam 629 165, Tamil Nadu, India
**Assistant Professor, Department of Mathematics, Nesamony Memorial Christian College, Marthandam – 629165, Tamil Nadu, India
[1,2] Affiliated to ManonmaniamSundaranar University, Abishekapatti, Tirunelveli -627 012, Tamil Nadu, India.
E-mail:ganeswari247.gs@gmail.com, sudhanaarun1985@gmail.com

ABSTRACT

The current paper looks into a new concept of an odd average harmonious graph of disjoint union of graphs. The odd average harmonious graph G with n vertices and m edges if there exists an injective function $L: V \to \{0, 1, 2 \ldots, 2n-1\}$ and the induced function $L^*: E \to \{0, 1, 2, \ldots, m-1\}$ is distinguished by $L^*(uv) = \left\lceil \frac{L(u)+L(v)+1}{2} \right\rceil \ (mod m)$ is a bijection, the resulting edge labels should be distinct. A graph which admits an odd Average harmonious labeling is called odd Average harmonious graph. We prove that the graph $C_n \cup P_3$, $C_4 \cup C_{bn}$, $C_n \cup H_{q,q}$, $C_n \cup K_{p,q}$, $C_n \cup S_p$, $K_{m,n} \cup k_{p,q}$, $C_n \cup K_{p,q}$ is an odd average harmonious graph graph and also the line graph $K_{p,q}$, $H_{q,q}$ and C_{bn} are not an odd average harmonious graph.

1. INTRODUCTION

Subject to certain conditions, in graph labeling, integers are assigned to vertices, edges, or both. Labeled graphs are useful a variety of mathematical models for a variety of applications, including theorem of Coding. In the field of graph theory, labeling a graph is a vast and enormous field of study. It is concerned with how the vertices and graph's edges are 1 labeled in relation to a few mathematical conditions [1]. A graph labeling, also known as a graph valuation, is a map that carries graph elements onto numbers known as labels. Graham and Sloane's [3] investigation of modular variants of additive base problems arising from erroneous coding resulted in the formation of harmonious graphs. We are particularly interested in the harmonious labeling of graphs. This type of graphs labeling has many applications, including social networking, a high probability event, resolving typical coloring algorithm problems, and network of transmission. Harmonious graphs are also very interesting. [5,6] introduced both odd and even harmonious labeling. Graphs that are odd harmonious and have applications were discussed by P. Jeyanthi, S.Philo[4] and S.K. Vaidya, N H Shah [8,9]. Gallian [1] collects and updates the results of a survey on graph labeling on a regular basis. If $f : V \rightarrow \{0, 1, 2, \ldots, q + p\}$ is injective and the induced function $f^* : E \rightarrow \{0, 1, \ldots, (q-1)\}$ is defined as $f^*(uv) = \begin{cases} f(u) + f(v) + 1 \ (mod\ m) & if \quad f(u) + f(v)\ isodd \\ f(u) + f(v)(mod\ m) & if \quad f(u) + f(v)\ is\ even \end{cases}$ is bijective, the resulting edge labels should be distinct. A graph which allows an average harmonious labeling is called Average harmonious graph.[2] G is an Even odd average harmonious graph if it is a graph with n vertices and m edges if there is a bijective function $f : V(G) \rightarrow \{1,3,5, \ldots (2n-1)\}$ as the induced mapping $f^* : E(G) \rightarrow \{0,1,2, \ldots . m-1)$characterized as$f^*(uv) = \frac{f(u)+f(v)}{2}\ (mod\ m)$ is a bijection. The current paper looks into a new concept of an odd average harmonious graph of disjoint union of graphs. We prove that the graph $C_n \cup P_3$, $C_4 \cup C_{bn}$, $C_n \cup H_{q,q}$, $C_n \cup K_{p,q}$, $C_n \cup S_p$, $K_{m,n} \cup k_{p,q}$, $C_n \cup K_{p,q}$ is an odd average harmonious graph and also the line graph $K_{p,q}$, $H_{q,q}$and C_{bn} are not an odd average harmonious graph.

2. PRELIMINARIES

Definition 2.1:

The graph $H_{n,n}$ has the vertex set $V = \{u_i, v_i\colon\ 1 \leq i \leq n\}$ and the edge set $E = \{ u_i v_j\colon\ 1 \leq i \leq n,\ n - i + 1 \leq j \leq n\}$.

Definition 2.2:

A corona graph$C_n \odot K_1$ is formed by taking one copy of G, which is assumed to have p vertices, and p copies of the graph $H$and connecting the k^{th} vertex in the k^{th}copy of H with an edge to put it another way, given two graphs G and H, the corona of G with H denoted by $C_n \odot K_1$ is the graph with vertex setv_i of G with every vertex of H_i for $1 \leq i \leq n$

A comb is a graph formed by connecting each vertex of a path with $P_n \odot K1$ denoted by C_{bn}.

Definition 2.3:[2]

A bipartite simple graph,where $|v_1| = m\ and\ |v_2| = n$, is called a complete bipartite graph and it is denoted by $K_{m,n}$,if its size is mn.That is ,if each vertex v_1 is adcajent to all vertices in v_2 and each vertex v_2 is adcajent to all vertices in v_1.

Definition2.4:[7]

A path P of length$n$ in a graph G is a sequence of distinct vertices $v_t, 1 \leq t \leq n$in G where $v_t\ v_{t+1}$is an edge for every $i = 0,1, \ldots, n-1$.We also say that P is a $v_0 v_n$–path

Definition 2.5:

A cycle is a simple graph with n vertices and m edges that form a cycle of length $'n'$.

All of the vertices in a cycle graph are of degree 2.

Remark: Remark: For the sake of clarity, we use " OAHG" instead of "odd average harmonious graph labeling" in this paper.

3. ODD AVERAGE HARMONIOUS LABELING OF DISJOINT UNION OF GRAPHS

We will go over the definition of OAHG Labeling in this section.

Definition:

The odd average harmonious graph G with n vertices and m edges if there exists an injective function $L: V \to \{0,1,2...,2n-1\}$and the induced function $L^* : E \to \{ 0, 1, 2, \dots, m - 1 \}$ is distinguished by $L^*(uv)=\left\lceil \frac{L(u)+L(v)+1}{2} \right\rceil \ (mod m)$ is a bijection, the resulting edge labels should be distinct. A graph which admits an odd Average harmonious labeling is called odd Average harmonious graph.

4. MAIN RESULT

Theorem 4.1. The graph $C_n \cup P_3$ is an OAHG Labeling.

Proof: Let$U = \{u_t,\ 1 \leq t \leq n\}$represent the vertices of the C_n cycle$and\ V = \{v_t\ , 1 \leq t \leq 3\}$ represent the vertex set of the path P_3.

$E\ (C_n \cup P_3) = \{ u_t\, u_{t+1} : 1 \leq t \leq n - 1,\ , u_1\, u_n\ and\ \ v_t\, v_{t+1}, t = 1,2 \}$

The graph contains$n = p + 3$ vertices and $m = n + 2$ edges .

Define a injective function $L : V(C_n \cup P_3) \to \{ 0, 1, 2, \dots, 2(n+2) - 1 \}$ such that

$L(u_t) = t - 1 :\ 1 \leq t \leq \frac{n}{2}$

$L(u_t) = \begin{cases} t+1 & if\ \ t\ is\ odd \\ t-1 & if\ t\ is\ even \end{cases}, \frac{n}{2} + 1 \leq t \leq n$

$L(v_t) = 2n + t - \frac{n}{2} : t = 1,3\ , L(v_2) = \frac{n}{2}$

As a result, both the vertex labels and the edge labels are distinct. $L^* : E(C_n \cup P_3) \to \{ 0, 1, 2, \dots, m - 1 \}$ by

$$L^*(u_t\, u_{t+1}) = \begin{cases} t & if\ \ 1 \leq t \leq \frac{n}{2} - 1 \\ t+1 & if\ \frac{n}{2} \leq t \leq n-1 \end{cases}$$

$L^*(u_1\, u_n) = \frac{n}{2}, L^*(v_t\, v_{t+1}) = (n+t) mod m, \quad t = 1,2.$

As a result, both the vertex labels and the edge labels are distinct.Given the above-mentioned labeling pattern, we can conclude that $C_n \cup P_3$is an OAHG Labeling.

Example 4.2.

An OAHG Labeling of.$C_n \cup P_3$is shown in fig 1.

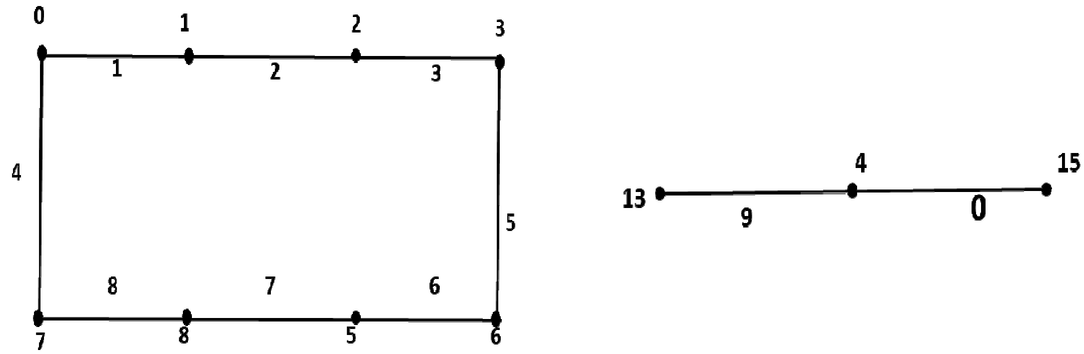


Fig 1.$C_n \cup P_3$ is an OAHG Labeling.

Theorem.4.3. The graph $C_4 \cup C_{bn}$ is OAHG Labeling.

Proof: Let$U = \{u_t, 1 \leq t \leq 4\}$represent the vertices of the C_4cycle and $V = \{v_t, 1 \leq t \leq n\}$, $W = \{w_t, 1 \leq t \leq n\}$ represent the vertex set of the comb C_{bn}

$E(C_4 \cup C_{bn}) = \{ u_1u_2, u_4u_1 \text{ and } u_t u_{t+1} : t = 2,3\}$ be the edges of cycle C_4 and

$E(C_4 \cup C_{bn}) = \{ v_t v_{t+1} : 1 \leq t \leq n-1, , v_t w_t, 1 \leq t \leq n-1 \text{ and } v_n w_n\}$

The graph contains$n = 2(2+p)$ vertices and $q = 3 + 2n$ edges.

Define a injective function $L : V(C_4 \cup C_{bn}) \to \{0, 1, 2, \ldots, 2(3+2n) - 1\}$ such that

$L(u_t) = t - 1, t = 1,2$

$$L(u_t) = \begin{cases} t+1 & if\ t\ is\ odd \\ t-1 & if\ t\ is\ even \end{cases}, 3 \leq t \leq 4$$

$$L(v_t) = \begin{cases} 2t & : 1 \leq t \leq n \quad if\ t\ is\ odd \\ 2t+3 & : 1 \leq t \leq n \quad if\ t\ is\ even \end{cases}$$

Case 1. If n is odd.

$$L(w_t) = \begin{cases} 2t+7 & : 1 \leq t \leq n-2 \quad if\ t\ is\ odd \\ 2t+4 & : 1 \leq t \leq n-1 \quad if\ t\ is\ even \end{cases}$$

$L(w_n) = 2n + 5$

Case 2. If n is odd.

$$L(w_t) = \begin{cases} 2t+7 & : 1 \leq t \leq n-1 \quad if\ t\ is\ odd \\ 2t+4 & : 1 \leq t \leq n-2 \quad if\ t\ is\ even \end{cases}$$

$L(w_n) = 2n + 2$

As a result, both the vertex labels and the edge labels are distinct. $L^* : E(L_n) \to \{0, 1, 2, \ldots, m - 1\}$ by

$L^*(u_t u_{t+1}) = t + 1, \quad \frac{n}{2} \leq t \leq n - 1$

$L^*(u_1 u_2) = 1, L^*(u_1 u_4) = 2$

$L^*(v_t w_t) = 4 + 2t, \quad 1 \leq t \leq n - 1$

$L^*(v_n w_n) = (3 + 2n)(mod m), L^*(v_t v_{t+1}) = 2t + 3, \quad 1 \leq t \leq n - 1.$

As a result, both the vertex labels and the edge labels are distinct. Given the above-mentioned labeling pattern, we can conclude that $C_4 \cup C_{bn}$is an OAHG Labeling.

Example .4.4.

An OAHG Labeling of.$C_4 \cup C_{bn}$ is shown in fig.2.

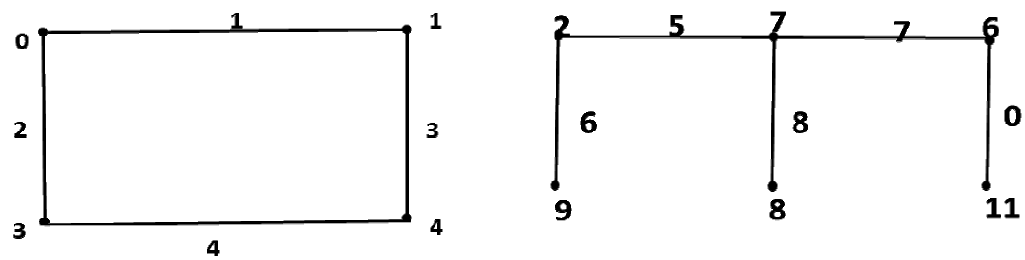


Fig 2. $C_4 \cup C_{bn}$ is an OAHG Labeling.

Theorem 4.5. The graph $C_n \cup H_{q,q}, \frac{n}{4} + 1 \ \ q$ is an OAHG Labeling.

Proof:

Let $U(C_n \cup H_{q,q}) = \{u_t, \ 1 \le t \le n\}$represent the vertices of the C_n cycle and $V(C_n \cup H_{q,q}) = \{v_t, 1 \le t \le p \ and \ w_t, 1 \le t \le q\}$ represent the vertex set of the graph $H_{q,q}$.

