Introducing A Hybrid Data Mining Model to Evaluate Customer Loyalty

Hossein Alizadeh
Department of Computer Engineering, Buinzahra Branch,
Islamic Azad University, Buinzahra, Iran
hossein7alizadeh@gmail.com

Behrouz Minaei-Bidgoli
School of Computer Engineering, Iran University of
Science & Technology, Tehran, Iran
fhm1390@yahoo.com

Abstract—**The main aim of this study was introducing a comprehensive model of bank customers' loyalty evaluation based on the assessment and comparison of different clustering methods' performance. This study also pursues the following specific objectives: a) using different clustering methods and comparing them for customer classification, b) finding the effective variables in determining the customer loyalty, and c) using different collective classification methods to increase the modeling accuracy and comparing the results with the basic methods. Since loyal customers generate more profit, this study aims at introducing a two-step model for classification of customers and their loyalty. For this purpose, various methods of clustering such as K-medoids, X-means and K-means were used, the last of which outperformed the other two through comparing with Davis-Bouldin index. Customers were clustered by using K-means and members of these four clusters were analyzed and labeled. Then, a predictive model was run based on demographic variables of customers using various classification methods such as DT (Decision Tree), ANN (Artificial Neural Networks), NB (Naive Bayes), KNN (K-Nearest Neighbors) and SVM (Support Vector Machine), as well as their bagging and boosting to predict the class of loyal customers. The results showed that the bagging-ANN was the most accurate method in predicting loyal customers. This two-stage model can be used in banks and financial institutions with similar data to identify the type of future customers.**

Keywords-Loyalty; data mining; clustering; classification; evaluation

I. INTRODUCTION

Customer Relationship Management (CRM) is one of the most important areas in which data mining techniques has been widely used [1]. CRM is a set of processes and systems which support business strategies in order to create beneficial relationships between current customers and future customers [1-3]. With the increasing importance of CRM in all domains of industry, researchers need a standardized framework to achieve satisfactory results by using effective data mining processes. There are many companies which could retain their customers and increase their potential profit by evaluating customer value (potential participation) and loyalty [5]. Database management systems, data mining techniques and classification methods have changed the entire process of

marketing [6]. Research conducted in this context noted unbalanced classes in customer dataset [7]. This means that customers who are loyal have formed a smaller proportion of total customers. In general, management of unbalanced classes is a key factor for success in direct marketing and loyal customer analysis [8].

Customer retention and customer development may prove more attractive and less expensive for many organizations compared to finding new customers [9, 10]. The main challenge in customer segmentation is the presence of different customer related variables and the high degree of heterogeneity in customer behavior [11]. Customer variables are generally divided into two categories: behavioral variables and characteristic variables [12]. Characteristic variables included demographic, geographic and psychological variables, while behavioral variables encompassed reaction and tendency of customers to products and brands [12]. A large number of studies in the field of customer value analysis and value based customer segmentation through data mining techniques have used RFM variables [13]. Using three parameters, R (Recency: How recently did the customer purchase?), F (frequency: How often do they purchase?) and M (monetary value: How much do they spend?), these studies have attempted to value and segment customers in terms of their loyalty [11, 14]. There are other influential variables for customer segmentation. It has been claimed that RFM model does not differentiate customers who have long-term relationships with the organization and those with short-term relationships. That is why the variable duration of the first purchase to the last purchase of the customer (L) is added to the model. The combination of demographic variables such as year of birth and gender, as well as behavioral variables of customers, may be used to analyze customer value and loyalty.

In [15] authors used customer clustering for categorization of different industries for exploring the consumption patterns of electronic services. After successful modeling, they were able to designate the clusters correctly and extract the relevant patterns. In [16], authors used the K-means, RFM variables and RFM weighted variables for clustering of the customers. Finally they employed the multi-objective genetic algorithm to determine the score of each group of customers that showed greater accuracy in comparison to other methods. However, there is no comprehensive study on the factors influencing the

customer loyalty in banking. On the other hand, most researches in this area have only used clustering methods rather than other data mining methods such as classification.

