

Identification of Customers in the CRM system using Data Mining and Fuzzy AHP Method

Bahareh Farhangian^a, Mahboubeh Shamsi^b, Reza Ahsan^c

^a *Master Student of Information Technology, Taali University of Qom, Iran.*

^b *The Professor Assistant of Department of Electrical and Computer Engineering, Qom University of Technology, Iran*

^c *The Professor Assistant of Qom Branch of Islamic Azad University, Iran*

Abstract

Today, the interaction between organizations and especially the banks and customers in the form of customer relationship management has considerably changed. One of the main challenges for customer-centric organizations in the CRM system is to identify customers, to differentiate between different groups and allocate resources to them value and profitability. As well as providing appropriate services to potential customers, with the increased competition between banks and financial institutions is very important. This is possible by identifying customers and analyzing their behavior. In this regard, this study has tried to identify different groups of Iranian banks by taking advantage of WRFMLP developed model. The transaction data of 19500 customers in the period 2014/03/21 to 2015/04/20 is used. The two-step data-mining algorithms was used to identify different groups of customers. A comparison conducted among three clustering algorithms showed the highest quality. The Fuzzy AHP method was used to recognize the customer's value in created groups and weighting input parameters. Finally, cluster analysis was performed using decision tree algorithms.

Keywords: CRM, Customers Identification, customer segmentation, data mining, Fuzzy AHP

Introduction:

Banks and financial institutions are dramatically expanding so that customers have many options in order to carry out their commercial transactions. Thus there is no long-term guarantee to continue business of customers in the bank. Also considering the fact that customers today are known as a key resource in achieving the purposes of banks, the importance of identifying the different groups of customers in order to better understand the customers and identify their banking needs is essential. So all customer-centric organizations that are leading, try to establish a strong CRM system.

CRM is a newly developed management mechanism which aims to increase communication between organizations with customers, and as a business strategy, which includes all processes to use modern techniques management and information. It is used to select, evaluate, retain and gain customers and other activities related to customers. This is of great importance for CRM in banks, because customers are considered as a survival and continuity of any financial and banking institution (Bhatnagar, 2012). Identification of different customer groups has an important role in the area of CRM in banks and financial institutions. In fact, the segmentation is the process of dividing customers into distinct, meaningful and homogeneous subgroups. The segmentation offers an overall view of your customer database and will help managers to have appropriate and policy and behavior in dealing with customers of every sector. (Tsiptsis & Chorianopoulos, 2009).

In the context of customer relationship management, data mining can be used as one of the axes of business to discover and use continuously profitable knowledge from the raw data. Data mining can be used to guide decision-making and predict the effects of decisions and also to increase response rates to marketing campaigns (Jayasree & Vijayalakshmi, 2013). Techniques and data mining algorithms are a great tool designed to analyze data and detect patterns and trends. It is as a guide for banks and decision-makers to be able to more accurately make decisions (Imran & Ahmed, 2013). The robust CRM is very important. Using data mining techniques, we can increase CRM productivity and to better understand customers and markets.

Given the importance of identification of different groups of customers in the banks CRM, in this study, we tried to divide the customers of one of Iranian banks into different groups using the REMLP developed model RFMLP which is done using data mining algorithms and to determine the value of customers in the created sections, fuzzy AHP method has been used.

Theoretical basis of research

CRM Elements

We can divide the CRM technology into three parts: operational, Collaborative, analytical.

Operational CRM: includes a part of applications, which are essential for CRM. The process related to customers is made automatic and includes marketing automation, sales and customer service automation. An important goal of CRM is to improve and optimize customer service and sales force performance. In

most cases, operational and analytical CRM components are inserted in an integrated package to save regularly customers data. (Rababah and et al, 2011, Mishra, 2009, Wahlberg et al, 2009, Torggler, 2008).

Collaborative CRM: includes control and integration of all communication channels between organizations and customers. On the communication channels of the customers, suppliers and business partners are focused and creates a multi-channel communication, including the facilitating customer service systems through a variety of interactive and communicative channels of communication such as telephone, SMS, email and so on. It is possible the customers to select a process which in their view is the easiest and best choice (Mishra, 2009, (Wahlberg et al, 2009, Torggler, 2008).

