



CHANGING TRENDS OF PREFERENCES IN MODE OF TRANSACTIONS-A PREDICTION USING ROUGH SET THEORY

Preethi M.

MSc Information Technology

M.O.P Vaishnav College for Women

Chennai-,600034 India

e-mail: m.preethiy@gmail.com

ABSTRACT

Rough Set Theory is a new technique that deals with fuzziness and improbability stressed in decision making. Data mining is a discipline that has an important contribution to data analysis, discovery of new significant knowledge, and independent decision making. The rough set theory offers a feasible approach for decision rule extraction from data. The introduction of demonetisation resulted in elimination of high valued currency notes. It aimed to achieve the goal of a 'less cash' society. Digital trades bring in better scalability and responsibility. Recently RBI has also disclosed its document- "Payments and Settlement Systems in India: Vision 2018" boosting the electronic payments and to help INDIA grow from cash to cashless society in the long run. Thus giving this model an overlook, this paper focuses on studying the views of people on evolution of cashless economy and their comfort level with it. The study was conducted in Chennai; data was collected with the help of organised questionnaire and analysed using rough set theory.

Keywords: Data mining, analysis, Rough Set Theory, cashless transaction, objects

I. INTRODUCTION

To generate information it needs enormous collection of data. It is essential, to develop powerful tool for interpretation of such data and for the abstraction of exciting knowledge that could help in decision-making. The only solution is 'Data Mining'. Data mining is a process of extraction of useful information and patterns from huge and unstructured data. The motive of data mining effort is either to generate a descriptive model or a predictive model. A descriptive model presents the main features of the data set and a predictive model is to predict an unknown value of a specific variable; the target variable. There are many techniques available in data mining in which regression is of the method. This strategy is adopted for prediction. Independent variables are features which are already known and are used for prediction of decision values for a new object. Rough Set Theory is one of the powerful tools which can be implemented to find hidden patterns and generate rules. This technique is used to deal with imprecise data.

Cash is any legal medium of exchange that is free of restrictions. The demonetisation effect has given incredible enhancement to the cashless transactions because of the hindrance of high denomination currencies. The circulation of hard cash became minimal. The spread of digital payment started to fulfil the goal of demonetisation. Apart from that, it has become very advantageous to common people like convenience, discounts, budget discipline

etc. Some of the cashless modes are cheque, Demand draft, debit card, credit card, net banking, mobile and e-wallets etc.

The main objective of the paper is to find the preferences of people regarding the mode of transactions and their comfort level with the cashless transactions. People preferences were studied by collecting data from them in the form of questionnaire and analysed by means of data mining approaches. Regression method is used to identify the type of attributes and rough set theory has been used for prediction purposes.

II. REVIEW OF LITERATURE

[1] Discusses about the situation of the country after the arrival of demonetization. The nature of the paper is very descriptive. [1] has covered the consequences, advantages and the disadvantages which were brought by demonetisation. [1] also deliberates the impacts of going cashless and concludes by giving suggestions on what kind of steps should the government take for improving the cashless economy. [2] remarks the exact information about how and when did the tremendous change evolved. [3] gives the picture of many countries of the world which has cashless economy and its features. Then the strengths of India going cashless, weaknesses faced and the available opportunities are conversed and concluded with the threats in going cashless [5][7] explains the rough set theory basic concepts and terminologies with suitable examples and enlightens the process of deriving decision patterns for predictions.

III. METHODOLOGY

Knowledge Discovery process consists of various steps. The steps are collecting data, analysing data sets to discover patterns, building model, evaluation the model using one of the data mining techniques. In this paper, Rough Set Theory is used for evaluating the model and for prediction of the mode preferred by people.

IV. ROUGH SET THEORY

One of the mathematical approaches in data mining is rough set theory. Observations show that the application of this theory has grown world-wide. ZdzislawPawlak projected this theory. This concept is considered as a mathematical tool for imperfect data analysis. It is first non-statistical approach in data mining. With every object in the universe, we subordinate some other object which is comparable to it. This acts as the basic concept for the rough set theory. Objects are categorised based upon the information available with that of the objects whose data are already available. All such objects form a set which is known as elementary sets. With the available knowledge of data, a boundary has to be assigned for each set. The boundary region of the rough

set is given by the variance of the lower and upper approximation. Lower approximation is the depiction of objects which belongs to the subset whereas the upper approximation is the description of objects which most likely belong to that set. The main benefit of rough set theory is that it helps in generating rules from data. In this method, data is structured as a table which is known as information systems. The records denote the objects and the columns represent the attributes. The attributes are broadly classified as condition and decision attributes. The decisions are derived from the rows of the table. This method is very easy to understand. Indiscernibility Relation acts as the chief concept in Rough Set Theory which is an equivalence relation.