Hence, one of the main challenges in this study is to use group classification methods (ensemble learning) in order to deal with this problem in dataset. This study presents a model for customer value analysis based on customer segmentation, using data mining techniques such as clustering and classification, to find loyal customers. The main objective of this study is customer segmentation using different clustering methods such as K-means as well as their existing four classes of variables to identify the variables effective in customer segmentation in respect with loyalty. In the second step, this study is aimed at presenting a classification model based on mass classification methods to predict the loyalty of new customers.

II. MATERIALS AND METHODS

There are two standard methods of data mining projects [15]. The first method is cross-industry standard process for data mining (CRISP-DM) and the second method is SEMMA (sample, explore, modify, model and assess). The former was developed by a consortium of the companies SPSS, DaimlerChrysler and NCR (National Cash Register). The latter was also suggested by the Institute of SAS [17]. SEMMA is a five-step method consisting of sampling, data exploration, data modification, modeling and model assessment. In addition to these steps, CRISP-DM involves a two-step business understanding in the first phase and deployment in the last phase. CRISP-DM compared with SEMMA, is closer to knowledge discovery method in databases proposed by Fayyad et al. (1996) [18]. Moreover, CRISP-DM has been widely used in customer segmentation projects by using RFM [19, 11, and 7]. Thus, this project also uses CRISP-DM. Futher, the RapiMiner data mining software, the featured software of the year 2013 for data mining was used.

The CRISP-DM phases include:

1. Business understanding: understanding the aims, investigating the situation, understanding the data mining goals and creating the plan project;

2. Data understanding: collecting primary data, describing data, data exploration, data quality review;

3. Data preparation: data selection, data cleaning and replacement (if necessary), and ultimately producing new variable data integration;

4. Modeling: choosing modeling method, building the model prototype, building the final model, and initial evaluation of the model.

5. Evaluation: evaluation of results and process review.

6. Deployment: a specific plan for application of the model in real world is be prepared.

III. RESULTS

Initially the data is derived from the Tourism Bank of Iran's database. The stored transactional data contained information related to 2124 customers. Then the customer data is integrated; that is the primary dataset containing transactional customer variables is incorporated into a table. Thus, every customer should have a data record. Three variables, R (time of the last transaction), M (total monetary volume of transactions for both creditor and debtor) and F (the number of transactions) are extracted for each customer. In this new generated dataset, only one record exists per customer.

A. Descriptive Statistics

In order to explain data as well as basic knowledge extraction from data collection in this section, some univariate statistics such as mean, standard deviation and bivariate statistics such as correlation are used. Table I summarizes the univariate statistics pertaining to some of the variables in the dataset. According to this table, the average age of customers is about 39 years and SD=10 indicate that the customers are mostly young. Majority of customers have a bachelor's degree and most of them are male. Given that SD=125 for F (the number of transactions) ranging from 1 to 3150, obviously the customers' behavior is highly diverse.

TABLE I.   STATISTICS RELATED TO VARIABLES

Variable	Type	Mean and SD.	Min. /Max.	Iteration of discrete variables
Age	Continuous	avg=39.131 +/- 10.967	[4 ; 87]	
Education	Discrete	mode=-9999 (1470), least=19046 (2)		19043 (112), 19040 (284), 19052 (158), 19037 (40), 19055 (31), 19049 (27), 19046 (2)
Gender	Discrete	mode=1180 (1888), least=1181 (236)		1180 (1888), 1181 (236)
Marriage	Discrete	mode=1184 (1230), least=1185 (894)		1185 (894), 1184 (1230)
M (bedehkar)	Continuous	avg=7077957513.45 +/- 53016182057.36	[-619938.000 ; 1223605648286.000]	
M (bestankar)	Continuous	avg=5543003252.14 +/- 41457587768.67	[-389961.000 ; 935889627450.000]	
R	Continuous	avg=16307.65 +/- 83.24	[16011.000 ; 16372.000]	
F	Continuous	avg=33.18 +/- 125.25	[1.000 ; 3150.000]	

