Presents an approach based on clustering and developed RFM methods to assess the level of customer loyalty in banking system

Shamsyeh Naghi-Malekshahi¹, Ali Harounabadi²*, Seyed Javad Mir Abedini²

¹Department of Computer Engineering, Ilam Science and Research branch, Ilam, Iran ²Department of Computer Software, Central Tehran Branch, Islamic Azad University, Tehran, Iran

* Corresponding Author: Email: a.harounabadi@gmail.com (a.harounabadi)

Abstract

Recognition of loyal customers and understanding customer needs, effective in achieving excellence in customer service. Retain such customers. In the long run than to attract new customers, more beneficial. In this paper we present a new invoice called length of customer relationship with the bank will add to classic RFM model. And utilizes fuzzy clustering and neural network techniques try Obtain better results than the classic mode. Simulation results on a real example shows the performance of our approach is better.

Keywords: Data Mining, Clustering, RFM, ANN, Loyal Customers

1. Introduction

Knowledge Today one of the great challenges of customer-centric organizations Understands customers, differentiate between various customer groups and their rankclassification (Eppler 2001). Organizations and banks just do not attract new customers and to ensure that customer satisfaction are loyal (Alhaji et al. 2014). Therefore, finding ways and models for clustering customers Based on their loyalty, for banks is essential. One of the techniques of data mining (DM), is clustering the data (Jafarzadeh et al. 2015). Clustering method is used to describe the various clusters in the data set (Gheorghe and Florina 2015). In clustering, customers with similar characteristics and conditions are in a cluster and clusters are formed, there are similarities with the lowest. This leads to a better understanding of the needs of consumers and is involved in the recognition of loyal customers and causes them to offer more services.

2. Expression Problem

Today's business environment is very complex, and strong competition among organizations providers of products and services is undeniable. Customers are the key factor in strengthening the organizations, And orientation of all the goals, strategies and resources are centered on attracting and keeping customers. Recognition of loyal customers and understanding customer needs, effective in achieving excellence in customer service.

Today, policy and competitive strategies of the experts emphasized. Therefore, credit institutions (banks, etc.) cannot achieve the main objectives of their work, such as creating competitive advantage or benefit granted. Identify different groups of customers and determine their needs and desires can cause customer satisfaction for financial institutions and this leads to more loyal consumers. Customers that banks can draw their consent, they express their positive experiences for others and so the banks are tools for promotion and thus reduced the cost of attracting new customers. Accordingly, in this paper we will attempt to RFM classical approach to improve and Obtain more detailed information from bank customers, To do this We will add one Other index to approach RFM called length client relationship And we expect to obtain a more accurate result. Remainder of the paper is as follows: In Section 3, a brief description of the clustering approach, RFM and artificial neural network (ANN) is presented. Section 4 deals with related work in this area. Section 5 describes the proposed approach is fully And evaluation of the proposed approach and conclusions are described respectively in sections 6 and 7.

3. Research Methodology

3.1 Clustering

One of the techniques that will help you be measured customer Satisfaction and enhance the quality of services and products using clustering techniques (Bose and Chen 2015). Clustering useful tool for analyzing the needs and demands of our customers and satisfy them. Humans intuitively by observing natural phenomena to compare and classify them into different groups (Bose and Chen 2015). One of the problems is crisp clustering is Boundary sharp problem that to deal with it should be use fuzzy clustering (Rahmati Torkashvand et al. 2014). In fact, the human mind to deal with complex issues, begin to classify the observations using their features. Clustering in banking processes, the process by which customers distinct sub-divided that same needs and specifications. Clustering helps to split a large market into smaller parts, optimize resource consumption for main customers.

3.2. RFM Model

The most popular model is used to determine loyalty, the model is called RFM (Recency, Frequency, Monetary), with emphasis on the parameters. Recently transactions, recurring transactions (turnover) and deposit rates, to determine each client's loyalty (Hu and Yeh 2014). Since the RFM, in some cases, have not been careful to distinguish accurately between clients and customers who have different attitudes to the same consideration, in this paper, by introducing modifications in the calculations, the basic model RFM, improved.

3.3. ANN

We ANN is an idea for information processing That is inspired by the biological nervous system so that as the human brain focuses on information processing (Erkaymaz et al. 2012). In this paper, a neural network is used to reduce the response time to customer and to determine the appropriate cluster for customers. The output of the artificial neural network (ANN), which is a cluster, customer uses their parameters to be entered. Clusters were classified based on customer loyalty, In other words, customer loyalty, will be validated.

4. Related research

Clustering techniques using RFM approach is now considered the world as one of the best methods of segmenting customers. The RFM techniques combined with techniques DM and new models are manufactured with great precision. Methods of using RFM are diversity and variation. In this section, some of the research done in this area is mentioned.

Chen et al in a paper titled "Enhancing Consumer Behavior Analysis by Data Mining Techniques" a two stage framework of customers based on the behavior and predict future behavior offer based on SOM waterfall and ANN, customers have become homogeneous graphs. Then, based on the knowledge obtained well begin to and evaluate clustering (Chen et al. 2009).