$E(C_n \cup H_{q,q}) = \{u_t u_{t+1} : 1 \le t \le n-1, , u_1 u_n \ and \ v_t w_s, 1 \le t \le p \ and \ 1 \le s \le q\}$

The graph contains$n = p + 2m$ vertices and $m = n + \frac{q(q+1)}{2}$ edges.

Define a injective function $L : V(C_n \cup H_{q,q}) \to \{0,1,2,\dots,2(n + \frac{q(q+1)}{2}) - 1\}$ such that

$$L(u_t) = t - 1 : \ 1 \le t \le \frac{n}{2}$$

$$L(u_t) = \begin{cases} t+1 & if \ t \ is \ odd \\ t-1 & if \ t \ is \ even \end{cases}, \frac{n}{2} + 1 \le t \le n$$

$$L(v_t) = (q + t - 1)(q - t) + \frac{n}{2} \ : 1 \le t \le q - 1$$

$$L(w_t) = 2(t-1) - \frac{n}{2} + 2n + 1 \ : 1 \le t \le q$$

As a result, both the vertex labels and the edge labels are distinct. $f^* : E(C_n \cup H_{q,q}) \to \{0, 1, 2, \dots, m - 1\}$ by

$$L^*(u_t u_{t+1}) = \begin{cases} t & if \ \ 1 \le t \le \frac{n}{2} - 1 \\ t+1 & if \ \frac{n}{2} \le t \le n-1 \end{cases}$$

$$L^*(u_1 u_n) = \frac{n}{2}$$

$$L^*(v_t w_s) = [\tfrac{n}{2}[(q + t - 1)(q - t)] + (n + s)](mod m), \ : 1 \le t \le p \ and \ 1 \le s \le q.$$

As a result, both the vertex labels and the edge labels are distinct. Given the above-mentioned labeling pattern, we can conclude that $C_n \cup H_{q,q}$is an OAHG Labeling.

Example 4.6.

An OAHG Labeling of.$C_n \cup H_{q,q}$is shown in fig 3

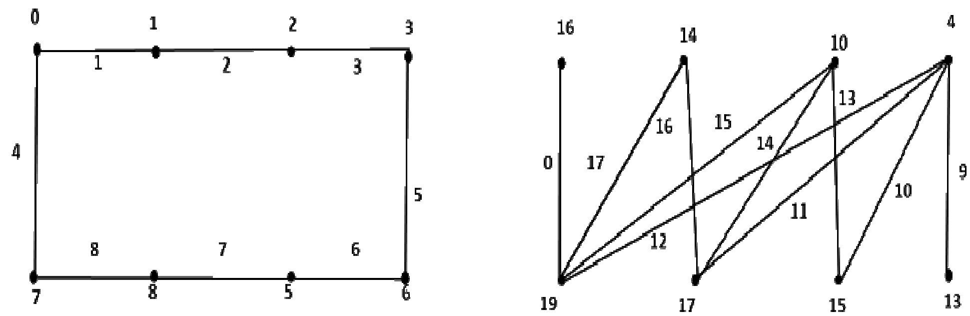


Fig 3.$C_n \cup H_{q,q}$ is an OAHG Labeling.

Theorem 4.7. The graph $C_n \cup K_{p,q}$ is an OAHG Labeling.

Proof: Let $U(C_n \cup K_{p,q}) = \{u_t,\ 1 \le t \le n\}$ be the vertices of cycle $C_n and$ $V(C_n \cup K_{p,q}) = \{v_t, 1 \le t \le p\ and\ w_t, 1 \le t \le q\}$ represent the vertex set of the complete bipartite graph$K_{p,q}$
$E\,(C_n \cup K_{p,q}) = \{u_t\, u_{t+1}\ : 1 \le t \le n-1,\ , u_1\, u_n\ and\ \ v_t\, w_s, 1 \le t \le p\ and\ 1 \le s \le q\}$
The graph contains$n + p + q$ vertices and $m = n + pq$ edges.
Define a injective function $L : V(C_n \cup K_{p,q}) \rightarrow \{0, 1, 2, \dots, 2(n + pq) - 1\}$ such that

$$L(u_t) = t - 1 : \ 1 \le t \le \frac{n}{2}$$

$$L(u_t) = \begin{cases} t+1 & if\ \ t\ is\ odd \\ t-1 & if\ t\ is\ even \end{cases}, \frac{n}{2} + 1 \le t \le n$$

$$L(v_t) = 2q(t-1) + \frac{n}{2} \quad : 1 \le t \le p$$

$$L(w_t) = 2(t-1) - \frac{n}{2} + 2n + 1 \quad : 1 \le t \le q$$

As a result, both the vertex labels and the edge labels are distinct. $f^* : E(C_n \cup K_{p,q}) \rightarrow \{0,\ 1,\ 2, \dots, m - 1\}$ by

$$L^*(u_t\, u_{t+1}) = \begin{cases} t & if \quad 1 \le t \le \frac{n}{2} - 1 \\ t+1 & if\ \frac{n}{2} \le t \le n-1 \end{cases}$$

$$L^*(u_1\, u_n) = \frac{n}{2}$$

$$L^*(v_t\, w_s) = \big((t-1)q + n + s\big)(mod m), \quad : 1 \le t \le p\ and\ 1 \le s \le q.$$

As a result, both the vertex labels and the edge labels are distinct. Given the above-mentioned labeling pattern, we can conclude that $C_n \cup K_{p,q}$is an OAHG Labeling.

Example.4.8.

An OAHG Labeling of.$C_n \cup K_{p,q}$is shown in fig 4.

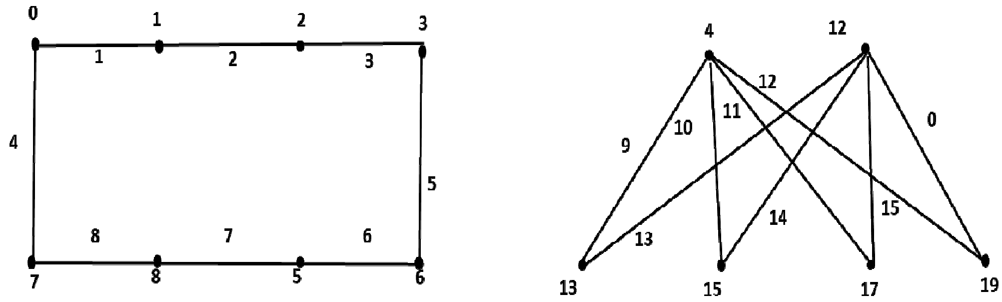


Fig 4.$C_n \cup K_{p,q}$ is an OAHG Labeling.

Theorem 4.9. The graph $C_n \cup S_p$ is an OAHG Labeling.

Proof: Let $U(C_n \cup S_p) = \{u_t,\ 1 \le t \le n\}$ be the vertices of cycle $C_n and\ V(C_n \cup K_{p,q}) = \{v_0,\ v_t\ , 1 \le t \le m\ \}$ represent the vertex set of the star graph s_p

$E\ (C_n \cup S_p) = \{\ u_t\ u_{t+1}\ : 1 \le t \le n - 1,\ , u_0\ u_t\ \}$

The graph contains$n + m + 1$ vertices and $m = p + n$edges.

Define a injective function $L :\ V(C_n \cup S_p) \rightarrow \{\ 0, 1, 2, \ldots, 2(\ n + p) - 1\ \}$ such that

$L(u_t) = t - 1 :\ 1 \le t \le \frac{n}{2}$

$L(\ u_t) = \begin{cases} t + 1 & if\ \ t\ is\ odd \\ t - 1 & if\ t\ is\ even \end{cases}, \frac{n}{2} + 1 \le t \le n$

$L(\ u_0) = \frac{n}{2}, L(\ u_t) = 2(t - 1) - \frac{n}{2} + 2n + 1\ \ : 1 \le t \le p$

As a result, both the vertex labels and the edge labels are distinct. $f^* : E(C_n \cup S_p) \rightarrow \{\ 0,\ 1,\ 2, \ldots, m - 1\ \}$ by

$$L^*(\ u_t\ u_{t+1}) = \begin{cases} t & if\ \ 1 \le t \le \frac{n}{2} - 1 \\ t + 1 & if\ \frac{n}{2} \le t \le n - 1 \end{cases}$$

$L^*(\ v_0\ v_t) = (n + t)(mod m) : 1 \le t \le p$

As a result, both the vertex labels and the edge labels are distinct. Given the above-mentioned labeling pattern, we can conclude that $C_n \cup S_p$ is an OAHG Labeling.

Example.4.10

An OAHG Labeling of.$C_n \cup S_p$is shown in fig 5.

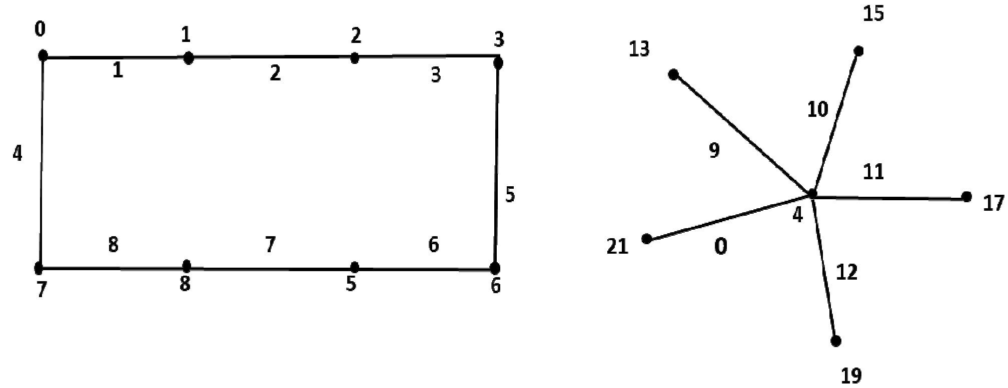


Fig 5.$C_n \cup S_p$is an OAHG Labeling.

Theorem 4.11.

The graph $K_{m,n} \cup k_{p,q}$ is an OAHG Labeling .

Proof:

Let $U(K_{m,n} \cup k_{p,q}) = \{u_t,\ v_t : 1 \leq t \leq m, 1 \leq t \leq n\}$ be the vertices of $K_{m,n} and\ V(K_{m,n} \cup k_{p,q}) = \{r_t, 1 \leq t \leq p\ and\ w_t, 1 \leq t \leq q\}$ represent the vertex set of the graph $k_{q,q}$.

$E\left(K_{m,n} \cup k_{p,q}\right) = \{u_t\, v_s : 1 \leq t \leq m, 1 \leq s \leq n\ ,\ and\ r_t\, w_s, 1 \leq t \leq p\ and\ 1 \leq s \leq q\}$

The graph contains$n + m + p + q$ vertices and $mn + pq$ edges.

Define a injective function $L : V(K_{m,n} \cup k_{p,q}) \rightarrow \{0, 1, 2, \ldots, 2(nm + pq) - 1\}$ such that

$L(u_t) = 2n(t-1) :\ 1 \leq t \leq m$

$L(v_t) = 1 + 2(t-1),\quad 1 \leq t \leq n$

$L(r_t) = 2[m + q(t-1)]\ :\ 1 \leq t \leq p$

$$L(w_t) = 2m(n-1) + 2(t-1) + 1\ :\ 1 \leq t \leq q$$

As a result, both the vertex labels and the edge labels are distinct. $f^* : E(K_{m,n} \cup k_{p,q}) \rightarrow \{0, 1, 2, 3, \ldots, m - 1\}$ by

$L^*(u_t\, v_s) = n(t-1) + 1(s-1) : 1 \leq t \leq m\ and\ 1 \leq s \leq n$

$L^*(r_t\, w_s) = q(t-1) + m[(n-1) + 1] + (s-1) + 1\ :\ 1 \leq t \leq p\ and\ 1 \leq s \leq q.$

As a result, both the vertex labels and the edge labels are distinct. Given the above-mentioned labeling pattern, we can conclude that $K_{m,n} \cup k_{p,q}$ is an OAHG Labeling.

Example.4.12.

An OAHG Labeling of.$K_{m,n} \cup k_{p,q}$is shown in fig

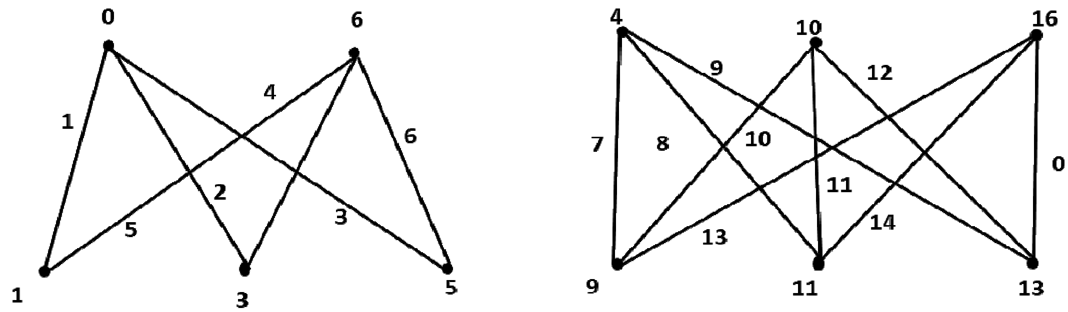


Fig 6. $K_{m,n} \cup k_{p,q}$ **is an OAHG Labeling.**

Theorem 4.13.

The line graph of $K_{m,n}$, $m, n > 2$, $H_{n,n}$ and C_{bn}, $n > 2$ are not OAHG Labeling .

Proof: Since the line graphs $L(K_{m,n})$, $m, n > 2$, $L(H_{n,n})$ and $L(C_{bn})$, $n > 2$ hence they are not odd average harmonious graph.

Theorem 4.14.

The line graph of path P_n, C_n ($n \equiv 0 (mod\ 4)$), $k_{2,2}$, $H_{2,2}$ and C_{b2} are not OAHG Labeling .