IV. CORRELATION ANALYSIS OF VARIABLES

Table II shows the results of correlation analysis between different variables. This correlation is presented by the well-known Pearson coefficient. The Pearson correlation ranges from -1 to +1; closer values to +1 indicate direct linear correlation between the two variables and closer values to -1 indicate indirect correlation between the two variables. The closer values to zero mean no correlation between the two variables.

According to this table, there is highly positive correlation or highly negative correlation between some variables. For example, it is clear that M creditor (M bedehkar) and M debtor (M bestankar) representing the total credit and debt have a highly positive correlation with F, as expected. Interestingly, there is a correlation between M debtor (M bedehkar) and marital status.

TABLE II. PEARSON CORRELATION TEST ON VARIABLES

	Age	Education	Gender	Marriage	M (bedehkar)	M (bestankar)	R	F
Age	1.000	0.186	0.001	-0.537	0.098	0.085	-0.022	0.108
Education	0.186	1.000	0.200	-0.340	0.034	0.046	-0.001	0.103
Gender	0.001	0.200	1.000	-0.123	-0.001	0.004	0.033	0.024
Marriage	-0.537	-0.340	-0.123	1.000	-0.130	-0.126	0.026	-0.181
M (bedehkar)	0.098	0.034	-0.001	-0.130	1.000	0.899	0.095	0.761
M (bestankar)	0.085	0.046	0.004	-0.126	0.899	1.000	0.096	0.704
R	-0.022	-0.001	0.033	0.026	0.095	0.096	1.000	0.171
F	0.108	0.103	0.024	-0.181	0.761	0.704	0.171	1.000

A. Extraction of New Variables

Using banking expert opinions, we decided to generate more variables from the current dataset. Since the customers deposit (credit) and withdraw (debit) their accounts, it was decided to calculate F and R twice per customer. That is, they are calculated once for debit accounts of each customer and once for credit accounts of the same customer. Given the significance of account balance, it is decided to subtract M creditor (M bestankar) of each customer including total deposits from M debtor (M bedehkar) including total withdrawal; the result is recorded as a variable. Thus, behavioral dataset includes the following variables:

• R Debtor (Rbedehkar): time span from the last withdrawal day relative to a certain date (01.01.1970) in days.

• R Creditor (Rbestankar): time span from the last deposit relative to a certain date (01.01.1970) in days.

• F Debtor (Fbedehkar): Number of withdrawals of a customer from his/her bank account.

• F Creditor (Fbestankar): Number of deposits of a customer to his/her bank account.

• M Debtor (Mbedehkar): Number of withdrawals of a customer from bank accounts in general.

• M Creditor (Mbestankar): Number of deposits of a customer to bank accounts in general.

• Mtotal: total amount of credit minus total amount of debit.

These variables are incorporated in the next clustering.

V. DISCRETIZATION OF VARIABLES

In this section, data is discretized. The reason for this is that the variables have different ranges. For example, F ranges from 1 to 3150, while M has a relatively broader range. Thus, M will have a greater effect on clustering if the dataset is incorporated unchanged in the clustering.

At this point, frequency-based discretization is used. In this way, each of Fbestankar, Rbestankar, Rbedehkar and Mtotal is divided into five ranges, provided that equal number of records is set in each range. Fbedehkar is divided into four ranges, because it includes many equal values. Accordingly, all variables are normalized prior to clustering.

A. Modeling

A two-stage modeling is performed in this study. In the first stage, the behavioral data including Fbestankar, Fbedehkar, Rbestankar, Rbedehkar and Mtotal are incorporated in clustering and customers are labeled in accordance with their value. Based on demographic and geographic variables, such as age, education, gender, marriage and place of birth (birth_city), the labels are modeled through classification, and the top model is identified and selected in each stage.