V. FROM BARTER TO DIGITAL

In earlier days, barter system used goods as a medium of exchange for purchasing goods or for services. Upon the introduction of currency, the barter system slowly transformed to an economy where currency played a vital role for transactions. The transactions were done using currency notes and coins. It greatly facilitated the transactions. Another way of transaction was the digital method which included internet banking, mobile banking etc. Transactions by hard cash paved the way for growth of black economy. This led to severe economic downfall in India. The government made many efforts to bring all the transactions into banking channel for the prevention of circulation of black money and, systematic and orderly growth of India’s economic development. One such effort was the introduction of demonetisation which resulted in elimination of high valued currency notes. This gave rise to the control of cash economy. The cashless transaction is not only safer than the cash transaction but is less time consuming, convenient and lowers risk. People are witnessing the effects of demonetization and many people are left in confusion. Going cashless not only facilitates our lives but also helps authenticate and formalize the transactions. This results in increasing the economic growth by controlling corruption and the flow of black money.

VI. PROPOSED SYSTEM

Table I: Attributes and their values

Sr. No	ATTRIBUTE	DESCRIPTION	POSSIBLE VALUES	CORRESPONDING VALUES
1	NBA	Number of Bank Accounts	1,2,3,4,>4	1,2,3,4,5
2	PACH	Preferred Amount of Cash in Hand	<500,500-1000,1000-2000,2000-3500,>3500	5,4,3,2,1
3	MBTR	Mobile Banking Transaction Regularity	Never, Rare, Sometimes, Mostly, Always	1,2,3,4,5
4	RWAA	Regularity in Withdrawal of Amount from ATM	Never, In case of emergency, monthly once, weekly, daily	5,4,3,2,1
5	OPMB	Option for Paying Monthly Bills	Cash, Online Payment	1,2
6	AASE	Average Amount Spent in a month through E-Payment	<1000,1000-2000,2000-2500,2500-4000,>4000	1,2,3,4,5
7	PASC	Percentage of Amount Spent as Cash in a month	20%,20-35%,35-50%,50-70%,>70%	5,4,3,2,1
8	CCDE	Convenience for Cashless payment for Daily Expenditure	Yes, No, Maybe	3,1,2
9	PIE	opinion about Positive Impact in Environment about minimising cash	Yes, No, Maybe	3,1,2
10	CRT	Opinion on Reduction in Theft and reduction in fake currency	Yes, No, Maybe	3,1,2
11	OCD	Opinion on Cashless transactions increment after Demonetisation	Yes, No, Maybe	3,1,2
12	OUCF	Opinion on Usage of Cash in Future	Strongly disagree, disagree ,neutral, agree, strongly agree	1,2,3,4,5

In this paper, we put forward a technique to predict the preference of people in the mode of transaction after the implementation of demonetisation. Euclidean distance measure is used along with the rough set theory in this proposed approach. Rendering the RST, the collected data is organised in the form of table in which the column represents the attributes. The characteristics considered are: Preference of amount of cash having in hand, Mobile banking transactions, Mode of Payment for monthly bills, Convenience, Amount spent as cash in a month, Safety opinion.

Each person’s response is considered as an object. For each object ‘i’ a vector V_i is produced consisting of n-tuples, where n is the count of variables. $V_i = (a_1, a_2, a_3, \dots, a_n)$ where a_i varies from 0 to n. Using the RST, the redundant and inconsistent values are eliminated and the decision table is formed.

Each value is assigned with the corresponding 5 scale point and the threshold value is determined. To predict the mode preferred for any given object X, the distance formula is used to find the closeness of the object with the member of the vector.

The Euclidean distance formula is

$$\text{Distance (d)} = \min (\text{sqrt} ((x_2-x_1)^2 + (y_2-y_1)^2)) \quad (1)$$

Where x_1, x_2, y_1, y_2 are the points in vectors V_i and V_j

The decision attribute of vector V_i is assigned to the object X whose d is minimum.