Analytical CRM: In the past, organizations were focusing on Collaborative and operational tools, but this trend is changed. Decision-makers have realized that analytical tools to advance strategies and tactical decisions are necessary, the decisions concerning the identification of customers, attracting, development, maintenance, marketing improvement and all activities related to provide services and products to customers. As a result, advanced information technology and large databases must be concluded. This is a part of the CRM primarily focused on the analyze of the collected and stored data in order to create a more meaningful and profitable interaction with customers. The focus is on the preparation, simulation, analyze and optimization of decision-making processes related to customers. To achieve these objectives, the data collected is interpreted and disseminated by processing tools. According to some researchers the most vital and important part of CRM is the analytical one. Analytical CRM solutions allow managers to communicate effectively with customers of an organization. An organization can understand customers behaviour only by analyzing the customers data. Analysis is the key word of analytical CRM (Mishra, 2009, Wahlberg et al, 2009, Torggler, 2008, Ranjan, 2009).

This part of CRM follows the following subjects (Ranjan, 2009).

- A definition of different segments of customers and their key differences
- Prediction of the characteristics of a particular segment
- Identification of customers transfer from a segment to another segments
- Identification of the most profitable segment of customers
- Identification of the segments which are difficult to retain long-term relationship
- Identification of the segments which respond to marketing messages
- Prediction of future needs And expectations of customers

Analytical CRM Dimensions

This research is done on identification of different groups of customers as the first dimension in analytical CRM and our goal is to improve CRM using data mining techniques. A description of the analytical CRM were in the previous section.

According to (Swift, 2001, Ngai et al, 2009, Kracklauer et al, 2004) references, the analytical CRM is in four dimensions: 1) customers identification, 2) customers attraction, 3) customers maintenance, 4)

customers development. These four dimensions can be seen as a cycle of customer management systems. Figure 1 shows the sequence of analytical CRM.



Figure 1: Analytical CRM stages (Kracklauer and et al, 2004)

In table 1, the analytical CRM dimensions and the elements related to each dimension are shown.

Table 1: Analytical CRM Dimensions

Dimensions	Elements
Customer Identification	Customer Segmentation, Target Customer Analysis
Customer Attraction	Direct Marketing
Customer Maintenance	Loyalty Program, One-to-One Marketing
Customer Development	Customer Lifetime Value, Up/Cross Selling, Market Basket Analysis

Data Mining

The instrumental data mining to analyze customer data within the analytical CRM analysis. Many organizations have a great capital of the collected and stored data on existing customers, potential customers, suppliers and business partners. However, the inability to discover valuable information hidden in the data is an obstacle for organizations to transform the data into valuable information and useful knowledge (Ngai and et al, 2009). The data mining is new and powerful technology that helps to the managers of an organization by focusing on the most important information in the data warehouse of any organization. Data mining tools predict future trends and behaviors, and allows managers to make decisions based on that knowledge. Data mining tools can be answers and questions that are very time consuming to be solves with traditional methods (Nirkhi, 2010).

Data Mining Methods

Data mining methods are divided in two categories of supervised learning and unsupervised learning (Chitra et al, 2013). Supervised learning is considered as guided learning and also predictive. Learning process is directed by a dependent goal or feature that has already been recognized, directed. In this goal, the relationship between input traits that is named as independent variable and one or some target traits named as dependent variable is explored and this relationship is shown with a structure as presentation model. Driven data mining tries to explain and study the target behavior as a function of the independent or predictor explain traits (Chitra et al, 2013). Its main method includes classification, regression, forecasts and estimation.

The most important decision tree used to predict include: C5.0, QUEST, C&R and CHILD.

Unsupervised learning is known as undirected learning. There is no distinction between dependent and independent features. In fact, there is no preset output, which can lead the algorithm to create a model (Chitra et al, 2013). It aims to find patterns among existing data without having the target variable or without having preset set. The main methods include clustering, discovering dependence rule and discover sequential pattern.

The most important clustering algorithms used include Two-Step, K-Means and SOM.

Fuzzy AHP Method

When some indices are considered in the evaluations, the evaluation becomes complex and when multiple criteria are together in space and of different genders, the complexity will be more.

The evaluation and comparison are exited from an analytical simple mode and a strong practical analysis tools will be required, therefore, in recent decades, a variety of multiple criteria decision making methods have been presented and developed.