B. Clustering

K-means, c-means and k-medoids are used for clustering of variables generated previously which in fact represent customer behavior. In using k-means and k-medoids which are unable to identify the optimal number of clusters, clustering was performed in a range from 2 to 9 clusters and Davies-Bouldin index was calculated in each stage. Davies-Bouldin index (V_{DB}) which has many applications uses similarity measure of two clusters. This index calculates the average similarity of each cluster to the most similar cluster. The higher the index, the better clusters are generated. Many other indices have been designed in recent years to overcome shortcomings of previous indices. However we refrain from mentioning them here due to their lengthy calculations.

$$V_{DB} = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i$$

Four clusters are generated in k-means and seven clusters are generated in k-medoids. On the other hand, x-means is able to identify the number of clusters; thus, the dataset is divided into four clusters. Table III shows the value of Davies-Bouldin

index for optimal number of clusters for the three clustering methods.

TABLE III. RESULTS OF CLUSTERING

Method	Optimal number of clusters	Davis-Bouldin Index
K-means	4	0.92
X-means	4	0.94
K-medoids	7	1.208

Four clusters are selected as optimal clusters for grouping customers in both k-means and x-means, while seven clusters are selected as optimal clusters in K-medoids. However, the Davies-Bouldin index is equal to 0.92 in k-means and 0.94 in x-means for 4 clusters and 1.121 in K-medoids for 7 clusters. As a result, k-means is chosen as a better method for clustering.

In order to label the clusters, the table pertinent to the information of cluster centers is analyzed. The results are as follows: Given that creditor refers to the customer who deposits in his account, obviously, Rbestankar and Fbestankar are higher in cluster_2 than other clusters (mean=4.39 and 4.34, respectively). Moreover, customers of cluster_2 gain higher scores in terms of Mtotal than other clusters. Fbedehkar and Rbedehkar are relatively lower than members of other clusters. For this reason, the members of this cluster are labeled as loyal customers. Customers of the cluster_0 are the second highest following cluster_2 in terms of Rbestankar and Fbestankar, while Mtotal is equal to 1.18, on average, which is the lowest compared to members of other clusters. In cluster_0, Fbedehkar and Rbedehkar are lower than other clusters. Thus, customers can be labeled as active, but with little capital. Certainly, cluster_1 includes the customer class with weak loyalty, because Rbestankar and Fbestankar are lower than other clusters. Moreover, Mtotal is slightly higher than cluster_0. Fbedehkar and Rbedehkar are higher in the cluster_1. Fbedehkar and Rbedehkar are evidently higher in cluster_3. Mtotal is equal to 3.92, on average. Moreover, Rbestankar and Fbestankar are medium to other clusters. Thus, members of cluster_3 can be potentially labeled. In order to calculate the scores of customers in each cluster, the Mtotal, Fbestankar and Rbestankar values are summed up and subtracted from Fbedehkar and Rbedehkar variables' values. The following table shows scores of members of each cluster. For clarity, the obtained values are normalized.

TABLE IV. SCORES OF EACH CLUSTER

	cluster_0	cluster_1	cluster_2	cluster_3
Total score	5.47	-2.69	8.81	0.59
Normalized Score	0.71	0.00	1.00	0.28

Apparently, customers of cluster_1 belong to class of customers with weak loyalty (total score=-2.69 and normalized score=0) and customers of cluster_2 are very loyal customers (total score=8.81 and normalized score=1). Two other clusters include medium customers.