VII. PERFORMANCE ANALYSIS AND RESULTS

For the proposed system, data was collected from various students, IT professionals and business people. The data objects are associated with both data and knowledge. The preference of mode of transaction by people is the decision variable. The abbreviations and the description of the attributes used in the technique are given along with their values in Table I.

Table II: Training data

SR.NO	NBA	PACH	MBTR	RWAA	OPMB	AASE	PASC	CCDE	PIE	CRT	OCD	OUCF	Average	Mode Preferred
1	2	1	1	4	1	1	1	2	3	3	3	5	2.25	Cash
2	1	2	2	4	1	2	2	3	2	3	3	5	2.5	Cash
3	2	1	3	4	2	5	4	3	3	3	3	3	3	Cashless
4	1	4	2	2	1	1	1	1	1	2	2	3	1.75	Cash
5	2	5	4	2	2	5	4	3	3	2	3	4	3.25	Cashless
6	2	4	4	4	2	3	4	2	3	3	3	5	3.25	Cashless
7	2	3	2	4	1	1	1	2	3	3	1	3	2.166	Cash
8	1	5	2	4	1	2	1	2	3	2	2	4	2.41	Cash
9	2	4	4	4	2	2	5	2	2	1	1	3	2.66	Cashless
10	2	4	4	4	2	4	4	2	2	3	3	4	3.16	Cashless

Table II is obtained by applying the assigned values to the collected data objects and the average is calculated. Threshold value should be determined for the training data. It is the minimum value to decide the preference. The threshold value for the sample data is 2.5. If the average is greater than 2.5, then the preference is predicted as cashless mode.

From table II, the following vectors are constructed. For this table, the vectors are V_1, V_2, \dots, V_{10} .

$$V_1=(2,1,1,4,1,1,1,2,3,3,3,5)=2.25$$

$$V_2=(1,2,2,4,1,2,2,3,2,3,3,5)=2.5$$

$$V_3=(2,1,3,4,2,5,4,3,3,3,3,3)=3$$

$$V_4=(1,4,2,2,1,1,1,1,1,2,2,3)=1.75$$

$$V_5=(2,5,4,2,2,5,4,3,3,2,3,4)=3.25$$

$$V_6=(2,4,4,4,2,3,4,2,3,3,3,5)=3.25$$

$$V_7=(2,3,2,4,1,1,1,2,3,3,1,3)=2.16$$

$$V_8=(1,5,2,4,1,2,1,2,3,2,2,4)=2.41$$

$$V_9=(2,4,4,4,2,2,5,2,2,1,1,3)=2.66$$

$$V_{10}=(2,4,4,4,2,4,4,2,2,3,3,4)=3.16$$

The decision attribute for the objects are determined based on the threshold value. Given a new object X, the distance between X and V_i is calculated using Euclidean distance formula. The vector with the minimum distance from X is chosen and the decision variable value of vector V_i is the decision for X.

Let X contain the values (1,4,4,4,2,5,4,3,2,3,3,4). By calculating the distance using the Euclidean distance formula, it is found that the distance between X and V_{10} is minimum.

The vectors and object X are represented in the form of graph to identify the patterns.