The hierarchical analysis process and fuzzy hierarchical analysis process are of MCDM methods. In AHP, although experts enjoy their competencies and intellectual abilities for comparisons in determining the most important parameters, but it should be noted that AHP is not able to reflect perfectly human thinking (Kahreman et al, 2004). AHP method is a hybrid method that combines fuzzy set theory and AHP. The main advantage of this method is to remove the ambiguity and imprecision of the human mind that we see this ambiguity and imprecision in the AHP method (Habibi, 2014). FAHP method has the following steps (Habibi, 1393 and Ishizaky et al, 2012):

The First Stage) the Hierarchical Chart

The Second Stage) The definition of fuzzy numbers in order to make paired comparisons: triangular fuzzy number is a fuzzy number with three real number that is shown as $F = (l, m, u)$.

The Third Stage) The formation of matrix of paired comparisons: at first, each pair-wise option is compared with other paired options, and then we form the matrix, so that the rows and columns of formed square matrix are the same decision criteria.

The Fourth Stage) Si calculation for each of the rows of the pair-wise comparison matrix: in fact, calculation of the sum total of all preferences of each element multiplied by the inverse of all preferences in which Formula 1 is used.

$$S_i = \sum_{gi}^m (\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j)^{-1} \quad (1)$$

$$\sum_{j=1}^m M_{gi}^j = (\sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m u_j) \quad (1.1)$$

$$(\sum_{i=1}^m \sum_{j=1}^m M_{gi}^j) = (\frac{1}{\sum_{i=1}^m u_i}, \frac{1}{\sum_{i=1}^m m_i}, \frac{1}{\sum_{i=1}^m l_i}) \quad (1.2)$$

The Fifth Stage) Calculation of Si magnitude than the other: for every one of triangular fuzzy numbers obtained in the previous stage, the comparison of other triangular numbers obtained will be done to determine the feasibility of each element. If two M_1 and M_2 fuzzy numbers are shown as $M_1 = (l_1, m_1, u_1)$ and $M_2 = (l_2, m_2, u_2)$, the comparison of these two numbers are calculated according to a formula 2.

$$V(M_2 \geq M_1) = \begin{cases} 1 & \text{if } m_2 \geq m_1 \\ 0 & \text{if } l_1 \geq u_2 \\ \frac{l_1 - u_2}{(m_2 - u_1) - (m_1 - l_1)} & \text{otherwise} \end{cases} \quad (2)$$

The Sixth Stage) The calculation of criteria weight in pair-wise comparisons matrix: for each of the triangular fuzzy numbers M_i where $i = 1, 2 \dots k$, the comparison to determine the weight of each element is performed according to formula 3.

$$V(M \geq M_1, \dots, M_K) = V[(M \geq M_1), \dots, (M \geq M_K)] = \min V(M \geq M_i), i=1, \dots, k \quad (3)$$

If $d(A_i)$ is the minimum grade of all comparisons M_i element is another element with k , that is:

$$(d(A_i) = \min V(S_i \geq S_k)) \quad (4)$$

Then the weight vector of n elements is as follows. However, gained weight vector is not normal we should normalize it in normal linear approach. In the linear normalization method, each element is divided by the set of all elements. In Formula 5 is given.

$$W = (d(A_1), \dots, d(A_n))$$

$$\text{normal} = w \left(\frac{d(A_1)}{d(A_1) + \dots + d(A_n)}, \dots, \frac{d(A_n)}{d(A_1) + \dots + d(A_n)} \right) \quad (5)$$

By way of interviews with banks experts, we can calculate the weight of each research criteria i.e. RFMLP and determine the importance of the different sections of segmentation.

Research methodology

There are several techniques for data mining projects, but one of the powerful methods is CRISP (Cross-Industry Standard Process for Data Mining). This is an overview of data mining cycle, which is done using a six steps work. These steps include: understanding the business issue, understanding the data, data preparation, modelling, evaluation of results and application of the model.

CRISP methodology is flexible and it can be used to comply with the research and we can bring it to a dedicated mode. According to the CRISP methodology, the present research to identify the different segments of customers and determine the values of the different sectors in banking related to deposit accounts will be promoted. The research framework is shown in Figure 2. The proposed framework has four main stages, each of which includes a number of activities. This framework is based on standard CRISP method and each of the steps and activities of the framework are done according to CRISP stages.