C.   Classification

In order to develop a predictive model at this stage, customers who are in loyal cluster are labeled as Good and other customers are labeled as Bad. Using demographic variables including age, education, marital status, gender and birth city, modelling is performed by classification methods. At this stage, the decision-tree (DT), K-Nearest Neighbors (KNN), Naive Bayes (NB), Support vector machine (SVM) and artificial neural networks (ANN) as well as their bagging and boosting are used. Thus, 15 classifications are made on the dataset. It should be noted that the methods employed in this section are those that have achieved better results in most studies compared with other methods. Such methods and the relevant parameters are summarized below:

- Naive Bayes: This method has no parameters to set. In Bagging and Boosting models of this method, the Bag number parameter for Bagging and the number of repetitions for Boosting was set to 10.

- K-Nearest Neighbors: In this method, there is only one adjustable parameter: K. This value was changed in the range from 5 to 15. The best parameter combination was obtained here with the value 13 for the parameter K. In Bagging and Boosting models of this method, the number of Bag parameter for Bagging and the number of repetitions for Boosting was set to 10.

- Artificial Neural Networks: In this method, using the hidden layers of algorithm, a non-linear input-output mapping is performed. The main objective of this method is finding the appropriate weights for the network so that all the initial training data are accurately classified or predicted. The parameters set for this method in RapidMiner software are the learning rate of 3.0, momentum of 2.0, the number of training cycles equal to 500; the number of hidden layers was set by the software itself. In Bagging and Boosting models of this method also the number of Bag parameter for Bagging and the number of repetitions for Boosting was to 10.

- Decision Tree: The C.5 algorithm is used for building a decision tree or set of rules. This algorithm works through sample analysis based on the feature that receives the maximum information. The parameters set in this model are Gini index of branching, the tree depth equal to 15 and the minimum size of branching equal to 30. At this stage, also the decision tree Bagging and Boosting methods were used. The number of bags parameter in Bagging method and the number of iterations in boosting method was set to 10.

- Support Vector Machine: Another intelligent model having acceptable performance in predicting the complex and nonlinear features is the Support Vector Machine (svm) for which the number of bags parameter in Bagging method and the number of iterations in boosting method was set to 10. Accuracy and F-measure are used to evaluate the methods used. In the diagram below, vertical axis shows accuracy and the horizontal axis shows F-measure for the models used.

Obviously, bagging-ANN outperforms other methods in terms of F-measure. It is followed by bagging-DT, boosting-DT and DT in terms of accuracy and F-measure. Moreover, ANN and boosting-ANN never recognize minority class in this dataset. KNN-based methods have also unsatisfactory performance. One of the objectives of this study is to determine effectiveness of mass classification. Therefore, the methods

with the higher accuracy, on average, are determined. Thus, the following table calculates the mean of accuracy and F-measure for all the methods used as well as the used bagging and boosting methods. According toTable V, bagging techniques have the lowest accuracy (mean=78.42%) and highest F-measure (mean=15.49%). Clearly, the boosting techniques have failed in increasing the accuracy of modeling. It should be noted that the F-measure value is very low due to the serious imbalances in the relevant category.

TABLE V.   MEAN OF ACCURACY AND F-MEASURE FOR THE MODELS USED

Method	Mean accuracy	F-measure
Basic methods	81.00	8.55
Bagging methods	78.42	15.49
Boosting methods	81.00	8.55

D. Classification by Changing the Balance of Classes

Considering the imbalance of the classes in the under study problem, and because about 17% of the customers were loyal customers which was much less than other customers with weak loyalty, we decided the next phase that the two upper clusters having more loyal customers than other clusters, to adopt a class of loyal customers (instead of just one cluster) and thus, about 36% of customers were considered as loyal customers, and modeling was performed on this more balanced dataset. Finally after modeling, the mean of different indices for the base models, the Bagging and Boosting models are listed Table VI.