Nagendra and Rajendra in "Customer Behavior Analysis Uses CBA (Data Mining Approach)" have to analyze. CBA approach to customer behavior using DM. their initially dealt with the definition of customer relationship management and customer behavior analysis and to classify customers have made use of Naive Bayes Classifier Algorithm (Nagendra and Rajendra 2012).

Rad et al presented a paper entitled "Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran" (Rad et al. 2010).

Cheng and Chen, in a paper entitled "Classifying the segmentation of customer value via RFM model and RS theory" have to classify customers using the RFM approach and the theory of RS (Cheng and Chen 2009).

Jim et al in a paper titled "A fast k-means clustering algorithm using cluster center displacement" in the National Taiwan University have proposed an algorithm that with the changes have KM algorithm have create a fast algorithm KM, there have done this by of using displacement of the center of the cluster and have shown that using move the cluster center can reduce the computational time (Jim et al. 2009).

Alok and Srivastava in the article entitled "The Antecedents of Customer Loyalty: An Empirical Investigation in Life Insurance Context" in LIC companies in India with life insurance companies be paid customer clustering (Alok and Srivastava 2013).

Farajian and mohammadi in an article entitled "Mining the Banking Customer Behavior Using Clustering and Association Rules Methods" have to define DM and customer relationship management, then the clustering of bank customers using the DM. They first customer clustering and association rules after it have been used to make new customers in specific categories (Farajian and mohammadi 2010).

5. The proposed approach

In this study, given that the RFM model used in banking instead of buying term will use the word transaction and converted to the distance between transactions and the number of repetitions of the transaction and the amount of money deposited. As well as to enhance the accuracy of the algorithm and differentiate between those that are deposited at different times the parameter "L" (length the presence of customer) will be added to this method. Note that in this paper is a case study for banks, we have considered the financial one year. So then, recently transactions measure is the last time the customer goes to the bank and has done transactions and is a number between 0 to 365. Repeated measure, measures the frequency of referrals in the period of one year and criteria for evaluation of client money the remaining. At the end of year. And the duration of interaction with the bank is the day the transaction has done minus the first day's transactions. This model is used to measure customer loyalty.

Overall, this study pursues the following objectives: 1. a fuzzy clustering method for classification of customer deposits. 2. The use of ANN alongside RFM approach in the banking system. 3. Identification of loyal customers in the banking system.

5.1. L (length customer relationship)

We approach is to increase the ability of classic RFM, consider an index as well as to identify the loyal customers. This variable represents the time interval between the first and last customer relationship with the bank in a certain interval. Length client relationship with organization represents the amount of time that a customer has started his association with the organization. Equation (1) above points shows:

$$Lenght(L) = T_{(Last Transaction)} - T_{(First Transaction)}$$
(1)

5.2. Algorithm

Figure 1, shown an overview of the proposed method for evaluating the bank's customer loyalty:



Fig.1 the number of association rules generated by all four methods

In this system, initially user data extraction, data should be obtained from the bank and then the process should begin early on input data that is included RFML. In this paper for normalizing the input data from relations 2 to 4 is:

$$R_{N} = \frac{R_{MAX} - R}{R_{MAX} - R_{MIN}}$$
⁽²⁾

$$F_N = \frac{F_{MAX} - F}{F_{MAX} - F_{MIN}} \tag{3}$$

$$M_{N} = \frac{M_{MAX} - M}{M_{MAX} - M_{MIN}}$$
⁽⁴⁾

That R_{MAX} and F_{MAX} and M_{MAX} the respectively meaning is the largest value of R and F and M, and the purpose from R_{MIN} and F_{MIN} and M_{MIN} the respectively the lowest value of R and F and M. Finally, the purpose from R, F, and M is the actual index. Then should be obtained from normalized data input, the harmonic mean, to do this we use the following equation 5:

$$Mean(H) = \frac{4}{((\frac{1}{R}) + (\frac{1}{F}) + (\frac{1}{M}) + (\frac{1}{L}))}$$
(5)

After preprocessing step and get the harmonic mean of the input data, Users to create profiles, exploring the desired do and then using clustering. Will produce the desired profile, in the proposed system, this is done using ANN. This means that using ANN turns out the new person out from this period that the clustering is done enter the bank; given ANN training has been previously belong to which cluster. In this study, we convert the data into five clusters and the value of each of the clusters examined as follows.

- **Loyal cluster:** Highest value and its members both in terms of repeat transactions and the amount of money deposited, and the duration of interaction with the bank and recently going to the bank are high value.
- Active cluster: These customers are loyal, but their loyalty level is lower from the above cluster.
- **Semi-active cluster:** Customers cluster recently have to go to the bank and their financial value Too high But it's time to interact with the bank is very low during the last few months have gone by the depository.
- **Middle cluster:** These customers are the average value for the bank and in terms of monetary value and repeat transactions recently and the duration of interaction with the bank are moderate.
- **Case Cluster:** (Customer loyalty lowest) customer are not good. And the value of the repetition and duration of interaction with the bank and recently are the lowest. Figure 2 show an example of clustering was performed.