Proof: Since $L(P_n) = P_{n-1}$, $L(C_n) = C_n$, ($n \equiv 0(mod\ 4)$), $L(C_{b2}) = P_2$, $L(H_{2,2}) = P_2$, $L(K_{2,2}) = C_4$ contains atleast one odd cycle hence they are not odd average harmonious graph.

CONCLUSION

This paper examines an odd average harmonious labeling, which is one of the most important labeling techniques. Because not all graphs admit the odd average harmonious, it is very interesting to investigate the various types of graphs that do.. Using some mathematical derivations, we have reported the odd average harmonious labeling of various graphs.We have proved that,the graph $C_n \cup P_3$, $C_4 \cup C_{bn}$, $C_n \cup H_{q,q}$, $C_n \cup K_{p,q}$, $C_n \cup S_p$, $K_{m,n} \cup k_{p,q}$, $C_n \cup K_{p,q}$ is an odd average harmonious graph and also the line graph $K_{p,q}$, $H_{q,q}$ and C_{bn} are not an odd average harmonious graph.

REFERENCES

1. *J.A.Gallian, A dynamic Survey of Graph Labeling, The Electronics Journal of Combinatorics, 2017, # DS6.*
2. *Ganeshwari G S and Sudhana G, Even Odd Average Harmonious Graphs,Proceedings of International Conference on Recent Trends and Techniques in Computer Science(2021),E-ISBN:978-93-915532-7-2,165- 176.*

3. *R. L. Graham and N. J. A. Sloane, On Additive bases and Harmonious Graphs, SIAM J. Algebr. Disc. Meth., 4(1980), 382C-404.*
4. *P. Jeyanthi, S.Philo, Odd Harmonious Labeling of Some Cycle Related Graphs, Proyecciones Journal of Mathematics, Vol.35, No.1 (2016), 85–98.*
5. *Z.Liang, Z.Bai, On the Odd Harmonious Graphs with Applications, J.Appl. Math.Comput., 29 (2009), 105C-116*
6. *N. AdalinBeatress and P. B. Sarasija, Average Harmonious Graphs, Annals of Pure and Applied Mathematics Vol. 11, 7-10 (2016).*
7. *M. Kalaimathi, B.J Balamurugan, Computation of Even-Odd Harmonious Labeling of Graphs Obtained by Graph Operations, Recent Trends in Pure and Applied Mathematics, AIP Conf. Proc. 2177, 020030-1 - 020030-9, (2019).*
8. *S.K.Vaidya and N. H. Shah, Some New Odd Harmonious Graphs, International Journal of Mathematics and Soft Computing, 1(2011), 9C-16.*
9. *S.K.Vaidya, N. H. Shah, Odd Harmonious Labeling of Some Graphs , International J.Math. Combin., 3(2012), 105C-112.*

DOI No. : 10.5958/2394-9309.2022.00047.6
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

ARTIFICIAL NEURAL NETWORK FOR PREDICTION

Suchithra P* and S. Santhana Lakshmi**
*Suchithra P, Research Scholar, Department of Statistics, Presidency College, Chennai. Email
**S. Santhana Lakshmi, Assistant Professor, PG Department of Statistics, SDNB Vaishnav College, Chennai.
Email Id: spprps111996@gmail.com, swarsakthi@gmail.com

ABSTRACT

Artificial intelligence (AI) has become a buzzword that we hear and see every day, everywhere. I'm wondering how the term "Artificial Intelligence" is used in this current trend across various platforms according to business needs. This made my interest in working on Artificial Intelligence which can mimic the human brain to learn the data and make it possible through Machine Learning techniques. So, I liked to make use of this concept to explore the purpose of predicting the house price value of residential properties, either to evaluate the purchase price or to evaluate the sales price of an individual home. Prediction is an essential part of planning and research. We can do this technique using an Artificial Neural Network. It is a modern concept, which is similar to the regression model (in our research), but a Machine Learning technique.

The research aims to introduce Neural Network with some background and to check how well the Neural Network perform. To proceed with my research, I took "Zillow's Home Value Prediction data" from Kaggle. For an individual, a house is the largest and most expensive one-time purchase where they invest for a lifetime. Hence, consumers who purchase or sell validate the market price through data science with real-time information and help predict future prices. Using Python, the study attempted to check how the neural network was performing in prediction. In the post, the research leads to the conclusion that "The Neural Network performs well for predicting house value."

Keywords : Artificial Intelligence, Neural Network, Prediction

I. INTRODUCTION

Prediction is determining what is going to happen in the present or future by analysing what happened in the past. We usually use regression equations to make predictions. When a company uses a statistical technique to predict, it uses its historical sales or demand data to predict future sales. Because of the complex formulas used for prediction, most companies rely on advanced software to accomplish this task. In the current trend, a company would like to use advanced techniques such as artificial intelligence (AI). Artificial Intelligence is an approach to make a computer or a robot think, the way a smart human think. This approach includes statistical methods, computational intelligence, and traditional coding AI. Machine Learning is a subset of Artificial Intelligence that includes an abstruse statistical technique that enables machines to improve at tasks with experience.

An Artificial Neural Network (ANN) is one set of algorithms used in machine learning for modelling data using graphs of Neurons. A Neural Network is a flexible model that adapts itself to the shape of the data. It is the largest digital inventory and estimation of American homes in the world. The value of the American housing stock exceeds $27.5 trillion dollars. Zillow covers 110 million homes throughout the United States with 103 variables per house. The dataset can be either supervised or unsupervised, and ANN can be done in both the situations. It is described as a class of information processing systems that exhibit the capacity and capability to learn, recall, predict and generalize from historic data using a process called ‘LEARNING’. Use AI expertise to optimize existing algorithms and focus on developing frameworks and abstractions to make common tasks, such as feature generation, model training, and deployments, easier for both scientists and engineers.

2. LITERATURE REVIEW

Previous studies based on ANN characteristics and advantages of using such techniques in modern technology and real-time application benefits, innovation trends were reviewed.

E.O. Ezugwu, Arthur & Hines (1995) predicted the failure rate of the mixed–oxide ceramic cutting tool. Different values of feed rate and cutting speed have been used to train Artificial Neural Network, Multi-layer Perception using the backpropagation algorithm. The trained network has been used to predict tool lives where the failure mode for experiments was not used in training. The best result he found was an 87.5% correct failure-mode prediction.

Abbas Heiot (2002) compared the prediction performance of multilayer perception and radial basis function neural networks to that of regression analysis. The results of this study indicate that when a combined third generation and fourth generation language data set were used, the neural network produced improved performance over conventional regression analysis in terms of mean absolute percentage error.

Youness El Hamzaoui and Jose Alfredo Hernandez Perez (2011), used ANN to predict residential properties. They applied both Feed-Forward and Back Propagation algorithms for prediction. They interpret that since the prediction is good, the ANN model can be used for reliability prediction.

Rajat Gupta (2017) tried to minimise the error by updating weights and biases using backpropagation and gradient descent. Initially, his model is unstable with wrong values of weights and biases. So, he trained the model and then found a cost function to minimize the error or deviation from output to original output by updating weight biases using backpropagation and gradient descent.

Alfiyatin, Taufiq, Febrita & Mahmudy (2017) determined the selling price of a house using Regression Analysis and particle swarm optimization. The researchers used three factors that influence the price of a house, which include physical conditions, concept and location. The research aims to predict house prices based on the houses in Malang city with regression analysis and particle swarm optimization. The result of this research proved that a combination of regression and particle swarm optimization is suitable and got the minimum prediction error.

3. DATA DESCRIPTION AND METHODOLOGY

DATA DESCRIPTION:

To perform prediction using Neural Network, the dataset “Zillow’s Home Value Prediction data” was taken from Kaggle. The following information provides information about the variables that are used to perform Neural Networks.

Sample Size : 1460

Source : Secondary data

Attributes :11

Training Set : 80%

Test set : 20%

S.No	Variable	Description
1	Lot area	Area of a parcel of land
2	Overall Qual	Overall material and finish quality
3	Overall Cond	Overall condition rating
4	Total Bsmt SF	Total square feet of basement area
5	Full Bathroom	A full bathroom contains four key items a bathtub, a shower, a toilet, and a sink with running water
6	Half Bathroom	If we remove both the shower and the tub from Full Bathroom, we end up with a “half bathroom”
7	Bedroom Abv Gr	Number of bedrooms above basement level
8	Tot Rms AbvGrd	Total rooms above grade (does not include bathrooms)
9	Fireplaces	Number of fireplaces
10	Garage Area	Size of garage in square feet
11	Above Median Price	Sales Price is in dollars. This has been split into two as above median and below median. 0 as below median and 1 as above median

Table-3.1

4. METHODOLOGY

After exploring the data, Neural Network modelhas been applied for prediction. The Elements, Activation Function, Optimizer used in this research were described.

Neural Network:

Artificial neural networks have two main hyperparameters that control the architecture or topology of the network: the number of layers and the number of nodes in each hidden layer. We must specify values for these parameters when configuring our network.

Elements of a Neural Network:

- Input Layer:

This layer accepts input features. It provides information from the outside world to the network, no computation is performed at this layer, nodes here just pass on the information(features) to the hidden layer.

- Hidden Layer:

Nodes of this layer are not exposed to the outer world.They are part of the abstraction provided by any neural network. The hidden layer performs all sorts of computation on the features entered through the input layer and transfer the result to the output layer.

- Output Layer:

This layer brings up the information learned by the network to the outer world.

There are terms used to describe the shape and architecture of a neural network; for example

- **Width**: The number of nodes in a specific layer.
- **Depth**: The number of layers in a neural network.
- **Architecture**: The specific arrangement of the layers and nodes in the network.

In general, you cannot analytically calculate the number of layers or the number of nodes to use per layer in an artificial neural network to address a specific real-world predictive modelling problem. The number of layers and the number of nodes in each layer are model hyperparameters that you must specify. A deep model provides a hierarchy of layers that build up increasing levels of abstraction from the space of the input variables to the output variables.

1) Activation function:

The activation function is a mathematical “gate” in between the input feeding the current neuron and its output going to the next layer. It can be as simple as a step function that turns the neuron output on and off, depending on a rule or threshold.The activation function is mostly used to make a non-linear transformation which allows us to fit nonlinear hypotheses or to estimate the complex functions.

- Decides whether a neuron should be activated or not by calculating the weighted sum and further adding bias with it.
- Purpose to introduce non-linearity into the output of a neuron.

Many Activation Functions are available but, in this research,ReLu and Sigmoid Activation Function were used.

- **RELUActivation function:**

2) Rectified Linear Unit allows the network to converge very quickly. Although it looks like a linear function, ReLU has a derivative function and allows for back-propagation.A Recurrent Neural Network may be more useful to deal with the prediction problem in financial. It is the most widely used activation function. Chiefly implemented in hidden layers of Neural network.

- **Equation :** $A(x) = \max(0,x)$
- **Value Range :** [0, inf)
- **Nature :** Non-linear, which means we can **easily backpropagate**

the errors and have multiple layers of neurons being
activated by the ReLU function.

- **Uses :** Less computationally expensive than tanh and sigmoid

because it involves simpler mathematical operations and at a time only a **few neurons are activated** making the
network sparse making it efficient and easy for
computation

- **Sigmoid Activation Function:**

For a long time, through the early 1990s, it was the default activation used on neural networks.

- **Equation**: $\mathbf{A} = \frac{\mathbf{1}}{(\mathbf{1}+\mathbf{e}^{-\mathbf{x}})}$
- It is a function which is plotted as 'S' - shaped graph
- **Nature**:Non-linear. Small changes in X would also bring about

large changes in the value of Y.

- **Value Range** :0 to 1
- **Uses** : Usually used in the output layer of binary classification,

where the result is either 0 or 1 (predicted as 1if it is
greater than 0.5, Otherwise 0)

Flow - Feed-Forward Neural Network:

Information flows through a neural network in two ways namely Feed Forward Neural Network and Back Propagation Neural Network.Feed-Forward flow was used here. When learning (being trained) or operating normally (after being trained), patterns of information are fed into the network via the input units, which trigger the layers of hidden units, and these in turn arrive at the output units. This common design is called a feed-forward network.

Not all units "fire" all the time. Each unit receives inputs from the units to its left, and the inputs are multiplied by the weights of the connections they travel along. Every unit adds up all the inputs it receives in this way and (in the simplest type of network) if the sum is more than a certain threshold value, the unit "fires" and triggers the units it's connected to.For a neural network to learn, there has to be an element of feedback involved—just as children learn by being told what they're doing right or wrong. In fact, we all use feedback, all the time.

Optimizer - *Stochastic Gradient Descent***:**

During the training process, we tweak and change the weights of our model to try and minimize that loss function, and make our predictions as correct as possible. But how exactly do we do that? How do we change the parameters of our model, by how much, and when?**This is where optimizers come in.** Optimizers, shape and mould our model into its most accurate possible form by adjusting the weights. The loss function is the guide to the terrain, telling the optimizer when it's moving in the right or wrong direction.There are many types of optimizers that can be used such as Gradient Descent, Adam, etc., but here, Stochastic Gradient Descent was used.

Stochastic gradient descent is used in real-time on-line processing, where the parameters are updated while presenting only one training example, and so an average of accuracy levels and training costs are taken for the entire training dataset at each epoch. Instead of calculating the gradients for *all of our training examples* on every pass of gradient descent, it's sometimes more efficient to only use a subset of the training examples each time. Stochastic gradient descent is an implementation that either uses batches of examples at a time or random examples on each pass.

4. RESULTS

The results from our model,for the data, are discussed below.