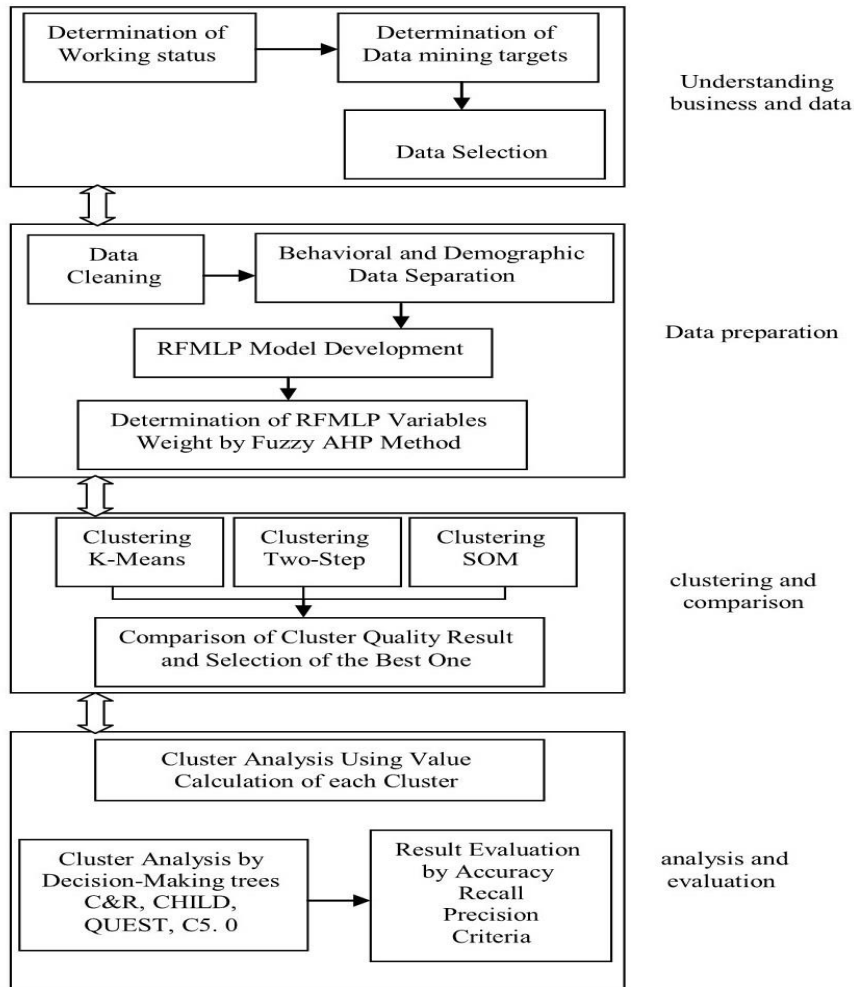


Figure 2: The Research Framework

The First Phase: Understanding Business and Data

In this phase, at first, we understood the goals and importance of research in terms of business which in the CRISP methodology, understanding business is the first step. This research is for customers of one of the Iranian banks. One the most important goals of any financial and banking institution is to retain and attract new customers, as well as provide the best customer service than competitors. In order to achieve these objectives we should have a correct understanding of the different customers and according to their characteristics, plan and present the services and products. The remarkable thing is that at any time the banks are required to identify and customers' needs to provide the different services for customers. With the results, we can identify customers with different levels of loyalty and profitability.

In order to achieve the stated objectives, this phase is continued and we will select the customer data and collection, which is as the second step in the CRISP methodology. In this study, to extract raw data the database in bank that are related to holders of bank deposit accounts at different investment was applied. This was done in accordance with existing standards. The input data related to transaction data and the accounts held by 19500 customers is from 2014/03/21 to 2015/04/20. Selected data are presented in Table 2.

Table 2: The selected data

Type of data	Selected fields	Notes
Transaction data (behavioral)	Recency (R)	The last transaction carried out in the selected time interval
	Frequency (F)	Total financial transactions in the time interval
	Monetary value (M)	The financial average of customer account in the time interval
	Relationship Length (L)	The years of customer account opening in bank up to 2015
	Customer Power (P)	The balance maximum of customer account during the thirteen month
Demographic data	Account holding	Savings interest-free
		Current interest-free
		Short-term investment deposit
		Long-term investment deposit
		Special investment deposit
	Sex	Female and male
	Province	31 province of Iran (the branch of the province in which the customer holds an account)

The selected data is derived from RFM model to assure the input data correctness that has contributed to the output. One of the most powerful models to implement CRM is RFM model that is based on recency, frequency and monetary value. According to Wei et al, 2010, RFM is a model based on behavior that is used to analyze customers behavior and then we can do their behaviors in the database. RFM model

expresses the present net present value of customers and is as customers evaluation system (Yanyan, 2011).

