IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

Behavioral Model for Banking Customer based on Neural Network with Time Series

V. Anbalagan[1], A. Malarvizhi[2]

M. Phil., Research Scholar, H.H The Rajah's College (Autonomous), Pudukkottai, India[1]

Research Guide, H.H The Rajah's College (Autonomous), Pudukkottai, India[2]

Abstract: Data mining is an essential tool for any banking CRM strategy to be successful. It not only recognizes patterns to make predictions, but can also highlight available opportunities. With the many advantages and new avenues that it offers, this is one tool that no bank can ignore if it wants to retain its customers and stand out on a highly competitive industry. This study presents a new stage frame work of customer behavior analysis that integrated a neural network with the help of time series algorithm. The time series mining function provides algorithms that are based on different underlying model assumptions with several parameters. The learning algorithms try to find the best model and the best parameter values for the given data. In time series algorithm which roles a main part o calculate a detailed forecast including seasonal behavior of the original tije series. The autoregressive part of the algorithm uses weighed previous values while the moving average part weights the previously assumed errors of the time series. The objective of my project of to identify the most value customer in the banking databases using decomposing the data, and their loyalty. Time series data is helpful characteristics of customer and facilitates marketing strategy development.

Keywords: Data Mining, Neural Network.

I. INTRODUCTION

The important resource in contemporary marketing strategies is Customers. Therefore, it is essential to enterprises and organization to successfully acquire new customers and retain high value customers. To achieve these aims, many enterprises plan to gain their own customers' data with many of database tools which can be analyzed to achieve the customer behavioral and applied to develop new business strategies. Accordingly, assuming there is same pattern between customer behaviors. Many method have been introduced to achieve better knowing of customer behaviors, the "behavioral scoring models" is one of the most successful technique that help decision makers to realize their customer behaviors. Behavioral scoring models help to analyze purchasing behavior of customers .The banking industry around the world has undergone a tremendous change in the way business is conducted. Leading banks are using Data Mining (DM) tools for customer segmentation and profitability, credit scoring and approval, predicting payment default, marketing, detecting fraudulent transactions, etc.

Today, data mining is being used by several industries including banking and finance, retail, insurance, telecommunications, etc. Other possible applications for data mining include database marketing, sales forecasting, call behavior analysis and churning management in telecommunications; forecasting of demand for utilities, such as energy and water; simulation of chemical and other process reactions; finding critical factors in discrete manufacturing (aerospace, automobile, electronics); CPU usage and forecasting. Data mining is often referred to as 'analytical intelligence' .Currently, huge electronic data repositories are being maintained by banks and other financial institutions across the globe. Valuable bits of information are embedded in these data repositories. The huge size of these data sources make it impossible for a human analyst to come up with interesting information that will help in the decision making process.Data mining can contribute to solving business problems in banking and finance by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because the volume data is too large or is generated too quickly to screen by experts.

Customer Behavior Modeling (CBM) or customer profiling is a tool to predict the future value of an individual and the risk category to which he belongs to based on his demographic characteristics, life-style and previous behavior. This helps to focus on customer retention.To achieve these aims, many enterprises plan to gain their own customers' data with many of database tools which can be analyzed to achieve the customer behavioral and applied to develop new business strategies. Economic theory has established that a business derives 80% of its income from 20% of its customers. However, instead of targeting all prospects equally or providing the same offers to all customers, enterprises select only those individuals that meet specified profitability levels based on previous behavior or individual needs. Accordingly, assuming there is same pattern between customer behaviors.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

II. LITERATURE SURVEY

Clustering the Time Series Data

T.W. Liao developed a two-step procedure for clustering multivariate time series of equal or unequal length. The first step applies the k-means or fuzzy c-means clustering algorithm to time stripped data in order to convert multivariate real-valued time series into univariate discrete-valued time series. The converted variable is interpreted as state variable process. The second step employs the k-means or FCM algorithm again to group the converted univariate time series, expressed as transition probability matrices, into a number of clusters. The traditional Euclidean distance is used in the first step, whereas various distance measures including the symmetric version of Kullback–Liebler distance are employed in the second step.

T.W. Liao applied several clustering algorithms including K-means, fuzzy c-means, and genetic clustering to multivariate battle simulation time series data of unequal length with the objective to form a discrete number of battle states.