Descriptive Statistics

	LotArea	OverallQual	OverallCond	TotalBsmtSF	FullBath	HalfBath	BedroomAbvGr	TotRmsAbvGrd	Fireplaces	GarageArea	AboveMedi
count	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460
mean	10516.828082	6.099315	5.575342	1057.429452	1.565068	0.382877	2.866438	6.517808	0.613014	472.980137	0
std	9981.264932	1.382997	1.112799	438.705324	0.550916	0.502885	0.815778	1.625393	0.644666	213.804841	0
min	1300.000000	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	2.000000	0.000000	0.000000	0
25%	7553.500000	5.000000	5.000000	795.750000	1.000000	0.000000	2.000000	5.000000	0.000000	334.500000	0
50%	9478.500000	6.000000	5.000000	991.500000	2.000000	0.000000	3.000000	6.000000	1.000000	480.000000	0
75%	11601.500000	7.000000	6.000000	1298.250000	2.000000	1.000000	3.000000	7.000000	1.000000	576.000000	1
max	215245.000000	10.000000	9.000000	6110.000000	3.000000	2.000000	8.000000	14.000000	3.000000	1418.000000	1

Table-4.1

The table-4.1 gives the count, mean, median, first quartile, third quartile, standard deviation, the minimum and maximum value of the data. The total number of observations used in the model is 1460.

Correlation Analysis:

To understand the relationship between the variables,Correlation Analysis is used.

Input	AboveMedianPrice
LotArea	0.189180249
OverallQual	0.672486281
OverallCond	-0.141429206
TotalBsmtSF	0.438300775
FullBath	0.615950512
HalfBath	0.308643517
BedroomAbvGr	0.144853307
TotRmsAbvGrd	0.417354918
Fireplaces	0.420297074
GarageArea	0.50506793
AboveMedianPrice	1

Table-4.2

Table-4.2 shows that Overall quality, Full Bathroom and Garage Area has a strong relationship between with the Above Median Price. This relationship is found based on the values that are greater than 0.5

Cross Tabulation:

After the Correlation Analysis, Cross tabulation was used to understand more about the variables. Cross tabulation is a method to quantitatively analyze the relationship between multiple variables. Here the Quality and ABmedian value's relationship is shown .

Quality * ABmedian

quality * ABmedian Crosstabulation

Count

		ABmedian		Total
		0	1	
Quality	1	2	0	2
	10	1	17	18
	2	3	0	3
	3	20	0	20
	4	113	3	116
	5	350	47	397
	6	198	176	374
	7	42	277	319
	8	3	165	168
	9	0	43	43
Total		732	728	1460

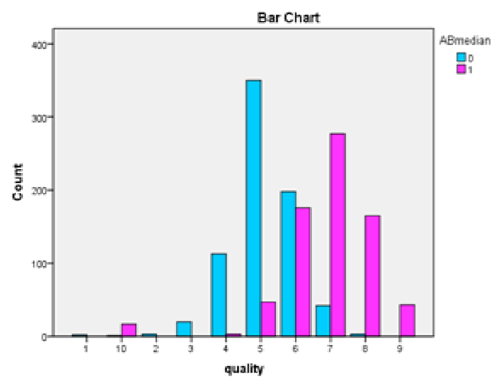


Table-4.3Figure-4.1

The above table and graph show the relationship between the quality and ABmedian value. This explains that 90.34% of the observation have good quality whichdetermines quality level 5-10. And half of the data are above the median sales price.

Fullbath * ABmedian:

fullbath * ABmedian Crosstabulation

Count

		ABmedian		Total
		0	1	
Fullbath	0	6	3	9
	1	554	96	650
	2	172	596	768
	3	0	33	33
Total		732	728	1460

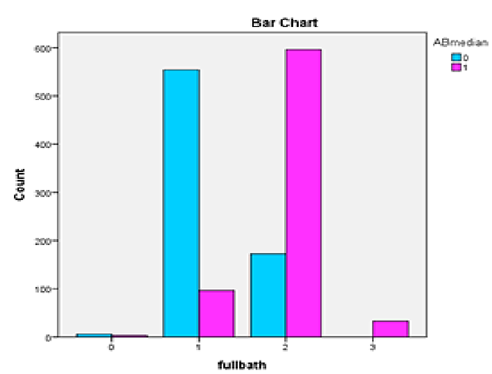


Table-4.4 Figure-4.2

The above table and graph show the relationship of the Full Bathroom and ABmedian value. This explains that a large amount of data falls in the Double full bathroom and its sales price is above the median. The Single full bathroom is the second highest that falls under the below-median sales price.

Neural Network:

To build a Neural Network, Python is used here. The data has been imported (Table-4.5) to the Jupyter Notebook using Python and the description of the data was obtained (Table-4.1)

	LotArea	OverallQual	OverallCond	TotalBsmtSF	FullBath	HalfBath	BedroomAbvGr	TotRmsAbvGrd	Fireplaces	GarageArea	AboveMedianPrice
0	8450	7	5	856	2	1	3	8	0	548	1
1	9600	6	8	1262	2	0	3	6	1	460	1
2	11250	7	5	920	2	1	3	6	1	608	1
3	9550	7	5	756	1	0	3	7	1	642	0
4	14260	8	5	1145	2	1	4	9	1	836	1
5	14115	5	5	796	1	1	1	5	0	480	0
6	10084	8	5	1686	2	0	3	7	1	636	1
7	10382	7	6	1107	2	1	3	7	2	484	1
8	6120	7	5	952	2	0	2	8	2	468	0
9	7420	5	6	991	1	0	2	5	2	205	0

Table-4.5

The data is converted into the array to process. Now, split the dataset into input features as (X) and the feature we wish to predict as (Y). To do that split, we simply assign the first 10 columns of our array to a variable called X and the last column of our array to a variable called Y.

The next step in our process is to make sure that the scale of the input features is similar. Features such as lot area are in the order of the thousands, a score for overall quality is ranged from 1 to 10, and the numbers of fireplaces tend to be 0, 1 or 2. This makes it difficult for the initialization of the neural network, which causes some practical problems. 'MinMaxScaler'was used to scale the data, so that all the input features lie between 0 and 1. Our scaled dataset is stored in the array 'X_scale'.

```
array([[0.0334198 , 0.66666667, 0.5       , ..., 0.5       , 0.        ,
        0.3864598 ],
       [0.03879502, 0.55555556, 0.875     , ..., 0.33333333, 0.33333333,
        0.32440056],
       [0.04650728, 0.66666667, 0.5       , ..., 0.33333333, 0.33333333,
        0.42877292],
       ...,
       [0.03618687, 0.66666667, 1.        , ..., 0.58333333, 0.66666667,
        0.17771509],
       [0.03934189, 0.44444444, 0.625     , ..., 0.25      , 0.        ,
        0.16925247],
       [0.04037019, 0.44444444, 0.625     , ..., 0.33333333, 0.        ,
        0.19464034]])
```

Figure-4.3

To split our dataset into a training set and a test set, 'train_test_split' was used. This function helps us to split our dataset into two. 20% of the data goes to the test set whereas 80% are for the training set, because if the training happened with more observations, then the test's accuracy will be more accurate.

To build our Neural Network, the first thing we have to do is to set up the architecture. That is, the number of hidden layers, number of neurons, activational function for the hidden layer and the output layer are to be specified. 32 neurons were used in the hidden layer with ReLU activational function. The output layer has 1 neuron with the sigmoid activation function. These instructions are described in 'Keras'.

Before we start our training, we have to configure the model by telling it, which algorithm we want to use to do the optimization, telling it what loss function to use, metrics we want to track apart from the loss function. Stochastic Gradient Descent is used here. The loss function for outputs that take the values 1 or

0 is called binary cross-entropy. Lastly, we want to track accuracy on top of the loss function. We are ready to train!

The function called ‘fit’ is used as we are fitting the parameters to the data. We have to specify what data we are training on, which is ‘X_train’ and ‘Y_train’. Then, we specify the size of our mini-batch and how long we want to train it for (epochs). The epoch is to tune our model. Here the epochs are 100.

```
Train on 1168 samples, validate on 292 samples
Epoch 1/100
1168/1168 [==============================] - 1s 1ms/step - loss: 0.7076 - acc: 0.5000 - val_loss: 0.7056 - val_acc: 0.4521
Epoch 2/100
1168/1168 [==============================] - 0s 92us/step - loss: 0.6977 - acc: 0.4777 - val_loss: 0.6951 - val_acc: 0.4384
Epoch 3/100
1168/1168 [==============================] - 0s 93us/step - loss: 0.6901 - acc: 0.5308 - val_loss: 0.6869 - val_acc: 0.6336
Epoch 4/100
1168/1168 [==============================] - 0s 96us/step - loss: 0.6838 - acc: 0.7072 - val_loss: 0.6807 - val_acc: 0.7397
Epoch 5/100
1168/1168 [==============================] - 0s 87us/step - loss: 0.6785 - acc: 0.7509 - val_loss: 0.6756 - val_acc: 0.7637
Epoch 6/100
1168/1168 [==============================] - 0s 113us/step - loss: 0.6737 - acc: 0.7748 - val_loss: 0.6709 - val_acc: 0.7945
Epoch 7/100
1168/1168 [==============================] - 0s 91us/step - loss: 0.6689 - acc: 0.7979 - val_loss: 0.6657 - val_acc: 0.8116
Epoch 8/100
1168/1168 [==============================] - 0s 111us/step - loss: 0.6638 - acc: 0.8057 - val_loss: 0.6603 - val_acc: 0.8185
Epoch 9/100
1168/1168 [==============================] - 0s 117us/step - loss: 0.6581 - acc: 0.8134 - val_loss: 0.6539 - val_acc: 0.8322
```

Figure-4.4

We can now see that the model is training! By looking at the numbers, we should be able to see the loss decrease and the accuracy increase over time.

```
Epoch 92/100
1168/1168 [==============================] - 0s 75us/step - loss: 0.2853 - acc: 0.8827 - val_loss: 0.2774 - val_acc: 0.8938
Epoch 93/100
1168/1168 [==============================] - 0s 85us/step - loss: 0.2848 - acc: 0.8827 - val_loss: 0.2820 - val_acc: 0.8904
Epoch 94/100
1168/1168 [==============================] - 0s 87us/step - loss: 0.2838 - acc: 0.8870 - val_loss: 0.2807 - val_acc: 0.8904
Epoch 95/100
1168/1168 [==============================] - 0s 75us/step - loss: 0.2835 - acc: 0.8853 - val_loss: 0.2770 - val_acc: 0.8904
Epoch 96/100
1168/1168 [==============================] - 0s 81us/step - loss: 0.2827 - acc: 0.8827 - val_loss: 0.2771 - val_acc: 0.8904
Epoch 97/100
1168/1168 [==============================] - 0s 81us/step - loss: 0.2799 - acc: 0.8878 - val_loss: 0.2702 - val_acc: 0.8973
Epoch 98/100
1168/1168 [==============================] - 0s 80us/step - loss: 0.2820 - acc: 0.8844 - val_loss: 0.2742 - val_acc: 0.8938
Epoch 99/100
1168/1168 [==============================] - ETA: 0s - loss: 0.2646 - acc: 0.889 - 0s 84us/step - loss: 0.2805 - acc: 0.8836
- val_loss: 0.2726 - val_acc: 0.8938
Epoch 100/100
1168/1168 [==============================] - 0s 86us/step - loss: 0.2798 - acc: 0.8861 - val_loss: 0.2751 - val_acc: 0.8904
```

Figure-4.5

Initially, the accuracy was 45.21% but after training the model at 100th Epoch, the accuracy was increased to 89.04%. This means that we can trust our model for prediction. This gives the prediction value more accurately because it has a greater accuracy value. Using matplotlib we can visualize the training and test accuracy. The improvements in our model to the training set looks matched up with improvements to the test set.

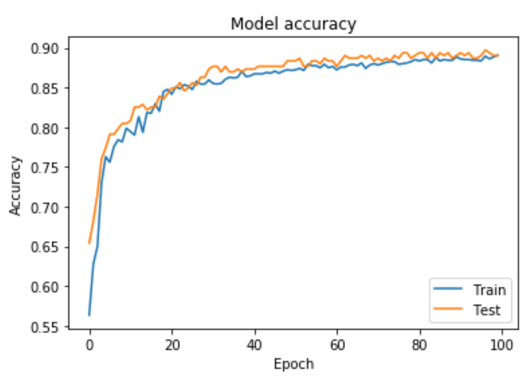


Figure-4.6

5. CONCLUSION

Artificial Intelligence provides endless opportunities for businesses to grow and improve their business operations. Through intelligent automation and using machine level techniques, great changes can be brought into daily life activities. The Neural Network is a Machine Learning Algorithm that performs better for our data. Here, the prediction accuracy is 89.04% at 100th Epoch, which is a good accuracy. Utilising intelligent automation will boost the growth process of any business and makes customers to make the best decision. The Neural Network's basics are merely like a regression model (in our case), which allocates weights and bias to fit the model for prediction which involves high quality and accuracy in Outputs that has less error and faster report generation. Businesses rely on everyday growth, its progress is based on advanced technology, which is their valuable resource for survival. With growing competition, the neural network is one of the paving ways for businesses to grow faster and work better.