In this study, we tried to use other parameters other than R, F and M parameters and develop RFMLP model according to the data provided by the banks. L can be a measure of customer loyalty. Some of the researchers, according to the studies done in this research, expressed that RFM model is not effective to develop some CRM strategies like Cross-selling. Therefore, we added the P parameter to RFM model, and according to customer power from the maximum account balance we can perform Up/Cross selling.

The Second Phase: Data Preparation

Preparation is the most time-consuming phase of data mining. In this phase the outlier, noisy and incomplete data is considered from the set of identified data and appropriate strategy for them. Also the data are converted into a format that is suitable for the algorithms. The number of incomplete data against the data available were very low, so the best strategy for sorting and preparing them was to delete the incomplete data and finally it is obtained after the sorting 19239 records of customer data.

For weighting the input parameters i.e. RFMLP, five banking experts were interviewed. In the interview, they were given a questionnaire and asked to assess these variables into pair wise comparisons. FAHP method stages was described in the research literature. After interviews and a pair wise comparison of the parameters, table 3 from geometric mean of response parameters was obtained.

Table 3: The matrix of the sum of experts' response

	L	R	F	M	P
L	(1,1,3)	(2.713,4.192,5.347)	(1.482,2.0701,2.993)	(2.368,3.3797, 4.418)	(1.272,1.789,2.7086)
R	(0.1867,0.2384,0.5822)	(1,1,3)	(0.2926,0.4029,0.6047)	(0.1286,0.1657,0.2507)	(0.3581,0.4781,1.082)
F	(0.333,0.482, 0.6791)	(1.643,2.479,3.401)	(1,1,3)	(0.319,0.3981,0.8670)	(0.5428,0.7247,1.584)
M	(0.2254,0.2951,0.4203)	(3.958,6.034,7.740)	(1.144,2.511,3.116)	(1,1,3)	(1.389,2.120,3.253)
M	(0.3689,0.555,0.8753)	(0.918,2.091,2.786)	(0.641,1.379,1.838)	(0.3071,0.4698,0.7188)	(1,1,3)

Si calculation for each pair-wise comparison matrix row: all Sis were calculated according to the formula 1.

$$S_1 = (8.835, 12.430, 18.466) * (0.0166, 0.0268, 0.391) = (0.1467, 0.3331, 0.722)$$

$$S_2 = (1.966, 2.285, 5.519) * (0.0166, 0.0268, 0.391) = (0.0326, 0.0612, 0.2158)$$

$$S_3 = (3.838, 5.083, 9.487) * (0.0166, 0.0268, 0.391) = (0.0637, 0.1362, 0.3709)$$

$$S_4 = (7.716, 11.96, 17.5293) * (0.0166, 0.0268, 0.391) = (0.1281, 0.3205, 0.6854)$$

$$S_5 = (3.235, 5.4953, 9.1281) * (0.0166, 0.0268, 0.391) = (0.0537, 0.1473, 0.3569)$$

The calculation of S_i magnitude than each other and calculation of criteria weights are done according to formula 3 and 4 are as follows:

$$d^-(A1) = \min(S_1 \geq S_2, S_3, S_4, S_5) = (1, 1, 1, 1) = 1$$

$$d^-(A2) = \min(S_2 \geq S_1, S_3, S_4, S_5) = (0.1201, 0.4951, 0.1573, 0.2908) = 0.1201$$

$$d^-(A3) = \min(S_3 \geq S_1, S_2, S_4, S_5) = (0.3897, 1, 0.4357, 1) = 0.3897$$

$$d^-(A4) = \min(S_4 \geq S_1, S_2, S_3, S_5) = (0.9364, 1, 1, 1) = 0.9364$$

$$d^-(A5) = \min(S_5 \geq S_1, S_2, S_3, S_4) = (0.3654, 1, 1, 0.4106) = 0.3654$$

Finally, the weights (0.3654&0.9364&0.3897&0.1201&1) are obtained and these weights are not normal which is done by normalization method according to Formula 5. The final weight of parameters is given in table 4.

