Customer Segmentation Based on CLV Model and Neural Network

M. Ayoubi

Abstract

Today, for the customer oriented organization, identifying customers, creating a distinction between of them and ranking is one of the biggest challenges. In past times, separation of customers to the different group was taken based on customers' demands. But today's customer value as a measurable parameter, can be used for customers segmentation.so, by means of customer' segmentation based on their value and measuring of their segmentation, we can explain marketing strategy and customer relationship management. Customers segmentation based on effective factors on customer lifetime value is the main object of this paper. For this object, we collected information about 56000 customers of the "Taavon bank" between 1390 and 1391. Mixed data was removed from data collection. At first by means of two-step approach, the optimum number of clustering was determined. Then, for second level of segmentation," Kohonen neural network" was applied. Based on WRFM model, the value of each cluster was calculated. Finally golden customers, was recognized. Result of this paper can be used for marketing strategy and developing of new services for each segment of customers.

Keywords: customer value, customer segmentation, customer relationship management, customer life time value, neural network

Introduction

During last two decades in marketing, along with historical evolution in marketing topics a new topic as the relationship full marketing or relationship is formed. The first priority of businesses in relationship marketing is to take care of existing customers and trying to build a long term relationship with a reliable and long-term profitability target (Gomez, 1978). In this sense the necessity of having two-way communication with customers and place its managing that to a higher profitability along with customers' satisfaction and actual needs and demands.

The new concept which is born out of the relationship marketing and has created a lot of changes in this area is customer relationship management (CRM).

the ultimate goal of customer relationship management is to attract, maintain and develop long-term relationships with customers(Jane Vsyng, 2002), by advent and dissemination of customer relationship management on the one hand and changes in thinking and the paradigm of marketing principles considering customers long-term value (CLV) is increasingly becomes more important.

As noted above, customer relationship management is trying to establish a long-term and profitable relationship with customers. But always not all of the customers in an organization are profitable and sometimes they are costly. Therefore, companies should seek for their best customer; because profitability depends on this group of customer, but how to identify the best customers? Answer to this question can be obtained by the use of CLV pattern. This pattern provides an instrument which by the use of that can segment the customers and identify the best among them (Khajvand, 2010)

The objectives of this research separately include:

Determination of gold and valuable bank customers; respect to the goals and aspirations of relevant research

Customer classification using the results of analysis of long-term value of customer

Contribute to richness of literature in the field of customer relationship management (CRM), customer long-term value (CLV) and relationship marketing (RM) during the study, and help to promote a culture of value-based marketing.

The literature

Theoretical background

The concept of customer lifetime

The concept of customer lifetime value is derived from customer relationship management. So that piers et al. (2008) put the aim of CRM as a closer and deeper relationship with customers to maximize customer lifetime value to the organization. Concepts such as CRM are thought from the beginning of 1970 and established as an effective trading technique in a competitive business environment. Today company don't rely on the sale of goods/services, they are seeking to create profitable customers (Khajvand, 2010). According to Philip Kotler; companies not only wants to create profitable customers, but also they want to keep him all their life. So that research shows some very successful companies have their customer retention rate over 90% (Pfeifer, 2004). In past two decades, companies begun to focus on cost management and revenue growth. Many consumer-oriented companies believe that their own customers are more

valuable than profit in any financial exchange which they make to company. Customer value is simply the value that company receives during the period of his relationship with the client (Kumar and Lemon and Parasuraman, 2006).

The point is that each customer according to the type of business has a life time. Customer lifetime means the time that customer has relationship with a company or an organization or dose business with them (Blattberg and Deighton, 1996).

The definition of customer lifetime value

various definitions by researchers in the field of customer lifetime value (CLV) is provided and the reason for different definitions is the diversity on how to calculate the value of each various models that researchers have done calculation of CLV based on. Some models are posed and consequently the definitions provided by the organization or firm define CLV. Hereby some of these definitions are given.

Net present value of all overhead and benefits of client (Roberts and Berger, 1989).

Total net profit (by considering the time value of money) which over the lifetime of a customer relationship with the firm provide (Pearson, 1996).

The expected benefits of the customer, exclusive cost of customer relationship management (Blattberg and Deighton, 1996).

The present value of future benefits generated from interaction with a customer (Kumar and Lemon and Parasuraman, 2006)

Total cash flow accumulated during the relationship with a client and company (Kumar, 2006).

Customer long-term value

Once familiarity with the related concepts; represent customer long-term value. The main tool for customer relationship management, to answer fundamental questions about attracting, retaining and promoting customers can be customer long-term value (CLV) (Kumar & Lemon & Parasuraman, 2006). When it comes to talk about the customer long-term value, means a long-term process, with profitability for the company or organization. In this process organization attract customers and try to retain customers in the following years so that companies are trying to identify their customers and maintain their most valuable customers and the potential profitable one to boost their profitability and to eliminate weaker clients to reduce the company's costs hence increase profitability. The value as a result of customer's long-term relationships resulting from the attracting process, maintaining

and improving the customer called customer long-term value (khajvand, 2010).