REFERENCES

1. *Adyan Nur Alfiyatin, Hilman Taufiq, Ruth Ema Febrita & Wayan Firdaus Mahmudy (2017). Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization. (IJACSA) International Journal of Advanced Computer Science and Applications, Volume:8, No:10.*
2. *Abbas Heiat (2002). Comparison of artificial neural network and regression models for estimating software development effort.Elsevier, Volume 44, Issue 15, Pages 911-922.*
3. *Rajat Gupta (2017). Getting started with Neural Network for regression and Tensorflow. Medium*
4. *E.O. Ezugwu, S.J. Arthur & E.L. Hines (1995). Tool-wear prediction using artificial neural networks. Elsevier, Volume 49, Issues 3–4, Pages 255-264*
5. *Youness El Hamzaoui, Jose Alfredo Hernandez Perez (2011). Application of Artificial Neural Networks to Predict the Selling Price in the Real Estate Valuation Process. IEEE Computer Society.*
6. *Haykin S., (1999). Neural Networks a Comprehensive Foundation, 2nd edition, Pearson Education.*
7. *Limsombunchai, V., Gan, C. and Lee, M. (2004). House Price Prediction: Hedonic Price Model vs. Artificial Neural Network, American Journal of Applied Sciences 1 (3): 193–201. https://thescipub.com/pdf/10.3844/ajassp.2004.193.201*
8. *Núñez Tabales, Julia M.; Caridad y Ocerin, José María; Rey Carmona, Francisco J. (2013). Artificial neural networks for predicting real estate prices. Econstor. https://www.econstor.eu/bitstream/10419/113851/1/756619068.pdf*
9. *Chusheng Wang (2021). House Price Prediction Model Based on Neural Network. Research Square. https://assets.researchsquare.com/files/rs-805003/v1_covered.pdf?c=1632752256*

DOI No. : 10.5958/2394-9309.2022.00048.8
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

UNIQUE ISOLATED SIGNED TOTAL DOMINATING FUNCTION IN GRAPHS

Sunitha S[*] and S. Chandra Kumar[**]
[*]Research Scholar, Department of Mathematics, Scott Christian College, Nagercoil, Tamilnadu
[**]Associate Professor, Department of Mathematics, Scott Christian College, Nagercoil, Tamilnadu

ABSTRACT
An unique isolated signed total dominating function (UISTDF) of a graph G is a signed domination function $f : V(G) \to \{-1,+1\}$ such that there exists exactly one vertex $w \in V(G)$ with $f(N(w)) = +1$. The unique isolated signed total domination number of G is denoted by $\gamma_{ist}^{u}(G)$, is the minimum weight of an UISTDF OF G. In this paper, we study some properties of UISTDF and we give unique isolated signed total domination number of disconnected graphs and some special graphs.
Keywords: isolated domination, signed dominating function, isolated signed total dominating function, unique isolated signed total dominating function.

1. INTRODUCTION

In this paper, we consider only finite, simple and undirected graphs. The set of vertices and edges of a graph $G(p,q)$ will be denoted by $V(G)$ and $E(G)$ respectively, $p = |V(G)|$ and $q = |E(G)|$. For graph theoretic terminology, we follow [5].

For $v \in V(G)$, the open neighborhood v is $N_G(v) = \{u \in V(G): uv \in E(G)\}$ and the closed neighbourhood of v is $N_G[v] = \{v\} \cup N(v)$. The degree of v is $deg_G(v) = |N_G(v)|$. The minimum and maximum degree of G is defined by $\delta(G) = \min_{v\in V(G)}\{\deg(v)\}$ and $\Delta(G) = \max_{v\in V(G)}\{\deg(v)\}$ respectively.

A function $f : V(G) \to \{0,1\}$ is called a dominating function if for every vertex $v \in V(G), f(N[v]) \geq 1$ [6]. The weight of f, denoted by $w(f)$ is the sum of the values $f(v)$ for all
$v \in V(G)$.

Many kinds of domination parameter has been defined and studied by many authors, we founded in [6]. Various domination functions has been defined from the definition of dominating function by replacing the co-domain $\{0,1\}$ as one of the sets $\{-1,0,1\}, \{-1,+1\}$ and etc. one of such example is signed dominating function [3,4].

In 1995, J.E.Dunbar et al [4] defined signed dominating function. A function $f : V(G) \to \{-1,+1\}$ is a signed dominating function of G, if for every vertex $v \in V(G)$, $f(N[v]) \geq 1$. The signed domination number, denoted by $\gamma_s(G)$, is the minimum weight of a signed dominating function on G [4].

In 2001, BohdanZelinka and Liberec [1] introduced the con- cept of signed total dominating function. A function $f : V(G) \to \{-1,+1\}$ is a signed total dominating function (STDF) of G, if for every vertex

$v \in V(G), f(N(v)) \geq 1$. The signed total domina- tion number, denoted by γst (G), is the minimum weight of a signed dominating function on G [1]. The signed total dominating function has been studied by several authors including [2, 7, 8, 10, 13, 14].

A subset S of vertices of a graph G is a total dominating set of G if every vertex in $V(G)$ has a neighbour in S. The minimum cardinality of a total dominating set of G is said to be the total domination number and is denoted by $\gamma t(G)$.. A subset S of vertices of a graph$G$ is a 2-total dominating set of G if every vertex in $V(G)$ has at least two neighbours in S. The minimum cardinality of a 2-total dominating set of G is said to be the total domination number and is denoted by$\gamma_{2,t}(G)$.

In 2016, Hameed and Balamurugan [11] introduced the concept of isolate domination in graphs.A dominating set S of a graph G is said to be an isolate dominating set if $< S >$ has at least one iso-lated vertex [11]. An isolate dominating set S is said to be minimal if no proper subset of S is an isolate dominating set. The minimum and maximum cardinality of a minimal isolate dominating set of G are called the isolate domination number $\gamma 0(G)$ and the upper isolate domination number $\Gamma 0(G)$ respectively.

In 2019, SivagnanamMutharasu and V. Nirmala [12] defined the concept of unique isolate domination in graphs. A dominating set S of a graph G is said to be an unique isolate dominating set (UIDS) of G if $< S >$ has exactly one isolated vertex. An UIDS S is said to be minimal if no proper subset of S is an UIDS The minimum cardinality of a minimal UIDS of G is called the UID number, denoted by $\gamma_0^U(G)$.

By using the definition of signed to total dominating function and unique isolate domination, we introduced the concept of unique unique isolated signed total dominating function. An unique isolated signed total dominating function (UISTDF) of a graph G is a signed domination function $f : V(G) \rightarrow \{-1, +1\}$ such that there exists exactly one vertex $w \in V(G)$ with $f(N(w)) = +1$. The unique isolated signed total domination number of G is denoted by $\gamma_{ist}^u(G)$, is the minimum weight of an UISTDF of G. In this paper, we study some properties of UISTDF and we give unique isolated signed total domination number of disconnected graphs and some special graphs.

MAIN RESULTS

Lemma 1

Let G be any graph in which $\deg(v)$ is even for all $v \in V(G)$. Then, G does not admit UISTDF

1) Proof

Note that $|N(u)|$ is even for any vertex $u \in V(G)$. Thus there exist no vertex $u \in V(G)$ such that$f(N(v)) = 1$ for any function $f : V(G) \rightarrow$
$\{-1, +1\}$.

Lemma 2

For any graph G which admits UISTF, $\gamma_{st}(G) \leq \gamma_{ist}(G) \leq \gamma_{ist}^u(G)$.

Proof.

Since every ISTDF is a STDF and every UISDF is an ISTDF, we have $\gamma_{st}(G) \leq \gamma_{ist}(G) \leq \gamma_{ist}^u(G)$

Theorem 3

Let $n \geq 2$ be an integer and let G be a disconnected graph with n components $G_1, G_2, \ldots, G_n$such that the first $r(\geq 1)$ components $G_1, G_2, \ldots, G_r$ admit UISTDF. Then, $\gamma_{ist}^u(G) = \min_{1\leq i\leq r}\{t_i\}$, where $t_i = \gamma_{ist}^u(G_i) + \sum_{j=1, j\neq i}^{n} \gamma_{2st}(G_j)$.

Proof

With out loss of generality, we assume that $t_1 = \min_{1\leq i\leq r}\{t_i\}$. Let f_1 be an minimum UISTDF of G_1 and f_i be a minimum S2TDF of G_i for each $i$with $2 \leq i \leq n$. Then, $f: V(G) \rightarrow \{-1, +1\}$ defined by $f(x) = f_i(x)$, if $x \in V(G_i)$, is an UISTDF of G with weight $\gamma_{ist}^u(G_1) + \sum_{i=2}^{n} \gamma_{2st}(G_j)$ and so$\gamma_{ist}^u(G) \leq \gamma_{ist}^u(G_1) + \sum_{i=2}^{n} \gamma_{2s}(G_i) = t_1$. Let g be a minimum.
UISTDF of G. Then there exists an integer j such that $g|G_j$ is a minimum UISTDF of $G_j$for some$j$ with $1 \leq j \leq r$. Also for each $i$with $1 \leq j \leq n\ (i \neq j), g|G_i$is a minimum S2TDF of G_i.Therefore $w(g) \geq \gamma_{ist}^u(G_j) + \sum_{i=1, i\neq j}^{n} \gamma_{2st}(G_i) = t_j \geq t_1$ and
hence$\gamma_{ist}^u(G) = \min_{1\leq i\leq r}\{t_i\}$

Remark4.

(a) Let G be a grap[h which admits a 2- total dominating set S. Then, $N(v) \subseteq S$ whenever $|N(v)| = 2$ for any vertex
$v \in V(G)$.

(b) Let G be a graph which admits an UISTDF (or STDF), say f. Then the vertices of $N(v)$ are labelled with $+1$ sign whenever $|N(v)| \leq 2$ for any vertex
$v \in V(G)$.

Lemma 5

a) The cycle graph C_n of order $n \geq 3$ does not admit UISTDF.

b) The path P_n graph of order $n \geq 2$ does not admit UISTDF.

Proof.

a) Let f be an UISTDF of C_n. Note that $N|(v)| = 2$ for all $v \in V(C_n)$. By Remark 4(b) all the vertices of C_n must have $+1$ sign. In this case $f(N(v)) \neq 1$ for all $v \in V(C_n)$, a ontradiction.

b) Let g be an UISTDF of P_n. Note that $N|(v_i)| = 2$ for all $i \in \{2,3, \ldots, n-1\}$. By Remark 4(b) all the vertices of P_n must have $+1$ sign. Here, $g(N(v_1)) = g(N(v_n)) = 1$, a contradiction.

Definition 6

The graph $P_n^{(2)+}$ of order $n \geq 3$ is obtained from P_n by joining the interval vertices v to the one end v_1 when $d(v, v_1)$ is even.

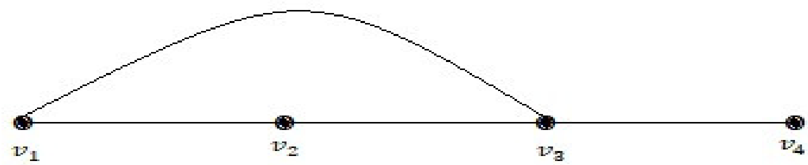


Figure 1 : $P_4^{(2)+}$

Let f be an UISTDF of the graph $P_4^{(2)+}$. Since $deg(v_2) = 2$, by Remark 4, we have $f(v_1) + f(v_3) = +1$. Suppose $f(v_4) = -1$. Then $f(N(v_4)) = 1$and $f(N(v_3)) \leq 1$, a contraction. Suppose $f(v_2) = -1$. Then, $f(N(v_1)) \leq 0$, a contradiction. Therefore all the vertices of $P_4^{(2)+}$ must have $+1$ sign and so $\gamma_{ist}^u(P_n^{(2)+}) = 4$.

B.

Proposition 7

Let H be any graph which does not admit UISTDF. Then, $G = H \cup rP_n^{(2)+} (r \geq 1)$ admits UISTDF if and only if $r = 1$. In this case $= 4 + \gamma_{2st}(H)$.

Proof

Suppose there exists an UISTDF of$G$, say$f$. Let $V(G) = V(H) \cup V(rP_4^{(2)+})$, where $V(H) = \{u_1, u_2, \dots, u_n\}$ and $V(rP_4^{(2)+})$ =$\{v_i^j \ : \ i = 1,2,3,4\}andj = 1,2, \dots, r$. Suppose $r \geq 2$.Since $\deg(v_2^1) = \deg(v_2^1) = 2$, by Remark 4, $f(N(v_4^1)) = f(N(v_4^2)) = +1$, which is a contraction to f is UISTDF. Thus $r = 1$. Conversely,

suppose$r = 1$. Then$G = H \cup P_4^{(2)+}$. By taking $G_1 = P_4^{(2)+} G_2 = H$ in Theorem 3, we can prove G admits UISTF and $\gamma_{ist}^u(G) = \gamma_{ist}^u(G_1) + \gamma_{ist}^u(G_2) = 4 + \gamma_{2st}(G_2)$.

Lemma 8

Let f be an UISTDF of G and let $S \subset V$. Then $f(S) = |S| (mod\ 2)$.

Proof.

S^- $\{v \mid f(v) = -1, v \in S$ Let $S^+ = \{v \mid f(v) = 1, v \in S\}$ and $=\}$ .Then, $|S^+| + |S^-| =$ and If both S^-
$|S|$ $|S^+| - |S^-| = f(S)$ and S^+ are either odd or even, then both $|S|$ and $f(S)$ must be even.

If either one of S^- and S^+ is odd and another one is even, then both $|S|$ and $f(S)$ must be odd. Therefore, $f(S) = |S| - 2|S^-|$.

Theorem 9

Let G be a connected graph of order $n$and $\delta \geq 2$. Then, $2\gamma_{2,t}(G) - n \leq \gamma_{ist}^u(G)$.

1) ***Proof***

Let f be a minimum UISTDF of G. Let $V^+ = \{u \in V : f(u) = +1\}$ and$V^- = \{u \in V : f(v) = -1\}$. If $V^- = \varphi$, then G does not admit UISTDF. Thus $V^- = \varphi$. Suppose there exists a vertex $v \in V^-$. Since $g(N(v)) \geq 1$and $\delta \geq 2$, v has at least two adjacent vertices in V^+. Similarly, when $v \in V^+$also v has at least two adjacent vertices in V^+.

γ_{ist}^{u}
n

Therefore $V^+$is a 2-total dominating set for G and so $|V^+| \geq \gamma_{2,t}(G)$. Since $\gamma_{ist}^{u}(G) = |V^+| - |V^-|$and $n = |V^+| + |V^-|$, we have $(G) = 2|V^+| - n$ans so $\gamma_{ist}^{u}(G)| \geq 2\gamma_{2,t}(G) -$n.