Table 4: final weight of RFMLP parameters Using FAHP

Parameters	Weights
Recency	0.0427
Frequency	0.1386
Monetary value	0.333
Relationship Length	0.3557
Customer Power	0.13

The Third Phase: Clustering and Comparison

Three widely used and clustering algorithms that is Two-Step, K-Means, and SOM are considered with different clusters for clustering. We will choose the algorithm with the highest quality. Table 5 shows the results of different algorithms clustering with the highest silhouette among 48 models.

Table 5: The best results of clustering

Model	Time (in min.)	Silhouette rate	Total clusters
Two-Step	<1	0.883	4

Two-Step	<1	0.868	5
K-means	<1	0.85	4
K-means	<1	0.832	6
K-means	<1	0.831	5
K-means	<1	0.812	9
Two-Step	<1	0.792	7
SOM	<1	0.504	12

Among the existing algorithms, Two-Step is considered as a final result, creating four clusters because of the highest quality. The final result of clustering is shown in Table 6.

Table 6: Result of Two-Step clustering

Cluster	% of Customer	Average R	Average F	Average M	Average L	Average P
4	54.1%	25	79	6652489	4	12854861
3	5.9%	214	87	2450565	4	9466265
2	38.1%	13	89	87373234	11	20910235
1	1.7%	19	345	271000000000	6	471000000000

The Fourth Phase: Analysis and Evaluation

FAHP method of customers value in each of the clusters is determined using the weights obtained to identify the most valuable customers group. Clusters value is calculated as follows:

$$V(C_i) = W_R * R(C_i) + W_F * F(C_i) + W_M * M(C_i) + W_L * L(C_i) + W_P * P(C_i) \quad (7)$$

$R(C_i)$, $F(C_i)$, $M(C_i)$, $L(C_i)$ and $P(C_i)$ show the normal value of date of the last transaction, the number of transactions, financial value, relationship length and i^{th} cluster power, respectively. W_R , W_F , W_M , W_L and W_P are the weights obtained from FAHP. Table 7 shows the value and rank of each of the clusters.

Table 7: and rank of clusters

Cluster	% of Customer	Value R	Value F	Value M	Value L	Value P	Value of each cluster	Segmentation Ranking
C ₂	38.1%	0.0411	0.0036	0.00069	0.1184	0.0034	0.1672	1
C ₁	1.7%	0.0396	0.0154	0.0189	0.0593	0.0077	0.1409	2
C ₄	54.4%	0.0396	0.0031	0.00056	0.0356	0.0002	0.0791	3
C ₃	5.9%	0.0152	0.0035	0.0002	0.0356	0.00015	0.0546	4

As a result of clustering and valuing of input parameters, the value of each cluster is calculated that the highest value is related to the third sector that covers 38% of customers.

To analyze the clusters obtained with different values and knowledge, decision tree algorithm C5.0, QUEST, CHILD and C&R was used to predict. Results of Two-Step clustering were considered as target variable and demographic data (Account Holding, Sex, Iran Province) in a table as a predictor field.

To run any four decision tree algorithms, data was divided into two categories of training and testing which 70% as training data and 30% as testing data were considered. Finally, to compare and evaluate algorithms, the three following criteria were used.

$$A = \frac{TP + TN}{TP + FP + TN + FN} \quad \text{Accuracy} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad \text{Recall} \quad (9)$$

$$P = \frac{TP}{TP + FP} \quad \text{Precision} \quad (10)$$

TP: the number of positive examples correctly predicted by the classification model.

TN: the number of negative examples correctly predicted by the classification model.

FN: the number of positive examples wrongly predicted as negative by the classification model.

FP: the number of negative examples wrongly predicted as positive number of negative examples wrongly predicted as positive by the classification model.

The results of comparison and evaluation from the decision tree algorithms are shown by table 8.

Table 8: Comparison among decision tree algorithms

Decision Tree Criteria	C5.0	QUEST	CHILD	C&R
Accuracy %	79.85%	77.94%	75.84%	1
Precision	0.766	0.747	0.723	0.723
Recall	0.722	0.769	0.746	0.746

According to the comparison of the table, the results from C5.0 algorithm will be considered. As we said, the second cluster with 38.1% had the highest value of the customers and is considered as the target customer and we should analyze them. The rules resulted from analysis with C5.0 for this cluster are as follows:

The rules related to the second cluster:

Rule 1) If account = special investment deposit and the province = Sistan and Baluchestan, Ghazvin, Alborz, Mazandaran, Fars, Qom, Tehran and Khuzestan → the second cluster

Rule 2) If account = current and province = Alborz, Hormozgan, Bushehr, Hamedan and Tehran → the second cluster

Rule 3) If account = interest-free → the second cluster

Rule 4) If short-term account and province = Tehran, Sistan and Baluchestan, Alborz and Qom → the second cluster

In determining the rules, the gender is not determining variable.