TABLE VI. THE MEAN ACCURACY AND F-MEASURE INDECES IN ALL MODELS FOR MORE BALANCED DATASET

Method used	Accuracy	F-measure
ANN	82.57%	0
Bagging – ANN	71.9	21.83%
Boosting – ANN	82.57%	0

As can be seen, the best F-measure obtained (21.83%) is related to Bagging-ANN; Obviously, ANN and boosting-ANN are unable to predict loyal (Good) customers; despite their accuracy (82.5%), these models are excluded from further analyses. Bagging-ANN is selected as the best model (71.9% accuracy). As indicated in Table VII, the F-measure value has decreased considerably which is due to the intense imbalance in the related classes. To solve the serious imbalance problem in the classes it was decided that the two no 2 and no 0 clusters which had more loyal customers compared to other clusters be rated as Good class customers and the existing customers in the other two clusters are considered as belong to the Bad class or class of customers with weak loyalty.

Thus, about 36% of customers are loyal. This leads to better balance between loyal customers and customers with weak loyalty. Finally, the modelling is done by all the listed modeling techniques used previously. Table VIII presents the results of evaluation on this new dataset. Obviously, the F-measure of neural networks has considerably improved with change in the class balance compared to pre-classification stage; accordingly in this stage considering better balance in the classes compared with the previous stage, NB has the best

recognition in the modelling (70% accuracy). Moreover, bagging-ANN is the best method in terms of F-measure (78.97%).

TABLE VII.   CUSTOMER CLASSIFICATION USING ANN

The method used	Accuracy	F-measure
ANN	82.57%	0
Bagging – ANN	71.9	21.83%
Boosting – ANN	82.57%	0

TABLE VIII.   CLASSIFICATION OF CUSTOMERS USING NEW DATASET

Model	Accuracy	F-measure
DT	66.88%	75.66%
Bagging-DT	68.45%	76.66%
Boosting-DT	67.82%	76.36%
ANN	65.15%	78.90%
Bagging-ANN	65.31%	78.97%
Boosting ANN	65.15%	78.90%
KNN	64.36%	72.22%
Bagging-KNN	61.54%	68.63%
Boosting KNN	66.41%	74.82%
NB	70.96%	77.52%
Bagging NB	69.07%	76.24%
Boosting NB	70.17%	77.70%
SVM	61.15%	71.65%
Bagging SVM	64.42%	70.35%
Boosting NB	63.34%	73.35%

VI.   DISCUSSION & FUTURE WORK

Considering the study results, it is clear that the customers' available demographic variables were not sufficient enough due to the lack of some very influential variables of customer loyalty such as the income variable, etc. Also, because the Bagging and Boosting collective classification methods in this study achieved better accuracy than the basic methods in detecting the loyal customers, it can be stated that it would be preferable to use other collective classification methods of predicting the customer loyalty class in the future studies. Among the methods used in this study, the K-Nearest Neighbors in all three basic, Bagging and Boosting methods in the first dataset has achieved better accuracy, although given the nature of modeling in which the F-measure index constitutes a more important indicator, using the Bagging method in artificial neural networks would be more appropriate. On the other hand, given that in the second modeling attempt was made to label two clusters of customers as "loyal", the modeling results on this dataset indicates that the F-measure index in this state can be brought up to 78.97%, which is much better than the previous state. For the purpose of development of the results of this study the following propositions can be made:

- Using other computational methods such as fuzzy logic and other methods to enhance the accuracy of modeling.

- It must be noted that in the present study the data pertaining to one of the banks which were more suitable than the data of other banks were used, although it suffers

from many shortcomings in terms of the lack of demographical variables like the employment, place of residence, etc., which is hoped to get better results in the future through data restoration.

- Implementation of the proposed model of the study on other datasets and reporting the results.

- Using more variables for modeling on such datasets.

VII. CONCLUSION

A comprehensive model of bank customers' loyalty evaluation based on the assessment and comparison of different clustering methods' performance was proposed and the different approaches were evaluated with significant accuracy being achieved. Results are shown in detail and further discussed whereas future work perspectives and improvements are proposed. The two-stage model proposed can be used in banks and financial institutions with similar data to identify the type of future customers.