Literature Survey of Representation-Based Clustering Approach

T.-C. Fu described the use of self-organizing maps for grouping data sequences segmented from the numerical time series using a continuous sliding window with the aim to discover similar temporal patterns dispersed along the time series. They introduced the perceptually important point (PIP) identification algorithm to reduce the dimension of the input data sequence in accordance with the query sequence. The distance measure between the PIPs found was defined as the sum of the mean squared distance along the vertical scale (the magnitude) and that along the horizontal scale (time dimension). To process multi resolution patterns, training patterns from different resolutions are grouped into a set of training samples to which the SOM clustering process is applied only once. Two enhancements were made to the SOM: filtering out those nodes (patterns) in the output layer that did not participate in the recall process and consolidating the discovered patterns with a relatively more general pattern by a redundancy removal step.

Dong Jixue for mining the financial time series uses the wave cluster, which is a kind of grid cluster and the density cluster unify. In this, basic details and methods of phase space reconstruction are analyzed in details. All of these provided the theoretical basis and technical feasibility to time series data mining based on phase space reconstruction. After contrasting the different means of time series pattern mining, the problem of Time Series Data Mining framework TSDM is pointed out, and the temporal patterns mining method based Wave cluster is systematically presented. By the multiresolution property of wavelet transformations and the grid-based partition method, it could detect arbitrary-shape clusters at different scales and levels of detail.

Huiting Liu proposed a new similar pattern matching method. Firstly, trends of time series are extracted by empirical mode decomposition, and the trends are translated into vectors to realize dimension reduction. Secondly, the vectors are clustered by a forward propagation learning algorithm. Finally, all the series that is similar with the query are found by calculating Euclidean distance between the query and the series that belong to the same category with it.

III.PROBLEM DESCRIPTION

EXISTING SYSTEMS

The existing system determine the new two-stage framework of customer behavior analysis using K-means clustering algorithm and an association rule inducer for analyzing bank databases. For differentiation purposes, the system grouped customers with shared customer behavior and RFM value. After briefly reviewed the customer profiles using the association rule inducer, the customers with a higher CP or RFM might be the target customer groups of precedence. The existing customers were divided into three profitable groups of customers according to their shared behavior and characteristics. Marketers then can infer the profiles of customers in each group and propose management strategies appropriate to the each group. This method of analyzing bank databases was analyzing slow process of the large system.

PROPOSED SYSTEM

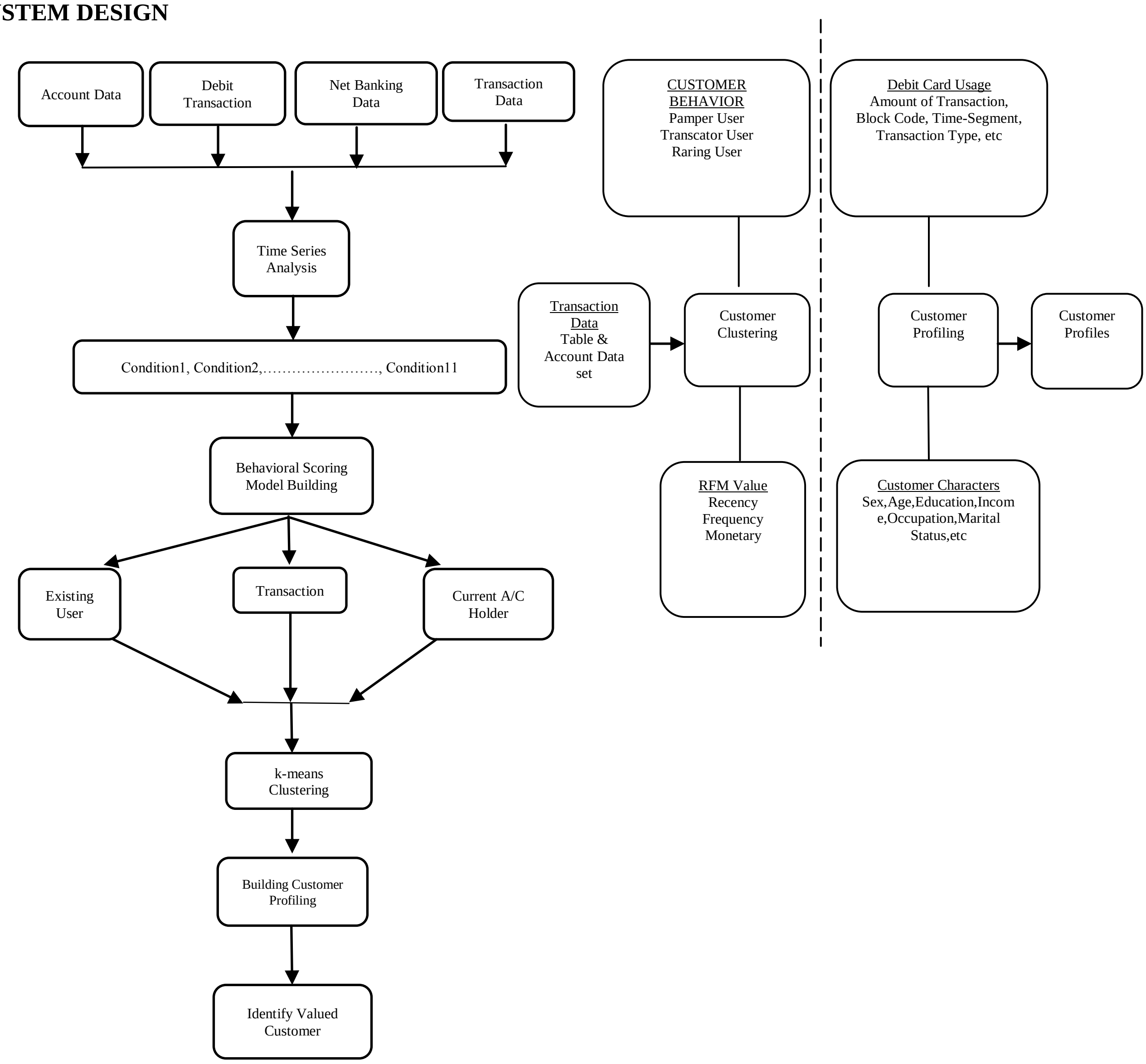
Sequence ranking is critical for many CRM applications such as customer churning behavior classification. A major problem for this research is how to resolve the many sequences which are ambiguous in that they confuse between positive and negative instances during classification.
Also to analyze customer characteristics and behaviors with appropriated criteria: access time, transaction access. The benefits are valuable for the bank to improve services. These results might be benefit for banking marketing team to launch a suitable promotion for appropriate clustering.To implement the Time series transformation approach which