Remark 10

The inequality given in Lemma 9 is sharp. For example, consider the following graph G. Every 2-total dominating set contain the vertices v_1, v_3, v_5, v_6 and v_7 (by Remark 4(a)). Since $|N(v_3)| = 3$, either v_2 or v_4 must be in every 2-total dominating set. Thus $\gamma_{2st}(G) \geq 6$. Also $V(G) - \{v_2\}$ is a 2- total dominating set with 6 vertices and so $\gamma_{2st} \leq 6$. For every UISTDF of G, it is true that $f(v_1) = f(v_3) = f(v_5) = f(v_6) = f(v_7) = +1$ (by Remark 4(b)). Since $|N(V_3)| = 3$, either $f(v_2)$ or $f(v_4)$ must have the label +1. Thus $w(f) \geq 5$ and so $\gamma_{ist}^{u}(G) \geq 5$. Now label the vertex v_2 of G by -1 and other vertices by +1. This gives a UISTDF with weight 5 and so $\gamma_{ist}^{u}(G) \leq 5$. Here, we have $2\gamma_{2,t}(G) - n = 2(6) - 7 = 5 = \gamma_{ist}^{u}(G)$.

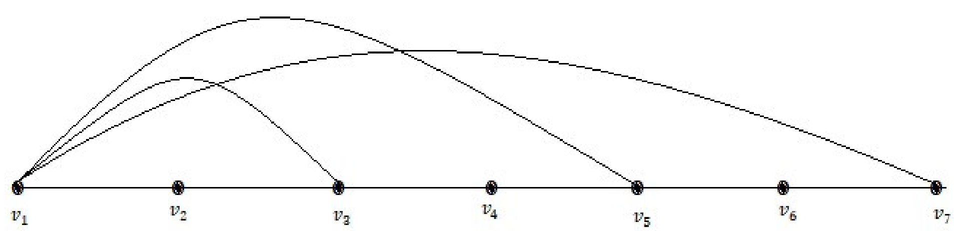


Figure 2: $P_7^{(2)+}$

Lemma 11

Let $n \geq 7$ be an odd integer. Then the graph $G = P_n^{(2)+}$ admits UISTDF with$\gamma_{ist}^{u}(G) = n - 2$.

Proof

Let $\{v_1, v_2, \dots, v_n\}$ be the vertex set of G. Let f be an UISTDF. Note that $|N(v_i)| = 2$for $i = 2, 4, \dots, n-1$. By remark 4(b) the vertices $v_1, v_3, \dots, v_n$ must have +1 sign.

Suppose $f(v_i) = f(v_j) = -1$ for some $i < j$and$i, j \in \{2, 4, \dots, n-1\}$. In this case, both $f(N(v_{i+1}))$ and $f(N(v_{j+1}))$ are less than are equal to 1, a contradiction.

Therefore, $f(v) = -1$ for maximum of one vertex in G. Thus $w(f) \geq n - 2$ and so $\gamma_{ist}^{u}(G) = n - 2$

Define a function $g : V(G) \to \{-1, +1\}$ as follows:

$$g(v_i) = \begin{cases} -1 & when\ i = 2 \\ +1 & otherwisw \end{cases}$$

In this case, $g(N(v_3)) = 1$ and $g(N(v_3)) \geq 2$for all$i \neq 2$. Thus g is an UISTDF with $w(g) = n -$2 and hence $\gamma_{ist}^{u}(G) = n - 2$.

Lemma 12

Let $n \geq 4$ be an even integer. Then the graph =$G = P_n^{(2)+}$ admits UISTDF with $\gamma_{ist}^{u}(G) = n$.

2) ***Proof***

Let $\{v_1, v_2, \dots, v_n\}$ be the vertex set of G. Let f be an UISTDF. Note that $|N(v_i)| = 2$ for $i = 2, 4, \dots, n-2$.By remark 4(b) the vertices $v_1, v_3, \dots, v_{n-1}$ must have +1 sign. Thus $f(N(v_n)) = 1$.

Suppose $f(v_i) - 1$ for some and $i \in \{2, 4, \ldots, n-2\}$, then $f(N(v_{i+1})) \leq 1$, a contradiction. Therefore, all the vertices of G must have +1. Thus $w(f) = n$ and so $\gamma_{ist}^{u}(G) = n$.

Remark 13

Let f be an UISTDF of the graph ${P_5}^{(2)+}$. Since $deg(v_2) = deg(v_4) = deg(v_5) =2$, by Remark 4(b), we have $f(v_1) = f(v_3) = f(v_5) = +1$. Suppose $f(v_2) = -1$. Then, $f(N(v_3)) = f(N(v_1)) = 1$, a contradiction. Therefore the graph ${P_5}^{(2)+}$ does not adsmit UISTDF.

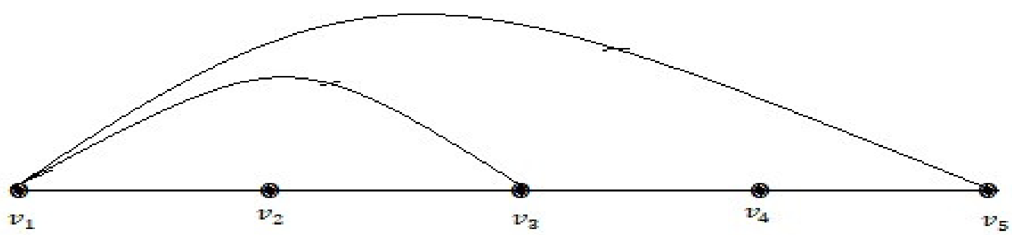


Figure 2: ${P_5}^{(2)+}$

When $n = 3$, the graph ${P_n}^{(2)+}$ is a cycle of length 3, which does not admit UISTDF(By Lemma 5). From Lemma 11, Lemma 12 and Remark 13, we can have the following theorem.

Theorem 14

Let $n \geq 2$ be an integer. Then the graph $G = {P_n}^{(2)+}$ admits UISTDF with $\gamma_{ist}^{u}(G) = n-2$ when n is odd and $\gamma_{ist}^{u}(G) = n$ when n is even.

Definition 15

The $E(H) \cup \{(x_1, y_1) : x_1 \in V(G)$ and $y_1 \in$ sum $G + H$ of two graphs G and H, is the graph with vertex set $E(G) \cup$ $V(G + H) = V(G) \cup V(H)$ and edge set $E(G+H) = V(H)\}$.

Remark 16

a) Since $P_2 + K_1 = C_3$, it does not admitUISTDF.

b) Suppose there exists an UISTDF f of the graph $P_3 + K_1$. Since $\deg(v_1) = 2$, by Remark 4(b), we have $f(v_0) = f(v_2) = +1$.

Suppose $f(v_1) = f(v_3) = +1$. Then, $f(N(v)) \neq 1$, for all $v \in V(P_3 + K_1)$, a contradiction Suppose v_1 and v_3 are of opposite sign. Then $f(N(v_0)) = f(N(v_1)) = 1$, a contradiction. Therefore the graph $P_3 + K_1$ does not admit UISTDF.

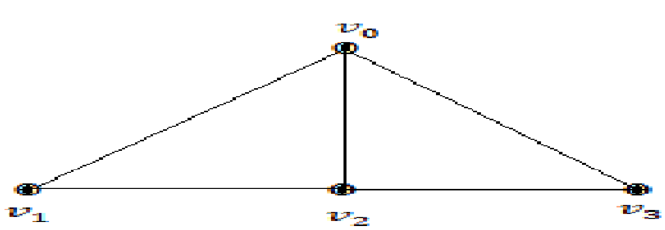


Figure 2: $P_3 + K_1$

Lemma 17

Let $n \geq 4$ be an odd integer. Then the graph $G = P_3 + K_1$ admits ISDF with $\gamma_{ist}^{u}(G) = n - 2$.

Proof

Let$\{v_0, v_1, v_2, \dots, v_n\}$ be the vertex set of$G$ such that v_0 is the full vertex. Let f be an UISTDF of G. Since$|N(v_i)| = 2$for $i = 1, n$, by remark 4(b), the vertices v_0, v_2 and v_{n-1} must have $+1$ sign.

Suppose $f(v_i) = -1$ for some $3 \leq i \leq n - 2$. Then both $f(N(v_{i-1}))$ and $f(N(v_{i+1}))$ are less than or equal to 1, a contradiction. Suppose $f(v_1) = f(v_n) = -1$. Then both $f(N(v_2))$ and $f(N(v_{n-1}))$ are less than or equal to 1, a contradiction.

Thus, $w(f) \geq n - 2$ and so $\gamma_{ist}^{u}(G) \geq n - 2$

Define a function $g : V(G) \rightarrow \{-1, +1\}$ as follows:

$$g(v_i) = \begin{cases} -1 & when\ i = 1 \\ +1 & otherwisw \end{cases}$$

In this case, $g(N(v_2)) = 1$ and $g(N(v_i)) \geq 2$ for all $i \neq 2$. Thus g is an UISTDF with $w(g) = n - 2$ and hence $\gamma_{ist}^{u}(G) = n - 2$.From lemma 17 and Remark 16, we can have the following theorem.

REFERENCES

1. *BohdanZelinka and Liberec, Signed total domination number of graphs, Czechoslovak Mathematical Journal, 51(126)(2001), 225 229.*
2. *By Maryam Atapour(Tabriz), Seyed Mahmoud Sheik- holeslami(Tabriz) and Abdollah Khodkar (Carrollton) Global signed total domination in graphs, Publ. Math. Debrecen, 79/1-2 (2011), 7-22.*
3. *J.E.Dunbar, S.T.Hedetniemi, M. A. Henning, and A. A. McRae. Minus domination in regular graphs, Discrete Math., 149 (1996) 311-312*
4. *J.E. Dunbar, S.T. Hedetniemi, M. A. Henning and P. J. Slater, Signed domination in graphs. In: Graph Theory, Combina- torics and Applications. Proc. 7th Internat. conf. Combina- torics, Graph Theory, Applications, (Y.Alavi, A.J.Schwenk, eds.). John Willey & Sons, Inc., 1 (1995) 311-322*
5. *T.W.Haynes, S.T.Hedetniemi and P.J.Slater"Fundamental of domination in graphs". Marcel Dekkerinc. New York-Basel- Hong Kong, 1998.*
6. *H. Karami, S.M.Sheikholeslami and AbdollahKhodkar, Lower bounds on signed edge total domination numbers in graphs, Czechoslovak Mathematical Journal, 58 (133) (2008), 595603*
7. *Laura Harris, The algorithmic complexity of certain functional variations of total domination in graphs, AUSTRALASIAN JOURNAL OF COMBINATIONS, 29 (2004), 143156*
8. *G. Marimuthu and M. Balakrishnan, Super edge magic graceful graphs, Inform. Sci., 287 (2014), 140151.*
9. *Michael A. Henning, Signed total domination in graphs, Discrete Mathematics, 278 (2004) 109 125*
10. *I.Sahul Hamid, S.Balamurugan, Isolate domination in graphs, Arab J Math Sci., 22 (2016), 232-241.*

11. *Sivagnanam Mutharasu1 and V. Nirmala, Unique isolate domination in graphs, Malaya Journal of Matematik, 7(4) (2019), 720-723*
12. *Wen-Sheng Li, Hua-Ming Xing, and Moo Young Sohn, On the signed total domination number of generalized petersen graphs P (n, 2) , Bull. Korean Math. Soc. 50 (2013), 20212096*
13. *Zhongsheng Huang, ZhifangFeng, and Huaming Xing, In- verse signed total domination numbers of some kinds of graphs, ICICA 2012, Part II, CCIS 308 (2012) 315321.*

DOI No. : 10.5958/2394-9309.2022.00049.X
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022
ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com
ISSN (Online) : 2394-9309

BENEFITS AND CHALLENGES FACED BY THE EMPLOYEES DUE TO WORK FROM HOMEA STUDY OF THE IT PROFESSIONALS IN CHENNAI DISTRICT, TAMIL NADU

Ida Sharon[*] and P. Aruna[**]

[*]Senior Post Graduate Student, Department of Statistics, Madras Christian College, Chennai

[**]Assistant Professor, Department of Statistics, Madras Christian College, Chennai.

ABSTRACT

Work from home (WFH) is not a new concept. From very beginning, people were trained in their respective occupations at their homes. But the rise of factories and offices, invented the term Workplace. During this pandemic, the idea of working from home received attention. We know that the information technology (IT) industry has become one of the fastest growing industries in India, because of which it has caught world attention. As a result, our study focused on IT employee's responses about working from home, particularly focusing Chennai region. We collected responses from people working in IT sector using Two stage sampling. Employees have a high autonomy in scheduling their work and therefore are assumed to have a higher intrinsic motivation. Thus, we expect working from home to positively influence work effort of employees. But on the other hand, Employees also feel isolation from workplace and miss the physical interaction with their co-workers. This in turn creates a void in their career growth. Quality time for family is compromised most times due to work. But family members are quite supportive in this.

Keywords : Work from Home, Chennai, IT employees, work-life balance.

1. INTRODUCTION

'Work from Home' is a strategy which allows workers to work securely and efficiently at home or anywhere else without being disturbed. Employees that work from home are able to complete their tasks. It also allows employees to work flexible hours, and it makes the job of the employer much easier. Working from home has been around for as long as people have been able to work. It's been a phenomenon for hundreds of thousands of years, not only in the last few decades with the arrival of telecommuting. Over the last 50 years, a range of factors have combined to result in more individuals working from home now than ever. By the year 2000, the necessity for remote work rules for both businesses and employees were acknowledged. Remote employment has risen in popularity as a result of technological advancements and more visibility on social media platforms. With the technical infrastructure created for telecommuting improving year after year, these technologies have become accountable for enabling "work from anywhere" rather than just "work from home." "Remote working does have a long and varied history, with its development affected by a variety of variables, including technological and societal shifts." In the current situation, the Corona virus has led to a significant, rapid trend towards people working from home. The rapid advent of COVID-19 prompted the World Health Organization to declare a pandemic to tackle the virus on March 11, 2020. This epidemic caused widespread social and economic disturbance throughout the world, resulting in the greatest global recession since the Great Depression of the 1930s. Staying out of the workplace was more crucial than promoting and allowing regular communication with employees because of COVID - 19. The only flexible approach an organization may use is the 'work from home' technique, which is why, particularly in the government sector, the actual location of a workplace has progressively reduced. The Indian government, led by Prime Minister Narendra Modi, announced a countrywide lockdown on March 24, 2020, to prevent the virus from spreading. A set of regulations and enforcement were followed throughout the country to safeguard the safety of the people. Since the beginning of the lockdown, about 90% of the Indian workforce has worked from home.