References

[1] E. W. T. Ngai, L. Xiu, D. C. K. Chau, "Application of data mining techniques in customer relationship management: a literature review and classification", Expert Systems with Applications, Vol. 36, No. 2, pp. 2592-2602, 2009

[2] R. Ling, D. C. Yen, "Customer relationship management: An analysis framework and implementation strategies", The Journal of Computer Information Systems, Vol. 41, No. 3, pp. 82-97, 2001

[3] D. K. Tiwary, "Application of data mining in customer relationship management (crm)", Advances in Computational Sciences and Technology, Vol. 3, No. 4, pp. 527-540, 2010

[4] S. Rosset, E. Neumann, U. Eick, N. Vatnik, "Customer lifetime value models for decision support", Data Mining and Knowledge Discovery. Vol. 7, No. 3, pp. 321-339, 2003

[5] Y. H. Liang, "Integration of data mining technologies to analyze customer value for the automotive maintenance industry", Expert Systems with Applications. Vol. 37, No. 12, pp. 7489-7496, 2010

[6] M. Re, G. Valentini, "Ensemble methods: a review" in Advances in Machine Learning and Data Mining for Astronomy, Chapman & Hall, 2012

[7] M. Alvandi, S. Fazli, F. S. Abdoli, "K-Mean Clustering Method for Analysis Customer Lifetime Value with LRFM Relationship Model in Banking Service", .International Research Journal of Applied and Basic Sciences, Vol. 3, No. 11, pp. 2294-2302, 2012

[8] S. H. Liao, Y. J. Chen, M. Y. Deng, "Mining customer knowledge for tourism new product development and customer relationship management", Expert Systems with Applications. Vol. 37, No. 6, pp. 4212-4223, 2010

[9] J. Pablo Maicas Lopez, Y. Polo Redondo, F. J. Sese Olivan, "The impact of customer relationship characteristics on customer switching behavior: Differences between switchers and stayers, "Managing Service Quality: An International Journal", Vol. 16, No. 6, pp. 556-574, 2006

[10] B. Fang, S. Ma, "Data mining technology and its Application In CRM of Commercial Banks", First International Workshop on Database Technology and Applications, pp. 243-246, 2009

[11] M. Khajvand, M. J. Tarokh, "Estimating customer future value of different customer segments based on adapted RFM model in retail banking context", Procedia Computer Science, Vol. 3, pp. 1327-1332, 2011

[12] P. Kolter, Marketing Management, 11th edition, Pearson Education, New Jerse,. 2003

[13] J. T. Wei, S. Y. Lin, C. C. Weng, H. H. Wu, "A case study of applying LRFM model in market segmentation of a children's dental clinic", Expert Systems with Applications, Vol. 39, No. 5, pp. 5529-5533, 2012

[14] C. H. Cheng, Y. S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory", Expert Systems with Applications, Vol. 36, No. 3, pp. 4176-4184, 2009

[15] H. Khodzi, E. Akhondzadeh-Noughabi, B. Minaei-Bidgoli, "A New Application of RFM Clustering for Guild Segmentation to Mine the Pattern of Using Banks' e-Payment Services", Journal of Global Marketing, Vol. 27, No. 3, pp. 178-190, 2014

[16] Z. Zalaghi, Y. A. Varzi, "Measuring customer loyalty using an extended RFM and clustering technique", Management Science Letters, Vol. 4, No. 5, pp. 905-912, 2014

[17] A. Azevedo, M. F. Santos, "KDD, SEMMA and CRISP-DM: a parallel overview", IADIS European Conference on Data Mining, Amsterdam, The Netherlands, July 24-26, 2008

[18] S. S. Rohanizadeh, M. B. Moghadam, "A proposed data mining methodology and its application to industrial procedures", Journal of Industrial Engineering, Vol. 4, pp. 37-50, 2009

[19] M. Bizhani, M. J. Tarokh, "Behavioral rules of bank's point-of-sale for segments description and scoring prediction", International Journal of Industrial Engineering Computations, Vol. 2, No. 2, pp. 337-350, 2011