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

one to find the Maximum Likelyhood (Calculate Mean) values with the use of EM (Expectation Maximization) Algorithm. Marketers then can infer the profiles of customers in each group and propose management strategies appropriate to the each group. This study provides a new method of analyzing bank databases. Beyond simply understanding customer value, the bank gains the opportunities to establish better customer relationships.

IV.METHODOLOGY

SYSTEM DESIGN



ANALYZING THE CONSUMER BEHAVIOR

Over the years Data mining (DM) can used to understand the consumer buying behavior using various techniques. Data mining has gradually increases many folds and today it is a giant 100 billion dollar industry. In data mining world every activity of a consumer in a supermarket is treated as a byte of data. How the consumer spends, which day what time normally he/she does the shopping, what they buy most often, how much they buy, in that locality etc. All this data which is gathered somewhere at the backend about which a consumer is not even aware and there is a big industry which is slicing & dicing this data & selling it at a premium price.Banks seeking newer and better ways to diffrentiate themselves from their competitors, customer clustering one of important way to rich this result; Customer clustering is the use of past transaction data to divide customer to the similar groups. The results produced are based on the assumptions that the customer behavior follows patterns similar to past pattern and repeats in the future. Therefore, there could not be a better time than now to analyze the importance of an effective new marketing strategy using the customer behavior analyze.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

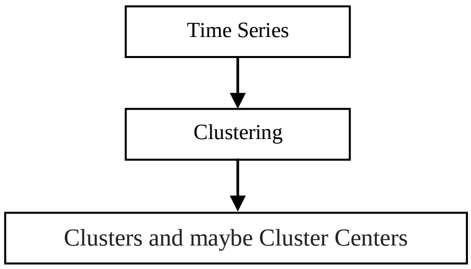
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

TIME SERIES DATA ANALYSIS

Just like static data clustering, time series clustering requires a clustering algorithm or procedure to form clusters given a set of unlabeled data objects and the choice of clustering algorithm depends both on the type of data available and on the particular purpose and application. As far as time series data are concerned, distinctions can be made as to whether the data are discrete-valued or real-valued, uniformly or non-uniformly sampled, unvaried or multivariate, and whether data series are of equal or unequal length.

Non-uniformly sampled data must be converted into uniformed data before clustering operations can be performed. This can be achieved by a wide range of methods, from simple down sampling based on the roughest sampling interval to a sophisticated modeling and estimation approach. Various algorithms have been developed to cluster different types of time series data. Putting their differences aside, it is far to say that in spirit they all try to modify the existing algorithms for clustering static data in such a way that time series data can be handled or to convert time series data into the form of static data so that the existing algorithms for clustering static data can be directly used.