According to the studies carried out and the analyze obtained from the customers value, the following results are deduced for each cluster.

Segment 1: The customers of the part have two ranks but it is the most sparsely populated part and only 1.7% of customers are included. The customers of the part have the highest financial value and power and they have significantly higher financial power. The relationship length parameter does not give customers the point one. Their relationship length with the bank is moderate. Banks must take care of them strongly not to make the smallest attrition in this group because the loss of a customer in this part may be equal to the loss of profit from fifty to one hundred customers of other groups. Due to their high potential of this group it should not be limited only to micro banking services and private banking services must be provided to them.

Bank should implement loyalty programs for the customers to retain their relationship with the bank and to rank higher in the future and to become powerful and loyal customers.

Due to the high financial power, the Cross/Selling strategies can be implemented on them and interest-free savings accounts can be offered to them.

Segment 2: This group comprises 38% of total customers. The most valuable customers belong to this group. There is a very strong relationship with the bank, which is indicative of their satisfaction of banking services. Because they have the financial value less than the cluster 1 and because of high loyalty, they are in the first place, because banks are focused to retain customers than to attract customers. If they try to increase their account balances, the value and profitability of the group will be significantly increased. For long-term profitability, long-term accounts can be offered to them. According to the rules extracted, most customers of the group are from Tehran and Alborz provinces that hold more accounts. The maintenance strategy must be considered for the group.

Since the their relationship length and the last transaction date value are higher than other parts, the customers of this group are very loyal customers and try to only use the services of these banks compared to other banks that is important for a bank because customers belong to this group are the first people to respond to marketing activities and they do not act negatively against the changes. Also they are resistant against other competitors advertisement.

With regard to the higher value of the bank group, we should consider promotion activities. In recognition of their loyalty, the internal special plans such as short-term and long-term accounts are designed. Although this type of activities can have the costs, but they lead to advertise by customers and other than a retention strategy, the costs of attracting new customers will be reduced, so the result (win-win) will bring to customers and banks.

Segment 3: The customers of this group have the lowest value for the bank, and they have four ranks in terms of value. In terms of financial value and power they are the lowest and the relationship length with the bank is very short. Recently, they have joined the bank. The average date of the last transaction carried out with the bank is 214 days before, which is indicative of loss of customer. However, given the number of transactions done in the early years, the customers were not transient customers and they intended to have continued cooperation. Because of dissatisfaction with banks, they left the bank.

They must be prevented from turning away because although they are not profitable for banks but such a disconnect with the bank can prevent to attract new customers and during the project, the causes of dissatisfaction will be extracted.

Segment 4: is the largest part. 57% of customers belong to this part. Their value has three ranks. Their relationship length with bank is not so much and recently they have joined banks. We should try to attract customers and plan the programs for long-term and valuable relationship to promote higher levels and ranks. The appropriate strategies must be implemented in order to attract and retain them and prevent them from joining to other rivals because they are likely to respond to advertising competitors, and due to the high volume, the studies should be done on them to customers become stable.

Conclusion

To identify the customers as the first step in the analytical CRM, two segmentation elements and the target customers analysis can be applied. In this research, we could execute the two identification dimensions of customers and know appropriately all customers in one of Iranian banks, using data mining method to segment the customers of bank and fuzzy AHP for analyzing the groups of customers. We developed the basic RFM model to segment the customers and RFMLP model was used. The complete segments indicating the customers' behavior were presented. The fuzzy AHP method was used to determine the weight of parameters and value of the groups. In this model, the L parameter had the highest value according to banking experts. Therefore, the parameter is used in the segmentations based on RFM model. The P parameter that indicates the power of the customer is used to propose the strategies of Cross/Selling.