2. LITERATURE REVIEW

A study conducted by BalazsAczelI, Marton Kovacs, Tanja van der Lippe, Barnabas Szaszi1 *(*March 25, 2021*)*investigates the systematic examination and support of researchers' productivity and work-life balance in an easier way to those working from home. According to the findings of this survey, the vast majority of people who would like to work more from home in the future believe that it is possible to do so (86 per cent). To summarise, greater WFH would help both researchers' work and non-work lives, and neither their employment responsibilities nor their household situations would be barriers to such a shift. Thus, concluding that WFH is becoming an increasingly important part of every worker’s lives.

According to Khristian S. Liwanag, one of the most difficult challenges for businesses, especially in pandemic situations like Covid-19, is maintaining employee overall productivity levels. The goal of this survey was to see if employees and companies are satisfied with the present WFH. The study included 340 teaching and non-teaching staff from the Bulacan DepEd Schools Section and found that 111 or 32.6 percent of employees felt that the most difficult obstacle of working from home is separating between work and home time, followed by 112 or 32.9 percent who said that feeling isolated from the workplace social network is also one of the challenges. 111 or 32.6 percent agree on Inadequate equipment or lack of technical support, 125 or 36.8 percent agree on Implications of limited interaction w/manager for career, and 107 or 31.5 percent are neutral on Feelings of hostility/resentment from co-workers, indicating that implications of limited interaction with the manager for career is the leading challenge of working from home among respondents.

The study for Dr.Sridevi.R and Sanjana.N (2021) aims to understand the impact of work from home among IT sector employees in the Coimbatore District, and it is based on the responses of 111 respondents from the Information Technology sector who worked from home during COVID 19 and indicated that most employees stop working from home owing to a lack of experience, a lack of equipment, and a communication gap. They concluded that if a solution was found to bridge the communication gap, employees' desire to work from home would grow, as would their productivity. As a result, the concept of working from home may grow more common in the future.

As said by K Garg and Jan van der Rijst (2015), the purpose was to determine employees' willingness to work from home as well as to assess the impact of working from home. The study population included 1612 employees, with 48 completed surveys returned by participants and discovered that the majority of the employees save enough money each month and work. Thus, it was established that job independence, interpersonal trust, and stress had a weakly favourable relationship with virtual work experience.

ItishreeMohanty *(*03 Mar 2021)claims that on the basis of demographic characteristics, the study assessed the numerous initiatives implemented by employers during the COVID-19 pandemic crisis, as well as employee skills level in implementing this change in various service and manufacturing sectors. Using the T test, it was also determined that there was a significant difference in the employees' attitudes on their quality of life when working from home during COVID-19. The organization needs to implement climate solutions and put in place in order to make today healthier while also protecting the future.

As stated by LinaVyas&NantapongButakhieo (2021) The study's foundation is based on an examination of WFH, teleworking, telecommuting, e–working, flexible workplace, and remote work. The methodology was created to facilitate the examination of WFH in Hong Kong during the COVID-19 pandemic. A survey reveals dissatisfaction with internal infrastructure, such as no or limited access to office records. Another study claims that the specific working conditions in Hong Kong make WFH less appealing to workers. The decision to cease in-person meetings and work was made quickly, but with no instructions on how to do so. If this technique is to become a viable choice or the new normal, proper training is essential. Perhaps the working balance will be seen after the pandemic, when WFH is no longer a forced mandate, but rather a flexible option.

Ms. G.Ragavi(october 2021) aim is to determine the work-life balance of WFH employees toward the company during the COVID19 Pandemic. The study's sample size was 120, and the sample design used was convenience sampling and revealed that a major obstacle in working from home was keeping children engaged in activities, because they, too, were trapped at home owing to the epidemic and were unable to attend school. Some tasks, such as cleaning, cooking, and other housework, have also had an impact on employees' effective performance, as well as their productivity while working from home. As a conclusion, the organisation should take certain steps to lessen the challenges that employees confront when working from home.

Teresa Galanti, Gloria Guidetti and Ferdinando Toscano examine the effect of family-work conflict, social isolation, distracting environment, job autonomy, and self-leadership on employee productivity, work engagement, and stress during the pandemic. The average age of the participants was 49.81 years old, according to this cross-sectional study, which assessed data acquired through an online questionnaire completed by 209 WFH workers during the pandemic (standard deviation 9.4, minimum 25, maximum 65). This study adds to the body of knowledge about remote work and remote worker well-being, which has been highlighted in the context of the COVID-19 epidemic, which has had significant emotional and health consequences. Furthermore, the results of the study are significant because it gives information on the demands of employees who have been forced to adjust to full-time WFH as a result of the epidemic, the majority of whom had no prior WFH experience. Managers, officers, and employees involved in distant activities should think about family-work conflict, social isolation, and distracting work settings as possible barriers to WFH involvement, and job autonomy and self-leadership as potential facilitators.

The most extensive examination of WFH productivity changes for knowledge workers is presented by Michael Gibbs, FriederikeMengel, and ChristophSiemroth (July 2021). WFO productivity changes with employee characteristics, presence of children at home, and WFO travel time, and give evidence on how WFO productivity varies with employee characteristics, presence of children at home, and WFO commute time. WFH may be more challenging for individuals with less experience, shorter tenure, and positions that need a lot of teamwork and coordination. Firms will need to design tools, training, and policies to emphasise human contacts during WFO, increase the efficiency of virtual communication, and teach managers and workers to better arrange work time at home.

3. OBJECTIVES OF THE STUDY

1. To assess the satisfaction and willingness of employees to work from home arrangement.
2. To identify the benefitsof WFH to employees.
3. To identify the challenges of WFH faced by the employees.
4. To analyse work-family-balance among the employees.

4. METHODS

4.1 Data collection

A study on the Benefits and challenges faced by the employees due to COVID-19 was conducted using a goggle form. As it was the time of pandemic, we couldn’t perform face to face personal interviews. Also, printed questionnaires are not possible to be distributed as entry to IT Professional campus is strictly for the IT Professionals only. So, google forms and phone calls are put to use.

4.2. Determination of sample size

Probability sampling and non-probability sampling are the two main types of sampling methodologies. In the first situation, each member has a fixed, known chance of being included in the sample, but in the second case, no person has a particular chance of being included in the sample. We performed simple random sampling in our project and used Fisher's method to compute sample size, which is

$N/1+Ne^2$; where n = sample size, e = error (0.05)

A survey was conducted with target population as IT employees of Chennai city. The information from total of 407 respondents are collected. Two stage sampling method, which is one of the probability sampling methods is adopted in this study. In the first stage the clusters are selected and in the second stage the simple random sampling method is used for selecting the sampling unit.

Two-stage cluster sampling, a simple case of multistage sampling, is obtained by selecting cluster samples in the first stage and then selecting sample of elements from every sampled cluster A survey was done with the target population of Chennai city IT personnel. Information was gathered from a total of 407 respondents. This study uses the two-stage sampling method, which is one of the probability sampling methods. Clusters are chosen in the first step, and the sample unit is chosen in the second stage using a simple random sampling method.

A simple case of multistage sampling is two-stage cluster sampling, which is obtained by selecting cluster samples in the first stage and then selecting a sample of elements from each sampled cluster in the second stage.

4.3 Data pre-processing

All the data pre-processing and analysis were conducted using SPSS. For the analysis of the survey responses, we read all the free-text comments to ascertain that they do not contain personal information and they are in line with the respondent's answer. A pilot study was conducted with 50 respondents. When tested for normality, our data was not normally distributed. Therefore, we opted for non-Parametric tests.

4.3.1 Demographic profile of respondents

Of all the respondents, 55% were male and 45% were female. The age group 21-30 years had the highest percentage of respondents with a total of 349 (85.7%), followed by the age group of 31-40 years with a total of41 respondents (10.1%), the age group of 41-50 yearswith a total of 16 respondents (3.9%) and finally the age group of 51-60 years with 1 respondent (0.2%).

In terms of education and qualification level, a total of 72.5% of the respondents had Under Graduate (UG) qualification, and 26.3% had Post Graduate (PG) qualifications, 1% had a Diploma degree and 0.2% had PhD.

4.3.2 Description statistics

There are several statistical measures of central tendency or "averages". In this study, we use two averages which are as follows:

- Arithmetic Mean
- Median

Arithmetic mean:

Arithmetic mean is the most commonly used measure of central tendency. It is defined as the sum of the values of all observations divided by the number of observations and is usually denoted by X. In general, if there are N observations as X_1, X_2, X_3, ..., X_N, then the Arithmetic Mean is given by

$$\overleftarrow{X} = \frac{X_1+X_2+X_3+\ldots\ldots+X}{N}$$

Median:

The median is the variable's positional value that splits the distribution into two equal parts: one with all values more than or equal to the median, and the other with all values less than or equal to it. When the data set is organised in magnitude, the Median is the "middle" element. Because the median is calculated by the relative positions of many values, it is unaffected by changes in the magnitude of the greatest value, for example.

Computation of median

By sorting the data in ascending order and finding the midway number, the median can be simply calculated.

4.3.3 Test of Independence of association tests:

CHI-SQUARE TEST

Chi-square is widely used to test the independence of attributes. It is applied to test the associate between the attributes when the sample data is presented in the form of a contingency table with any number of rows or columns. The hypothetical decision based on the chi-square test statistic.

$$\chi^2 = \sum_i \frac{(Oi - Ei)^2}{Ei}$$

Where, Ei= $\frac{Ri*Cj}{N}$

R_{ij}- is the sum total of the row in which E_{ij}lies

C_{ij}- is the sum total of the column in which E_{ij}lies

The characteristic of this distribution is completely defined by the number of degrees of freedom v which is given by v=(r-1) (c-1)

Where r = number of rows and c = number of columns

If calculated value of $\chi^2$is less than that theoretical value of χ^2at the level of significanceα, then accept the hypothesis H_0reject null hypothesis.

Chi-square tests were completed to test whether marital status had any relationship with benefits, having children with challengesand change in salary with work-family balance of the employees while working from home.

5. DATA ANALYSIS AND INTERPRETATION

5.1 Satisfaction of the Employees while working from home

From the total of 407 respondents, 135 employees (33.2%) are very satisfied with their work from home arrangement, 143 employees (35.1%) are satisfied with their work from home arrangement, 83 employees (20.4%) are neither satisfied nor dissatisfied with their work from home arrangement, 24 employees (5.9%) are dissatisfied with their work from home arrangement and 22 employees (5.4%) are very dissatisfied with their current work from home arrangement.

The results depict that most of the employees, on an average, are satisfied with their current work from home arrangement and show willingness to adopt this arrangement.

5.2 Benefits of the Employees while working from home

Based on the arithmetic mean 2.2582of the Benefits of working from home, on a 5-point scale, and median 2.2857, we performed Chi square test. Results show that there issignificant association or relationship between Marital status of the employees and Benefits of working from Home, which is χ^2 (2, N = 407) is 7.665, p = 0.022 at 5% level of significance.

Employees who are married, find working from home more beneficial and advantageous than the employees who are unmarried or single. (33% to 67%).

5.3 Challenges faced by the Employees while working from home

Based on the arithmetic mean 2.3390 of the Challenges of working from home, on a 5-point scale, and median 2.2857, we performed Chi square test. Results show that there is a significant association or relationshipbetween Parentingand Challenges faced by the employees, which is χ^2(1, N = 407) is 4.545, p = 0.033 at 5% level of significance.

Employees without children or those who are not parenting, agree that they face challenges while working from home compared to the employees who have children (86.4% to 13.6%).

5.4 Work-family Balance

Based on the arithmetic mean 2.4412of the Work-Family balance while working from home, on a 5-point scale, and median 2.4286,we performed Chi square test. Results show that there is a significant association or relationship between their change in salary during this period of working from home and work-family balance by the employees, which is χ^2(2, N = 407) is 8.214, p= 0.016 at 5% level of significance.

Employees whose salary increased during this pandemic, agree that they are able balance their family time and workmuch better than the employees whose salary did not change during this pandemic or the ones whose salary decreased during pandemic (30.4% to 66.3% and 3.3%).

6. CONCLUSION

The group studied under this research is mature, with 72.5 percent of respondents holding a under graduate degree and an average age of the employees was 25.96 years.

The findings of the study reveal that while working from home, there is more work independence and flexibility over their work scheduleand also helps in saving time and money compared to going office regularly.On an average, most of theemployees agree that they are highly confident about their work and feel very satisfied about their job. While working from home, employees find it easy to avoid politics at their workplace and do their job at ease, resulting in growth of their career to which most of them agreed.

But there were also many difficulties and challenges that the employees have to cope up in their home arrangement. Employees find it hard to distinguish between work time and home time which results in improper balance of priorities leading to physical health as well as mental health changes.They also agree that theyfeel isolatedwhile working, from their work place network. The biggest challenge among all these is to coordinatecomplex tasks with other employees via online platform and physical interaction with their mentors as well as co-workers posed a great difficulty to work. Even stress had a weak positive relation to experience with virtual work.

Amidst all these challenges, the study revealed that the employees had more conducive and relaxed environment at home. This helps them to balance the job, family as well as personal life while working at home. The employees agree that they are able to spend quality time with their familiesduring this pandemic. And most importantly, Families of the employees were more supportive and they too adapted soon to this work arrangement.