The former approach usually works directly with raw time series data, thus called raw-data-based approach, and the major modification lies in replacing the distance/similarity measure for static data with an appropriate one for time series. The latter approach first converts a raw time series data either into a feature vector of lower dimension or a number of model parameters, and then applies a conventional clustering algorithm to the extracted feature vectors or model parameters, thus called feature- and model-based approach, respectively.



	CONDITIONS
1	Account Opening Date to Current date has more than 5 years
2	If the account type is CURRENT means, account has minimum of 3 years else if the account type is SAVING has minimum of 5 years also satisfy the account balance of Rs.10000
3	Maximum of 10 Net Banking Transactions, Not more than 5 penalties paid in a Year.
4	Maximum 5 times deposit more than Rs.50000 in a year.
5	Below 3 times withdraw the maximum amount of Rs.50000 in his account in one year.
6	Maintain above Rs.50000 in last 6 months
7	Account type BUSINESS, Daily Transact Minimum of Rs.20000 and Account balance maintain Rs.25000
8	According to the Account holder’s Age
9	According to the Account holder’s Occupation like Government Staff, Businessmen
10	No account balance decrease Rs.5000 in more than 3 times in a year.
11	Minimum 10 times usage of ATM Debit card transactions in a year.

CUSTOMER BEHAVIORAL SCORING

Credit and behavioral scoring models (Thomas, 2000) are one of the most successful applications of statistical and operational research modelling in finance and banking, and the number of scoring analysts in the industry is constantly increasing. The main objective of both credit and behavioral scoring models is to classify customers into groups (Lancher et al., 1995). Hence scoring problems are related to the field of classification analysis (Hand, 1981; Johnson & Wichern, 1998; Morrison, 1990). Applying to bank databases, classification analysis for credit scoring is used to categorize a new applicant as either accepted or rejected with respect to his features such as age, income and martial status (Chen & Huang, 2003).

Until now, the building of both scoring models has been always based on a pragmatic approach; because of this, the best and most standard scoring models for every unique circumstance most certainly does not exist. Most previous studies have focused on building more accurate credit or behavioral scoring models and increasing the accuracy of the classification model with various kinds of statistical techniques.analyzing bank databases for customer behavior

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

management is difficult since bank databases are multi-dimensional, comprising of monthly account records and daily transaction records (Donato et al., 1999). Therefore, even with highly accurate scoring models, some misclassification patterns appear frequently.

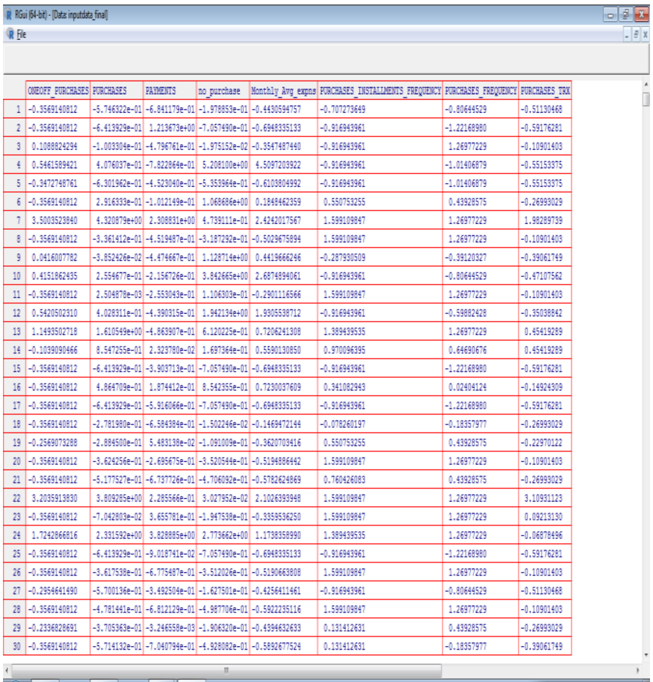
This study intended to draw much from data mining perspectives. Providing a general integrated data mining and behavioral scoring model for customer behavior analysis, which includes necessary preprocessing of the real-world data sets, scoring predicators derivation and customer profiling in order to support a standard model building process will be of great utility. The framework of two-stage behavioral scoring model serves as a tool to validate the effect of data mining techniques in practical scoring analysis applications. In the business world, the most successful application of behavioral scoring model is embodied into databases, which is an approach of analyzing customer histories, looking for similar behavioral patterns among existing customer preferences and using those patterns for a targeted selection of existing or future customers The decisions to be made include which target groups of customers will be encouraged to spend more, what credit line to assign, whether to promote new products to particular groups of customers, and, if the repayment ability turns bad, how to manage debt recovery.