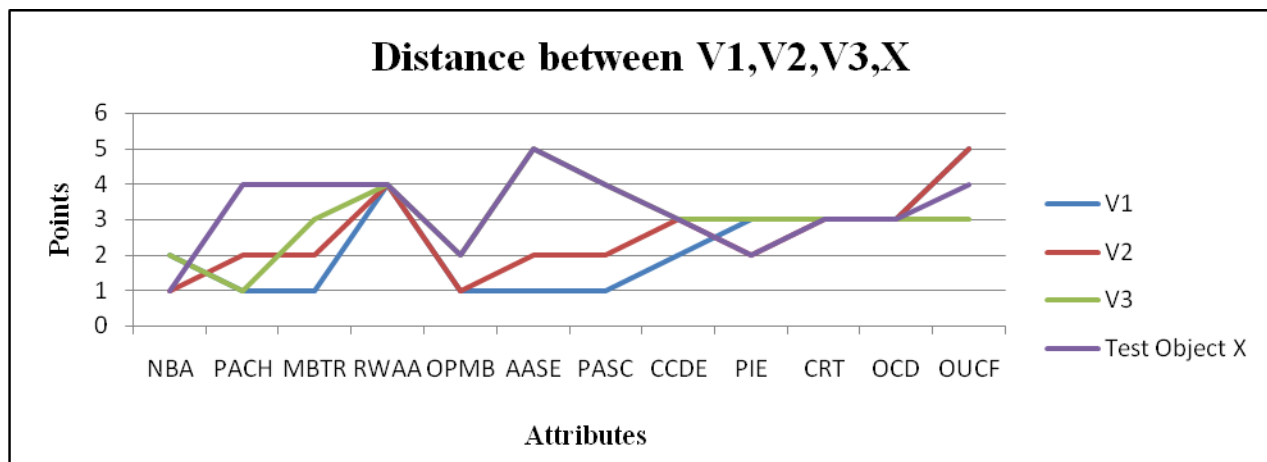


Figure1. Representation of vectors in the form of graph

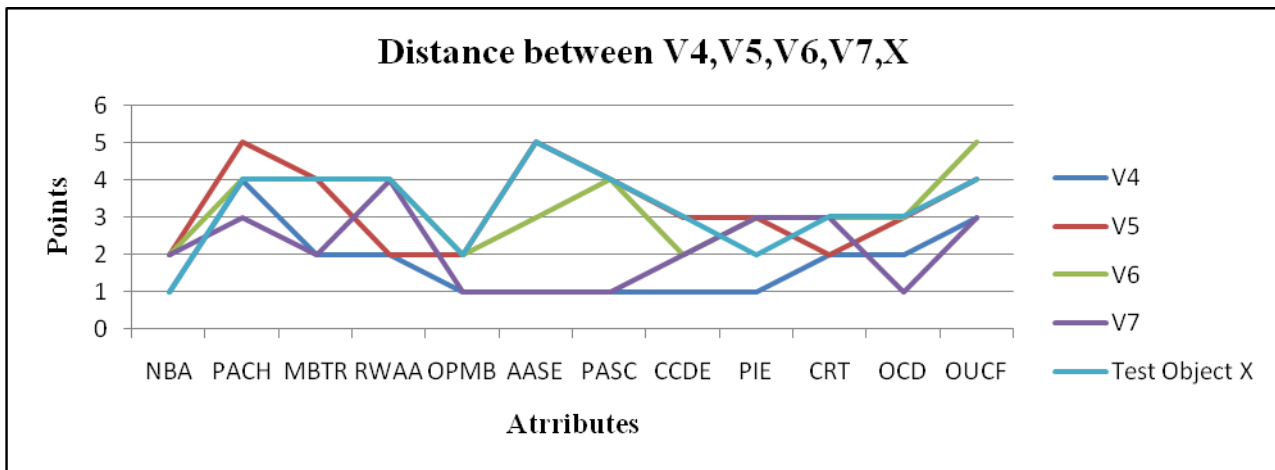


Figure2. Representation of vectors in the form of graph

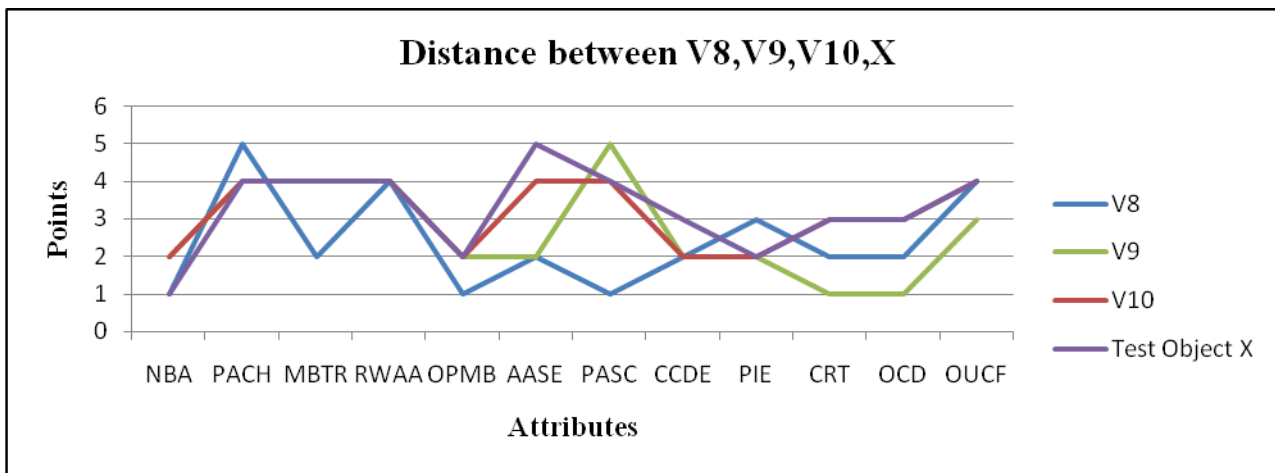


Figure3. Representation of vectors in the form of graph

Figure1, Figure2 and Figure3 shows that vector V_{10} and object X has similar pattern. Both Euclidean distance and the pattern prediction shows that vector V_{10} and object X are similar.

In the similar manner, more test data are evaluated and the decision variable is predicted.

So the decision value of V_{10} is the decision of X. The decision variable for object X is Cashless.

Table III: Sample Test Data

SR.No	NBA	PACH	MBTR	RWAA	OPMB	AASE	PASC	CCDE	PIE	CRT	OCD	OUCF	Mode Preferred
1	2	4	3	4	1	1	1	1	1	1	1	3	Cash
2	2	4	4	4	2	5	5	3	2	3	3	4	Cashless
3	2	4	2	3	1	1	3	1	1	2	3	3	Cash
4	1	5	4	4	2	3	4	2	3	2	3	1	Cashless
5	1	2	4	3	2	5	5	3	2	3	3	2	Cashless
6	3	4	3	4	2	2	4	2	2	3	1	5	Cashless
7	4	1	4	4	2	5	5	3	3	3	2	4	Cashless
8	3	4	5	2	2	5	4	3	3	3	3	4	Cashless
9	1	5	2	4	1	1	1	2	2	2	1	2	Cash

The above model is evaluated for 100 people.

VIII. CONCLUSION

From the above study, 60% of the people have preferred cashless transactions and are convenient with it. It gives the impression that many people agree with the government on the attempt to create a cashless economy which helps to fight against corruption, money laundering. But one of the biggest problems is security. It is important to strengthen Internet Security to protect against online fraud. For smooth implementation of cash less system, some