REFERENCES

1. *Aczel B, Kovacs M, van der Lippe T, Szaszi B, (2021), Researchers working from home: Benefits and challenges. PLOS ONE 16(3): e0249127, pp. 3-5.*
2. *Khristian S. Liwanag, 2021, A case study on the advantages and challenges of work from home (WFH) scheme during pandemic, International engineering journal for research development, Vol.6(5), pp. 4-16.*
3. *Dr.Sridevi.R and Sanjana.N, 2021,A study on the impact of Work From Home among IT employees during covid 19 in Coimbatore District, UTKAL Historical Research Journal, Vol.-34(X), pp. 172-178.*
3. *Ajay K Garg and Jan van der Rijst, 2015, The benefits and pitfalls of employees Working from Home: Study of a private company in South Africa, Corporate Board: Role, Duties & Composition, Vol 11(2), pp. 36-45.*
4. *ItishreeMohanty, 2021, Quality of work life during working from home: covid 19 pandemic situation, International Journal of Applied and Natural Sciences (IJANS), Vol 10(1), pp. 23-31.*
5. *LinaVyas&NantapongButakhieo, 2021, The impact of working from home during COVID-19 on work and life domains: an exploratory study on Hong Kong, Policy Design and Practice, Vol 4(1), pp. 59-76.*
6. *Heejung Chung, HyojinSeo, Sarah Forbes, and Holly Birkett, 2020, Working from home during the covid-19 lockdown: Changing preferences and the future of work, University of Birmingham, pp. 10-38.*
7. *G.Ragavi, 2021, A study on impact of Work From Home among employees during pandemic covid-19, International Journal of Scientific Research and Engineering Development, Vol 4(5), pp. 1206-1210.*
8. *Michael Gibbs, FriederikeMengel, and ChristophSiemroth, 2021, Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals, Becker Friedman Institute, pp. 4-35.*
9. *Teresa Galanti, Gloria Guidetti and Ferdinando Toscano, 2021, Work From Home During the COVID-19 Outbreak: The Impact on Employees' Remote Work Productivity, Engagement, and Stress, J Occup Environ Med, Vol 63(7), pp.*

DOI No. : 10.5958/2394-9309.2022.00050.6
Aryabhatta Journal of Mathematics & Informatics Vol. 14, Special Conference Issue 30-31 March 2022 ISSN (Print) : 0975-7139
Double-Blind Peer Reviewed Refereed International Journal ©www.abjmi.com ISSN (Online) : 2394-9309

AN OVERVIEW OF SIMILARITY MEASURESIN CLUSTERING

Dr. G .R. Kanmani, Dr. S. Dhivya Devi
Guest Lecturer, Department of Statistics, University of Madras, Chennai-600 005
Assistant Professor, Department of Statistics, Shri ShankarlalSundarbaiShasun Jain College For Women, Chennai-600 017
E-mail : kanmaniashok82@gmail.com, dhivyadevi@shasuncollege.edu.in

ABSTRACT
Clustering is an unsupervised learning technique used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is the process of grouping a set of similar objects into different groups, the partitioning of a data set into subsets, so that the data in each subset according to some defined distance measure. This paper discuss about the various clustering methodsthe current similarity measures based on distance based clustering, it alsoexplains the limitations analogous with the existing clustering techniques and propose that the combination of the advantages of the existing systems can help overcome the limitations of the subsisting systems.
Keywords- Similarity, Similarity measures, Distance based.

1. INTRODUCTION

1.1 Clustering

Clustering using distance functions, called distance based clustering, is a very popular procedure to cluster the objects and has given good results. The clusters are formed in such a way that any two data objects within a cluster have a least distance value and any two data objects across different clusters have a highest distance value.

1.2 Similarity of data

Similarity is an amount that reflects how data samples are close to each other. it represents how the data are having similar patterns are. Clustering is grouping similar an object together based on a similarity. This similarity measure is most usedin most applications based on distance functions such asChebychev, Mahalanobis, Spearman, Chi-Square, Euclidean distanceetc. to group objects in clusters. The clusters are established in such a way that any two data objects within a cluster have a minimum distance value and any two data objects beyond different clusters have a maximum distance value. In this paper many limitations associated with distance measures based clustering which have been addressed and are aiming to overcome in our research.

1.3 Types of Clustering

Clustering algorithms can be classifybroadly into the following categories:

1. Partitional Clustering

2. Density based Clustering
3. Hierarchical clustering

1.3.1 Partitional Clustering

It is a popular category of clustering algorithm. Partition clustering algorithm divides the data points into “k“ partitions, the partitions is done based on objective function where each segregation represents a cluster. “The clusters are formed such that theobjects that are “similar “to one another within the same cluster and are “dissimilar” to the objects in other clusters”.When the number of clusters required are static then Partitional clustering methods can be applied. K-means, PAM (Partition around mediods) and CLARA are a few popular examples for the partitioning clustering algorithms.

1.3.2 Density Based Clustering

Density-based clustering algorithms create a region like arbitrary-shaped clusters in which the density of data objects exceeds a particular threshold value. Anwell-known example for Density based clustering approach is DBSCAN algorithm.

1.3.3 Hierarchical Clustering

Hierarchical clustering algorithms work to divide or integrate a particular dataset into a sequence of nested partitions. The hierarchy of these nested partitions can be of two types, viz., agglomerative, i.e., bottom-up or divisive, i.e., top-down. In the agglomerative method, clustering started with a single data object in a single cluster and resume to cluster the closest pairs of clusters until all the data objects are grouped in conjunction as just one cluster. Divisive hierarchical clustering, Moreover it starts with all data objects in a single cluster and keeps splitting larger clusters into smaller ones till all the data objects are split into unit clusters. Balance Iterative Reducing and Clustering using Hierarchies CURE (Cluster Using Representatives) are examples of Hierarchical clustering approach

2. **METRIC:**

Similarity is measured with the help of Distance metrics among the data objects. Obtaining an appropriate distance similarity function is the basic main requirement of metric calculation in a specific problem. A metric function or distance function is a function that describe a distance between elementsobjects of a set. Metric spaceplays a critical role in clustering techniques. In this paper, an example of the basic algorithm for k-means clustering algorithm using Euclidean distance metric is given.

A given distance(e.g. dissimilarity) is signified to be a metric if and only if it satisfies the following four conditions:

1- Non-negativity: $d(x, y) \geq 0$, for any two distinct observations p and q.

2- Symmetry: $d(x, y) = d(y, x)$ for all x and y.

3- Triangle Inequality: $d(x, y) \leq d(x, r) + d(y, q)$ for all x,y, r.

4- $d(x, y) = 0$ only if $x = y$.

Distance measures are the fundamental principle for classification, like the k-nearest neighbour’s classifier algorithm, which measures the dissimilarity between given data samples.

2.1.1 Euclidean distance

Euclidean distance is considered as the standard metric for numeric attributes or features. It is simply the typical distance between two points. Euclidean distance is extensively used in clustering problems, including clustering text. The renege distance measure used with the K-means algorithm is also the Euclidean distance. The Euclidean distance is given by

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

One of the most prominent classification algorithms, the KNN algorithm, can benefit from using the Euclidean distance to classify data.

2.1.2 Manhattan Distance

This metric will measure the distance between two places in a given city. In Manhattan distance the distance can be measured in terms of the number of blocks that disparate two different places. Itis defined as the absolute differences between coordinates of pair of data objects.

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

2.1.3 Canberra Distance

It is a weighted form of manhattan distance used in Clustering, like Fuzzy Clustering, classification, computer security, and ham/spam detection systems. It is more robust to outliers in contrast to the foregoing metric.

$$d_{ij} = \sum_{k=1}^{n} \frac{|x_{ik} - x_{jk}|}{|x_{ik}| + |x_{jk}|}$$

2.1.4 Chebyshev Distance

The Chebyshev distance among two n-D observations or vectors is equal to the maximum absolute value of the variations allying the data samples' coordinates. In a 2-D world, the Chebyshev distance between data points can be set on as the sum of absolute differences of their 2-dimensional coordinates.

$$d(\mathbf{x}, \mathbf{y}) = \lim_{r \to \infty} \left(\sum_{k=1}^{n} |x_k - y_k|^r \right)^{1/r}.$$

3. EXAMPLE OF USING A DISTANCE METRIC IN CLUSTERING

Algorithm K-means using basic Euclidean distance metric.

Let X be the set of data objects and Let V= be the set of centers.

1. Randomly choose the cluster centers.
2. Calculate the distance between each data object and cluster centers.

3. Assign data object to the cluster center whose distance from the cluster centeris minimum of all the cluster centers.
4. Calculate new cluster center using equation.
5. Again calculate the distance between each data and new obtained cluster centers.
6. Repeat steps 3 to 5 till no data object was reassigned.

4. LIMITATIONS AND FINDINGS

1. Distance based metrics is used mostly in all types of clustering
2. Correlation among the data objects cannot be captured good enough using Distance metrics.
3. Similar data patterns will exist while measured by the distance metrics.
4. Very less research exist in clustering negative data.
5. Most of the similarity concepts for clustering are not robust.

REFERENCES

1. *Aggarwal, C. C.; Hinneburg, A.; and Keim, D. A. 2001. On the surprising behavior of distance metrics in high dimensional space. Database Theory—ICDT 2001 420–434.*
2. *Ankerst, M.; Breunig, M. M.; Kriegel, H.-P.; and Sander, J. 1999. OPTICS: ordering points to identify the clustering structure. ACM SIGMOD Record. Chen, M.-S.; Han, J.; and Yu, P. S. 1996.*
3. *Data mining: an overview from a database perspective. Knowledge and Data Engineering, IEEE Transactions on 8(6):866–883.*
4. *Defays, D. 1977. An efficient algorithm for a complete link method. The Computer Journal 20(4):364–366. Ester, M.; Kriegel, H.-P.; Sander, J.; and Xu, X. 1996.*
5. *A densitybased algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data mining, 226–231.*
6. *Laurent Galluccioa , Olivier Michelb, Pierre Comonb, Mark Kligerc, Alfred O. Herod, Clustering with a new distance measure based on a dual-rooted tree, Information Sciences Volume 251, 1 December 2013, Pages 96-113, Elsevier.*
7. *SuphakitNiwattanakul, JatsadaSingthongchai, EkkachaiNaenudorn, Supacha-nun Wanapu, Using of Jaccard Coefficient for Keywords Similarity, Proceedings of the International MultiConference of Engineers and Computer Scientists 2013, Vol I, IMECS 2013, March 13 - 15, 2013, Hong Kong.*
8. *Archana Singh, Avantika Yadav, Ajay Rana, K-means with Three different Distance Metrics, International Journal of Computer Applications, Volume 67, No.10, April 2013.*
9. *Jian Pei , Xiaoling Zhang , Moonjung Cho , Haixun Wang , Yu, P.S. , MaPle:a fast algorithm for maximal pattern-based clustering, Data Mining, 2003. ICDM 2003. Third IEEE International Conference, Pages 259 - 266.*

FORM IV

Statement of relevant particulars regarding **"ARYABHATTA JOURNAL OF MATHEMATICS & INFORMATICS".**

Published in accordance in Form -1, Rule-3 of the Press and Registration of News Paper/Journals/Books Act, 1867.

1.	Place of Publication	:	2255, Sector-17, Urban Estate, HUDA, JAGADHRI (Yamuna Nagar) - 135 003 Haryana - INDIA
2.	Periodicity of Publication	:	Twice a year
3.	Printer's Name	:	Dr. T.P. Singh
	Nationality	:	Indian
	Address	:	2255, Sector-17, HUDA, JAGADHRI (Yamuna Nagar) - 135 003 Haryana - INDIA
4.	Publisher's Name	:	Dr. T.P. Singh
	Nationality	:	Indian
	Address	:	2255, Sector-17, HUDA, JAGADHRI (Yamuna Nagar) - 135 003 Haryana - INDIA
5.	Editor's Name	:	Dr. T.P. Singh
	Nationality	:	Indian
	Address	:	2255, Sector-17, HUDA, JAGADHRI (Yamuna Nagar) - 135 003 Haryana - INDIA
6.	Name and Address of the owner	:	Aryans Research & Educational Trust (Regd. No. 2655)

I, Dr. T. P. Singh, hereby declare that the particulars given above are true to the best of my knowledge and belief.

Dr. T. P. Singh

Editor-in-Chief / Printer & Publisher

SUBSCRIPTION FORM

I/We wish to subscribe **Aryabhatta Journal of Mathematics & Informatics.** The Filled performa is :

NAME & ADDRESS OF SUBSCRIBER

Name ___________________________ Organisation__________________________

Designation________________________E-mail : ___________________________

Address of Communication____

____

Subscription : Annual/2 years/Life subscription

PAYMENT

I/We wish to make the payment through (Please tick ✓)

☐ Crossed Cheque ☐ Demand Draft

enclosed crossed cheque/demand draft No._______________Dated______________

Drawn on________________________________in favour of **ARYANS RESEARCH & EDUCATIONAL TRUST** for Rs._______________payable at Jagadhri/Yamuna Nagar

or Direct deposit through NEFT/RTGS : Name of Beneficiary : **ARYANS RESEARCH & EDUCATIONAL TRUST** A/C No. : 05522151010233, Bank Name : Punjab National Bank (PNB) Jagadhri (Haryana) IFSC Code : PUNB0055210

Address to be sent :
Dr. T.P. Singh
Chief Editor
2255, Sec-17, HUDA, Jagadhri - 135003 (Haryana) India

Signature of Subscriber

Subscription Information

	India	Abroad
Subscription for individuals/research scholars	Rs. 700/- per issue	US$ 30/-
Life subscription for individuals/research scholars	Rs. 7,000/-	US$ 200/-
Annual subscription for Libraries / Institutions	Rs. 2,000/-	US$ 40/-
Subscription for two years	Rs. 3,000/-	US$ 60/-
Life subscription for Libraries / Institutions	Rs. 8,000/-	US$ 400/-

Including mailing expenses.

Payments : Payments should be made in favour of **"ARYANS RESEARCH & EDUCATIONAL TRUST"** through **Demand Draft** of any Nationalized Bank payable at **JAGADHRI/YAMUNANAGAR** (Haryana) India.