IV. ALGORITHM

Time-Series Transformation Algorithm
ClusterTrans(BankDB, p+, p−)
Input **:** Training database BankDB, numbers of clusters p+ and p−;
Output **:** Transformed training data V ectors;
Steps Algorithm
Step 1: Partition BankDB(based on iterations) into a positive subset DB+ and a negative subset DB−;
Step 2: Clusters+ = K-MeansClustering(DB+, p+);
Clusters− = K-MeansClustering(DB−, p−);
Step 3: Model = (Clusters+,Clusters−); Vectors={};
Step 4: For each input datum $seq_i \in$ BankDB, do
Step 5: $vector_i$ = maxlikelihood(seq_i,Model);
Step 6: Vectors = V ectors ∪ v $ector_i$;
Step 7: End For
Step 8: Return V ectors

V.IMPLEMENTATION

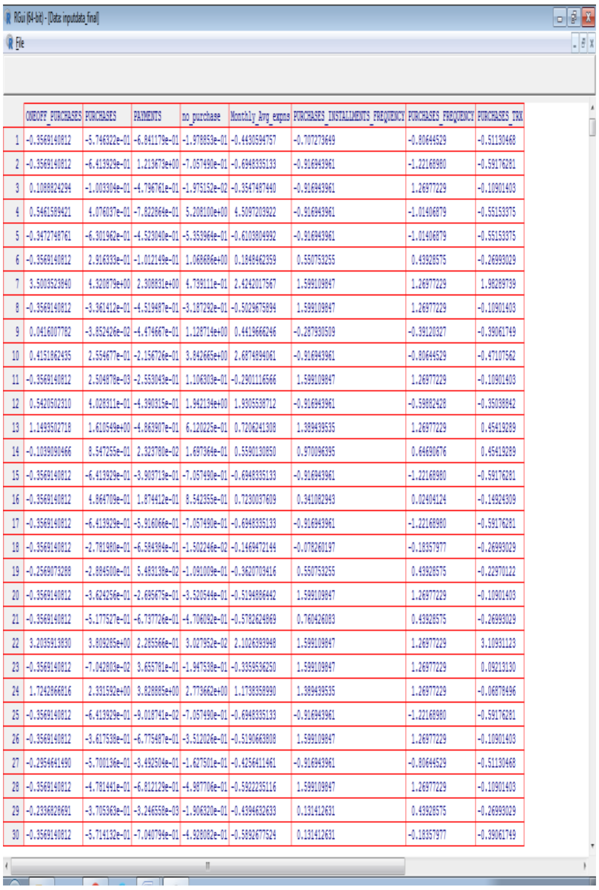
Bank Database is splitter into two subsets like DB+ and DB- depending upon the scoring mean values.