Data mining

Simply, data mining is automatic discovery process of useful information from massive resources. Imielinski and Virmani (1999) introduced data mining as searching for patterns in large databases. Even it is possible to name data mining as the exploration science from data. Hen and cumber (2006) considered data mining as a step to knowledge discovery process. Knowledge discovery is a rotating process and its steps provided as follows:
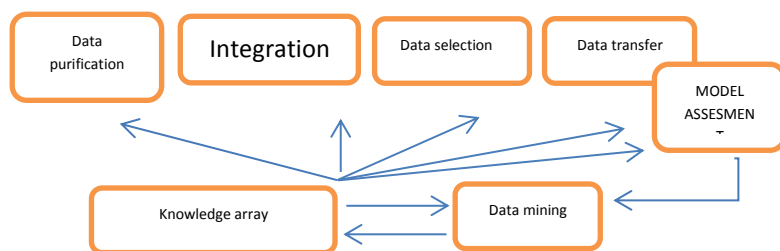

Figure 1. The process of data mining

In an overview data mining techniques are used for two purposes: to describe and to predict (Gupta et al., 2006). The uses of predictive techniques are to predict the value of a particular property based on the value of other features. Anticipated features called target attributes and are related to other characteristics and features that help to anticipate the explanatory variables and independent operation (Tan and Steinbach and Kumar, 2005). However, the purpose in using descriptive techniques is to extract patterns to summarize the data from communication between the layers. (Tan and Steinbach and Kumar, 2005)

Clustering

Clustering is an action which cause heterogeneous population segmentation (Anjay & Zeev & Chao, 2009) when a clustering process commence, the number of shapes and characteristics of clustering is not clear and there is no prior knowledge of clusters; the clustering technique is called an unsupervised (Khajvand, 2010). Clustering techniques usually are used for sectors and clusters that can be used for further analysis, such as segmenting customers into homogeneous sections which each section has unique characteristics and depending characteristics of each section; loyalty programs and marketing strategies for the sector are set.

Two-step algorithm

The algorithm is provided by SPSS software package to gets lower and upper limits for the number of clusters and provides

a clustering model based on the maximum and minimum number of clusters which is determined by user. This algorithm has globular clusters and provides clusters with arbitrary shape. As the name implies, this algorithm uses a two-step clustering method:

Pre-clustering. At first raw data compressed to provide a set of manageable clusters. At this stage a sequential clustering method is used to form a small sub clusters.

Clustering: at the second stage agglomerative hierarchical method is used to join clusters with each other to form larger clusters.

As an advantage of this method can point to the number of clusters which obtained by algorithm itself and can work with large databases (Rastegar, 2010).

Neural network

Neural network, usually called "network or education", is the simplified module of human nervous system. Its main unit is neurons, which usually are organized in layers. A neural network sometimes is called same named multilayer preceptors; which is a simplified pattern of human brain information processing. The neural network by simulating inner connection between the neurons works. Learning in neural network is gained by adjustment of each of these weights to the joints of these neurons. Weights adjustment received by measuring the amount of error in prediction. Over and over weights adjustment repeats in a neural network with various data output which they are known and at each repetition of the amount error measured and due amount of error weights adjustments for each connections corrected. These actions continue up to the time that proper function or the amount of risk in prediction reach to the acceptable limit and there are no limitations in the type of attributes. Neural networks can function with input data, output number, category or binary. To this end one or more attributes for input and one or several field as the output should be marked. Attributes which are set as "both" or "null" will not be considered.

Neural networks are powerful estimators in a way that do the prediction as well as other predictors (Freeman and Skapura, 1992)

Kohonen neural network:

Kohonen neural is developed neural network architecture. Kohonen creates a type of neural network that can be used in clustering of data set in separate segments. When the network is fully trained; similar records in output map come together while the different records come far from each other.

Acquired visit numbers by every single piece of pattern is visible so string units can be recognized. These provides a true witnesses of appropriate number of obtained clusters (Kwang and Yun-Chiang, 2006)

Another use of Kohonen network is to reduce dimensions. The distance characteristics of two dimensions network; provides a k mapping of the primary prediction to two extracted features which preserve similarity of the relationship between main predictors (Wang and Yom Chiang, 2006).

Experimental background

In general, several studies in the field CLV are done and different models are offered which can be categorized in 3 different groups. First group includes models that study on concepts such as marketing programs to attract and retain their customer. Conducted researches in this case include presented patterns by Kumar et al in 2006.

The second group includes models that show the relationship between different components of the CLV. For example, the relationship between customer acquisition and customer retention by Thomas in 2001 showed that these concepts have direct relationship with CLV. Reichheld in a study in 1996 showed that customer retention is the most sensitive factor which affects the CLV.

The third group is models who relate the CLV to the value of agency. One example of this pattern is Gupta et al study which by the use of 5 companies' data in 2004 proved CLV is a good representative to express the value of a firm. Another study is Kumar and Leo in 2006 that shows how the value of a brand can be specified for individual can affect his CLV.

In writings in the banking industry research has been done as well; and the glade et al pattern can be an example. In this study the value of customer life cycle as decreased value of future income margin is stated based on client activity. Another example in banking industry is the model which is provided by Hamlyn et al in 2007. Model is founded base on homogeneous grouping of customers based on 4 variables; age, lifestyle, services used by the client in bank and the level of activity.