The three most commonly used algorithms of K-Means, Two-Step and SOM with different clusters were used for segmentation and finally the Two-Step output with four clusters was accepted because it had the highest quality among other segments. Using the weights assigned to input parameters, output clusters were ranked and their value was determined. The analysis showed that there is in cluster 1, the powerful customers, cluster 2, the loyal customers, cluster 3, the turn away customers and in cluster 4, the new customers and specific strategies for each cluster should be considered. The second cluster of the loyal customers had the highest rank and value as the target customers. The decision tree algorithms include QUEST, C&R, CHILD and C5.0 were compared to predict and discover the knowledge of the relationships between cluster and demographic information and because the results from C5.0 had the higher accuracy, its rules were accepted. The following suggestions should be considered in the future research.

- Other clustering algorithms to segment customers and other classification algorithms or social rules should be applied to analyze the target customers and the results to be compared.
- We can apply RFMLP model in other fields of electronic banking.
- Using other demographic information and other information of customers behavior and model development.

Reference

- Bhatnagar, S. (2012), Customer Relationship Management in Banking". (Bancon) Innovation to unlock the next decade. 199-206
- Chitra, K., Subashini, B. (2013), Data Mining Techniques and its Applications in Banking Sector, (IJETA) International Journal of Emerging Technology and Advanced Engineering, Volume 3, 2250-2459
- Feng, D., Zhang, Z., Zhou, F., Jianheng, J. (2008), Application Study of Data Mining on Customer Relationship Management in E-commerce, IEEE
- Imran, K., & Ahmed, Q. (2013), Use of Data Mining in Banking, (IJERA) International Journal of Engineering Research and Applications. Vol. 2, Issue 2 pp.738-742
- Ishizaka, Alessio and Hoang, Nam Nguyen. (2012), Calibrated Fuzzy AHP for current bank account selection, Expert Systems with Applications
- Jayasree, V., & Vijayalakshmi, R. (2013), A Review on Data Mining in Banking Sector, American Journal of Applied Sciences 10 (10): 1160-1165
- Habibi, Arash, Eizadyar, Sedigheh, Sarafrazi, Azam. (2014). The fuzzy multiple criteria decision making. Katibeh Gil publications
- Kahraman, C., Cebeci, U., Ruan, D. (2004). Multi-attribute comparison of catering service companies using fuzzy AHP: the case of Turkey. Int J Prod Econ
- Kracklauer, A. H., Mills, D. Q., Seifert, D. (2004), Customer Management as the Origin of Collaborative Customer Relationship Management, Collaborative Customer Relationship Management - taking CRM to the next level, 3-6.
- Mishra, A., & Mishra, D. (2009), Customer Relationship Management: Implementation Process Perspective, Acta Polytechnica Hungarica, Vol. 6, No. 4
- Ngai, E.W.T., Li, Xiu., Chau, D.C.K. (2009), Application of data mining techniques in customer relationship management: A literature review and classification, An International Expert Systems with Applications 36 (2009) 2592-2602
- Nirkhi, M.S. (2010). Potential use of Artificial Neural Network in Data Mining, IEEE

- Rababah, K., Mohd, H., & Ibrahim, H. (2011). "Customer Relationship Management (CRM) Processes from Theory to Practice: The Pre-implementation Plan of CRM System". *International Journal of e-Education, e-Business, e-Management and e-Learning*, Vol. 1, No. 1
- Ranjan, J. (2009), *Role of Analytical CRM in CRM Systems: Importance and Benefites*, Management & Change, Volume 13, Number 1
- Swift, R. S. (2001), *Accelarating customer relationships: Using CRM and relationship technologies*, Upper saddle river. N.J.: Prentice Hall PTR.
- Tsiptsis, K., & Chorianopoulos, A. (2009), *Data Mining Techniques in CRM: Inside Customer Segmentation*. A John Wiley and Sons, Ltd, Publication
- Torggler, M. (2008), *The Functionality and Usage of CRM Systems*. World Academy of Science, Engineering and Technology, Vol:2 2008-05-22
- Wahlberg, O., Strandberg, C., Sundberg, H. (2009), *Trends, Topics and Underresearched Areas in CRM Research*, (IJPS) *International Journal of Public Information Systems*.
- Wei, J.T., Lin, Sh.Y., Wu, H.H. (2010). A review of the application of RFM model. *African Journal of Business Management* Vol. 4(19), pp. 4199-4206.
- Yanyan, P. (2011), *Evaluation and Classification of Commercial Bank Customer Value*, Fourth International Conference on Business Intelligence and Financial Engineering, IEEE