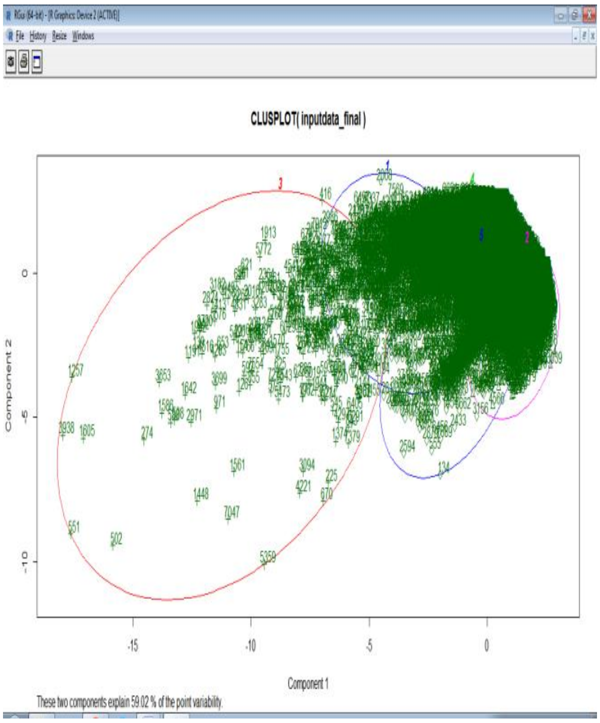
	ONEOFF_PURCHASES	PURCHASES	PAYMENTS	no_purchase	Monthly_Avg_expns	PURCHASES_INSTALLMENTS_FREQUENCY	PURCHASES_FREQUENCY	PURCHASES_TRX
1	-0.3569140812	-5.746322e-01	-6.841179e-01	-1.978853e-01	-0.4430594757	-0.707273649	-0.80644529	-0.51130468
2	-0.3569140812	-6.413929e-01	1.213673e+00	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
3	0.1088824294	-1.003304e-01	-4.796761e-01	-1.975152e-02	-0.3547487440	-0.916943961	1.26977229	-0.10901403
4	0.5461589421	4.076037e-01	-7.822864e-01	5.208100e+00	4.5097203922	-0.916943961	-1.01406879	-0.55153375
5	-0.3472748761	-6.301962e-01	-4.523040e-01	-5.353964e-01	-0.6103804992	-0.916943961	-1.01406879	-0.55153375
6	-0.3569140812	2.916333e-01	-1.012149e-01	1.068686e+00	0.1848462359	0.550753255	0.43928575	-0.26993029
7	3.5003523840	4.320879e+00	2.308831e+00	4.739111e-01	2.4242017567	1.599109847	1.26977229	1.98289739
8	-0.3569140812	-3.361412e-01	-4.519487e-01	-3.187292e-01	-0.5029675894	1.599109847	1.26977229	-0.10901403
9	0.0416007782	-3.852426e-02	-4.474667e-01	1.128714e+00	0.4419666246	-0.287930509	-0.39120327	-0.39061749
10	0.4151862435	2.554677e-01	-2.156726e-01	3.842665e+00	2.6874894061	-0.916943961	-0.80644529	-0.47107562
11	-0.3569140812	2.504878e-03	-2.553043e-01	1.106303e-01	-0.2901116566	1.599109847	1.26977229	-0.10901403
12	0.5420502310	4.028311e-01	-4.390315e-01	1.942134e+00	1.9305538712	-0.916943961	-0.59882428	-0.35038842
13	1.1493502718	1.610549e+00	-4.863907e-01	6.120225e-01	0.7206241308	1.389439535	1.26977229	0.45419289
14	-0.1039090466	8.547255e-01	2.323780e-02	1.697364e-01	0.5590130850	0.970096395	0.64690676	0.45419289
15	-0.3569140812	-6.413929e-01	-3.903713e-01	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
16	-0.3569140812	4.864709e-01	1.874412e-01	8.542355e-01	0.7230037609	0.341082943	0.02404124	-0.14924309
17	-0.3569140812	-6.413929e-01	-5.916066e-01	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
18	-0.3569140812	-2.781980e-01	-6.584384e-01	-1.502246e-02	-0.1469472144	-0.078260197	-0.18357977	-0.26993029
19	-0.2569073288	-2.884500e-01	5.483138e-02	-1.091009e-01	-0.3620703416	0.550753255	0.43928575	-0.22970122
20	-0.3569140812	-3.624256e-01	-2.695675e-01	-3.520544e-01	-0.5194886442	1.599109847	1.26977229	-0.10901403
21	-0.3569140812	-5.177527e-01	-6.737726e-01	-4.706092e-01	-0.5782624869	0.760426083	0.43928575	-0.26993029
22	3.2035913830	3.809285e+00	2.285566e-01	3.027952e-02	2.1026393948	1.599109847	1.26977229	3.10931123
23	-0.3569140812	-7.042803e-02	3.655781e-01	-1.947538e-01	-0.3359536250	1.599109847	1.26977229	0.09213130
24	1.7242866816	2.331592e+00	3.828885e+00	2.773662e+00	1.1738358990	1.389439535	1.26977229	-0.06878496
25	-0.3569140812	-6.413929e-01	-9.018741e-02	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
26	-0.3569140812	-3.617538e-01	-6.775487e-01	-3.512026e-01	-0.5190663808	1.599109847	1.26977229	-0.10901403
27	-0.2954641490	-5.700136e-01	-3.492504e-01	-1.627501e-01	-0.4256411461	-0.916943961	-0.80644529	-0.51130468
28	-0.3569140812	-4.781441e-01	-6.812129e-01	-4.987706e-01	-0.5922235116	1.599109847	1.26977229	-0.10901403
29	-0.2336828691	-3.705363e-01	-3.246558e-03	-1.906320e-01	-0.4394632633	0.131412631	0.43928575	-0.26993029
30	-0.3569140812	-5.714132e-01	-7.040794e-01	-4.928082e-01	-0.5892677524	0.131412631	-0.18357977	-0.39061749