Neural Network Training (nntrainto	ol)	
Neural Network		
	1	
hout and		Output
b C	b C	
Algorithms		
Training: Levenberg-Margu	ardt (trainim)	
Performance: Mean Squared Err	or (mse)	
Data Division: Random (dividera	ind)	
D		
Progress		
Epoch: 0	46 iterations	1000
Time:	0:00:02	
Performance: 20.0	0.00106	0.00
Gradient: 1.00	0.000111	1.00e-10
Mu: 0.00100	1.00e-11	1.00e+10
Validation Checks: 0	6	6
Plots		
Performance (plotperform)		
Training State (plottrainstate	•)	
Deservation (all the second	(alatraaraarian)	
(plotregression	n)	
Plot Intendel:	1 epoo	hs
herelendered and the second second	und and and and and a set of the	
Validation stop		
	Chan Training	Concel
	an scop training wi	Cancer

Fig. 3 learning neural network

Since the purpose of ANN, classification clusters of customers, the data that is considered as a cluster customer, in ANN as the target and the output of the system will be considered. According to the above description and the initial value of ANN that is given below, proposed ANN algorithm will be in the form of (3) and (4):



Fig. 4 learning data - test data - evaluation of performance

After creating ANN, should be using the training data, we train the network. The work done by writing ANN code in MATLAB environment. Finally, the network has the necessary training and should be evaluated by test data. Figure 5 illustrates the above statement.



Fig. 5 assess the performance of ANN With test data

6. Evaluation of the proposed approach

Precision and recall results of the proposed method is that from four variables (recently transactions, recurring transactions, the amount of money deposited and duration of customer interaction) is formed, we compared with three variables system (recently transaction, the transaction is repeated, the amount of money deposited). The results show that the proposed system has been improved. See Figure 6.



Fig. 6 compares the proposed approach and the classical approach

7. Conclusions

According to the results of experiments performed It was found that length of partnership index added to the other three indexes increase the accuracy of identifying loyal customers In the banking system. The use from ANN is capable of learning considering to its ability to learn able to propose an appropriate cluster for each customer.

References

Alhaji, A-A. and Tasmin, R. and Josu, T. (2014). Online Banking and Customer Service Delivery in Malaysia: Data Screening and Preliminary Findings, Procedia - Social and Behavioral Sciences, Volume 129, 15, Pages 562–570

Alok, K. and Srivastava, M. (2013). The Antecedents of Customer Loyalty: An Empirical Investigation in Life Insurance Context, Journal of Competitiveness, Vol. 5, Issue 2, pp. 139-163.

Bose, I. and Chen, X. (2015). Detecting the migration of mobile service customers using fuzzy clustering, Information & Management, Volume 52, Issue 2, Pages 227-238.

Chen-Hsieh, N. and Chung-Chu, K. (2009). Enhancing Consumer Behavior Analysis by Data Mining Techniques, International Journal of Information and Management Sciences 20, 39-53

Cheng, C-H. and Chen, Y-Sh. (2009). Classifying the segmentation of customer value via RFM model and RS theory, Expert Systems with Applications 36, 4176–4184.

Eppler, M.J. (2001). Making knowledge visible through intranet knowledge maps: concepts, elements, cases. In, IEEE, 9 pp.

Erkaymaz, O. Özer, M. Yumuşak, N. (2012). Performance Analysis of A Feed-Forward Artifical Neural Network With Small-World Topology, Procedia Technology, Volume 1, Pages 291-296.

Farajian, M-A. Mohammadi, Sh.(2010). Mining the Banking Customer Behavior Using Clustering and Association Rules Methods, Internatiionall Journall off Industriiall Engineeriing & Productiion Research, Volume 21, Number 4 pp.. 239--245

Gheorghe, G. Florina, S. (2015). An assessment of the renewable energy potential using a clustering based data mining method. Case study in Romania, Energy, Volume 81, 1, Pages 416–429.

Hu, Y-H. Yeh, T-W. (2014). Discovering valuable frequent patterns based on RFM analysis without customer identification information, Knowledge-Based Systems, Volume 61, Pages 76-88.

Jafarzadeh, H. Rahmati-Torkashvand, R. Asgari, Ch. Amiry, A. (2015). Provide a New Approach for Mining Fuzzy Association Rules using Apriori Algorithm, Indian Journal of Science and Technology, Vol 8(8), 707–714.

Jim, Z.C. Tsung-Jen, H. Yi-Ching, L. (2009). A fast k-means clustering algorithm using cluster center displacement, Pattern Recognition 42, 2551 -- 2556

Nagendra, K.V. Rajendra, C. (2012). Customer behaviour Analysis using CBA (Data Mining Approach), National Conference on Research Trends in Computer Science and Technology.

Rahmati-Torkashvand, R. Asgari, Ch. Jafarzadeh, H. (2014). Combining Apriori algorithm and C-Means clustering algorithm to improve mining association rules, Advances in Environmental Biology, 8 (22), pp. 226-234.

Rad, A. Naderi, B. Ghobadizadeh, M. (2010). Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran, Expert Systems with Applications, Vol. 34, No. 1, 2212.