measures are recommended to the Government to bring in transparency and effectiveness in e- payment system. Strategies should be used by government and RBI to encourage cashless transactions by promoting mobile wallets and withdrawing service charge on cards and digital payments. A financial literacy campaign should be conducted by government to make people aware of benefits of electronic payments.

IX. REFERENCES

- [1] Pradeep H. Tawade(2017),“FUTURE AND SCOPE OF CASHLESS ECONOMY IN INDIA”, IJARIII , Vol-2 Issue-3 , ISSN (O)-2395-4396.
- [2] Ramya, N & et.al(2017), “Cashless transaction: Modes, advantages and disadvantages”.
- [3] Mahima, ”CASHLESS ECONOMY: SWOT ANALYSIS FROM INDIAN PERSPECTIVE”, ISBN: 978-93-86171-37-5
- [4] Preetigarg & manvipanchal(Apr. 2017),”Study on Introduction of Cashless Economy in India 2016: Benefits & Challenge’s”, e-ISSN: 2278-487X, p-ISSN: 2319-7668, Volume 19, Issue 4. Ver. II, PP 116-120
- [5] Mohammad(2008),” Application of Rough Set Theory in Data Mining for Decision Support Systems (dsss) “, Journal of Industrial Engineering 1 (2008), PP. 25 - 34
- [6] Silvia Rissino1 ,”Rough Set Theory – Fundamental Concepts, Principals, Data Extraction, and Applications “
- [7] Smita, Priti Sharma(May-Jun. 2014), ”Use of Data Mining in Various Field: A Survey Paper”, e-ISSN: 2278-0661, p-ISSN: 2278-8727Volume 16, Issue 3, Ver. PP 18-21
- [8] Bharati M & et. al. ,”DATA MINING TECHNIQUES AND APPLICATIONS”, Vol. 1 No. 4 301-305.
- [9] Deshpande (2010),”DATA MINING SYSTEM AND APPLICATIONS: A REVIEW” ,International Journal of Distributed and Parallel systems (IJDPSS) Vol.1, No.1