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

The DB+ and DB- is assigned to input dataset values of K-Means. The K-means algorithm groups D-dimensional data vectors into a predefined number of clusters on the basis of the Euclidean distance as the similarity criteria. Euclidean distances among data vectors are minimum for data vectors within a cluster as compared with distances to other data vectors in different clusters. Vectors of the same cluster are associated with one centroid vector, which represents


RGui (64-bit) - [Data: inputdata_final]
File

	ONEOFF_PURCHASES	PURCHASES	PAYMENTS	no_purchase	Monthly_Avg_expns	PURCHASES_INSTALLMENTS_FREQUENCY	PURCHASES_FREQUENCY	PURCHASES_TRX
1	-0.3569140812	-5.746322e-01	-6.841179e-01	-1.978853e-01	-0.4430594757	-0.707273649	-0.80644529	-0.51130468
2	-0.3569140812	-6.413929e-01	1.213673e+00	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
3	0.1088824294	-1.003304e-01	-4.796761e-01	-1.975152e-02	-0.3547487440	-0.916943961	1.26977229	-0.10901403
4	0.5461589421	4.076037e-01	-7.822864e-01	5.208100e+00	4.5097203922	-0.916943961	-1.01406879	-0.55153375
5	-0.3472748761	-6.301962e-01	-4.523040e-01	-5.353964e-01	-0.6103804992	-0.916943961	-1.01406879	-0.55153375
6	-0.3569140812	2.916333e-01	-1.012149e-01	1.068686e+00	0.1848462359	0.550753255	0.43928575	-0.26993029
7	3.5003523840	4.320879e+00	2.308831e+00	4.739111e-01	2.4242017567	1.599109847	1.26977229	1.98289739
8	-0.3569140812	-3.361412e-01	-4.519487e-01	-3.187292e-01	-0.5029675894	1.599109847	1.26977229	-0.10901403
9	0.0416007782	-3.852426e-02	-4.474667e-01	1.128714e+00	0.4419666246	-0.287930509	-0.39120327	-0.39061749
10	0.4151862435	2.554677e-01	-2.156726e-01	3.842665e+00	2.6874894061	-0.916943961	-0.80644529	-0.47107562
11	-0.3569140812	2.504878e-03	-2.553043e-01	1.106303e-01	-0.2901116566	1.599109847	1.26977229	-0.10901403
12	0.5420502310	4.028311e-01	-4.390315e-01	1.942134e+00	1.9305538712	-0.916943961	-0.59882428	-0.35038842
13	1.1493502718	1.610549e+00	-4.863907e-01	6.120225e-01	0.7206241308	1.389439535	1.26977229	0.45419289
14	-0.1039090466	8.547255e-01	2.323780e-02	1.697364e-01	0.5590130850	0.970096395	0.64690676	0.45419289
15	-0.3569140812	-6.413929e-01	-3.903713e-01	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
16	-0.3569140812	4.864709e-01	1.874412e-01	8.542355e-01	0.7230037609	0.341082943	0.02404124	-0.14924309
17	-0.3569140812	-6.413929e-01	-5.916066e-01	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
18	-0.3569140812	-2.781980e-01	-6.584384e-01	-1.502246e-02	-0.1469472144	-0.078260197	-0.18357977	-0.26993029
19	-0.2569073288	-2.884500e-01	5.483138e-02	-1.091009e-01	-0.3620703416	0.550753255	0.43928575	-0.22970122
20	-0.3569140812	-3.624256e-01	-2.695675e-01	-3.520544e-01	-0.5194886442	1.599109847	1.26977229	-0.10901403
21	-0.3569140812	-5.177527e-01	-6.737726e-01	-4.706092e-01	-0.5782624869	0.760426083	0.43928575	-0.26993029
22	3.2035913830	3.809285e+00	2.285566e-01	3.027952e-02	2.1026393948	1.599109847	1.26977229	3.10931123
23	-0.3569140812	-7.042803e-02	3.655781e-01	-1.947538e-01	-0.3359536250	1.599109847	1.26977229	0.09213130
24	1.7242866816	2.331592e+00	3.828885e+00	2.773662e+00	1.1738358990	1.389439535	1.26977229	-0.06878496
25	-0.3569140812	-6.413929e-01	-9.018741e-02	-7.057490e-01	-0.6948335133	-0.916943961	-1.22168980	-0.59176281
26	-0.3569140812	-3.617538e-01	-6.775487e-01	-3.512026e-01	-0.5190663808	1.599109847	1.26977229	-0.10901403
27	-0.2954641490	-5.700136e-01	-3.492504e-01	-1.627501e-01	-0.4256411461	-0.916943961	-0.80644529	-0.51130468
28	-0.3569140812	-4.781441e-01	-6.812129e-01	-4.987706e-01	-0.5922235116	1.599109847	1.26977229	-0.10901403
29	-0.2336828691	-3.705363e-01	-3.246558e-03	-1.906320e-01	-0.4394632633	0.131412631	0.43928575	-0.26993029
30	-0.3569140812	-5.714132e-01	-7.040794e-01	-4.928082e-01	-0.5892677524	0.131412631	-0.18357977	-0.39061749



ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 8, August 2017

The result will be arrange the data with the use of Average Mean. This study provides a new method of analyzing bank databases. Beyond simply understanding customer value, the bank gains the opportunities to establish better customer relationship.

VI. CONCLUSION

This study proposes a new framework of customer behavior analysis using K-means Clustering algorithm for analyzing bank databases. For differentiation purposes, we grouped customers with shared customer behavior with RFM value. After briefly reviewing the customer profiles using the association rule inducer, the customers with a higher CP or RFM might be the target customer groups of precedence. The existing customers were divided into three profitable groups of customers according to their shared behavior and characteristics. Marketers then can infer the profiles of customers in each group and propose management strategies appropriate to the each group. This study provides a new method of analyzing bank databases. Beyond simply understanding customer value, the bank gains the opportunities to establish better customer relationships. In modern years, the management and processing of so called data streams has become a subject of dynamic research in numerous fields of computer science such as, e.g., distributed systems, database systems, and data mining. Lot of research work has been carried in this field to develop an efficient clustering algorithm for time series data streams. Time series data are frequently large and may contain outliers. Therefore, careful examination of the earlier proposed algorithms is necessary. In this paper we surveyed the current studies on time series clustering. These studies are structured into many categories depending upon whether they work directly with the innovative data. Most clustering algorithms are not capable to make a distinction between real and random patterns. In addition, this paper discusses about possible high dimensional problems with time series data. The application areas are summarized with a brief description of the data used. The uniqueness and drawbacks of past studies and some possible topics for further study are also discussed. The future work determines to develop an effective clustering algorithm for time series data streams.

REFERENCES

[1] Methods Mohammad Ali Farajian & Shahriar Mohammadi, "Mining the Banking Customer Behavior Using Clustering and Association Rules", December 2010,, Vol.. 21,, No.. 4
[2] Setiono, R., Thong, J.Y.L., Yap, C.S., "Symbolic Rule Extraction from Neural Networks-an Application to Identifying Organizations Adopting IT". Information and Management, 34(2), 1998, pp 91–101
[3] Madan Lal Bhasin, "Data Mining: A Competitive Tool in the Banking and Retail Industries", 2006, p.no.588-590
[4] T.W. Liao, "Mining of vector time series by clustering", Working paper, 2005.
[5] T.W. Liao, B. Bolt, J. Forester, E. Hailman, C. Hansen, R.C. Kaste, J. O'May, "Understanding and projecting the battle state", 23rd Army Science Conference, Orlando, FL, December 25,2002.
[6] T.-C. Fu, F.-L. Chung, V. Ng, R. Luk, "Pattern discovery from stock time series using self organizing maps", KDD 2001 Workshop on Temporal Data Mining, August 26–29, San Francisco, 2001, pp. 27–37.
[7] D. Jixue; , "Data Mining of Time Series Based on Wave Cluster," Information Technology and Applications, 2009. IFITA '09. International Forum on , vol.1, no., pp.697-699, 15-17 May 2009.
[8] H. Liu; Z. Ni; J. Li; , "Time Series Similar Pattern Matching Based on Empirical Mode Decomposition," Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on , vol.1, no., pp.644-648, 16-18 Oct. 2006.
[9] Nan-Chen Hsieh, Kuo-Chung Chu 2009 "Enhancing Consumer Behavior Analysis by Data Mining Techniques" in International journal of information management and sciences.
[10] K.V.Nagendra, C.Rajendra 2012 "Customer Behavior Analysis using CBA (Data Mining Approach)" in National Conference on Research trends in Computer Science and Technology.