As an Iranian investigation we can also refer to Sohrabi et al (1389) study which provides a model to determine that the value of customer life cycle in the banking industry, and Jafar Razmi in 1387 as a model to calculate the value of customer life in insurance industry.

Conceptual model

The different methodologies presented to implement data mining for projects, but one of the most powerful methodologies is the CRISP-DM (Larousse, 2006).

This methodology as a process method includes life cycle of a data mining project. Methodology of crisp is a flexible methodology and easily can adapt to the applications of data mining. The proposed framework of the present research is on the basis of these methodologies. Figure 2 shows the proposed framework of the research.

Figure 2. Research model-based on crisp

Methodology

the purpose of this study is development applied research; the reason is that this research has all characteristics of a developmental research including achieving a better understanding, develop an idea, use a combination of methods, duration (Larousse, 2006), and in the subject of study "analysis and analysis the customer long time value (CLV) » is true. This article aims to apply scientific principles and patterns to provide a framework for the analysis of customer long-term value of groups of customers and this research develop available knowledge in this field. On the hand this research is an applied research because recommended tool is used as an executive phase in organization and practically does the analysis of customer long-term value concerned bank groups.

Findings of the study

At first, banks actual customers current account opened before 89 and their transactions from the beginning of April 89 until the beginning of 91 received. The number of records was 56000, but after eliminating confused records and refining reduced to 33194 records. Information used categories in two general categories: demographic information and customers' behavioral data (transactional). Demographic data includes fields such as age, gender, education level and marital status. Transaction information includes data related to the monetary value of transactions, the time of referring to the bank at first time, the last refer to the bank by the end of 90, the number of transactions, the average time between the uses of services. Requests and receiving data in the form of these titles were handed to the research: elite customers, demographic data of customers, number of transactions, and date of the last transaction of deposit.

In the next phase the data was then normalized, meaning that it can be used in clementine software. Data in from of csv format prepared in excel and is encoded into clementine software.

SPSS clementine software is known as data mining software in commercial activities and some of the regarding clustering and classification of the results and related tasks in the research are

implemented by this software. Calculation of adjusted parameters M, F, R:

Parameter (R) or recency that called as recency in this study and calculated based on the last time reference of customer. The exact method of calculation is that the beginning period of supposed data since the last visit deducted and the number of days shows the recency of customer. It is clear the larger value indicates the date is closer to present time therefore the customer has a higher rank and will regarded as loyal.

Parameter (f) is called in the literature as the frequency of purchase. in this study, given that the case study is a bank, it is the number of transactions, as the title express the total number of transactions of all deposits for a customer during the period under consideration (for 1 year from march 90 to march 91).

parameter (m) it is also the financial factor or monetary among these three factors, in literature it is the amount paid to purchase a product or service that in the present study it is measured and valued by the total balance of all deposits of customers at the end of one-year period (Amir, 2010)

These parameters are the same parameters affecting the amunt of the customer value and by using them value and rating determined and calculated (Amir, 2010-Rynartz and Kumar, 2004-van Kansan and Kumar, 2004, Khajvand, 2010).

Modeling

There are various algorithms and different factors for customer clustering, which influence the choice of algorithm and have been mentioned earlier. But one of the widely used algorithms is the two-step algorithm is. This algorithm is an unsupervised algorithm which simply and easily can count the clusters based on its training. This algorithm due to the lower and upper limits which are given by user to the application determines the optimal cluster. Similarly, when we want another redistricting process in segmentation, this pattern can be a suitable model. Also in this study obtained number is by two-step model and entered into neural network, which based on neural network architecture it re-segmented. Another algorithm is optimized SOM algorithm which is the same Kohonen. This algorithm is based on neural network to algorithms and it is widely used algorithm in clustering. One of the problems in this algorithm is its disability in determining the hidden relationships and patterns. The goal of the study of is not to determine patterns and relationships so this can be suitable tool for this study.

Before entering data into the node of Kohonen neural network we should optimized clusters. To this end two-step node is

used. Using this; the appropriate number of nodes and to the optimized clusters is obtained.

Then the number of obtained clusters enters to Kohonen node to obtain the number of optimized cluster.

The achieved clusters evaluate to identify valuable clusters. To do this, CLV pattern is used. Obtained cluster each have a central number as an indicator of recency of transactions, a central number as an indicator of the volume of transactions and a central number as an index of frequency. To calculate CLV; these numbers placed in formula (1), and calculate value of each cluster. The clusters which has more CLV; are identified as more valuable cluster. As the subject of study was bank and the bank used as a financial service provider firm so CLV pattern is suitable to calculate the value of WRFM cluster pattern. "W" in the WRFM represents the weight of each parameter. To gain weight of each parameter the method of AHP and opinion of three branches, the marketing, resources equipment, plan and programing experts have been used. Based on this calculated weight for parameters are provided in the following table:

Table 1: parameters' weight

parameter	R	F	M
weight	087 / 0	0 345 / 0	568 / 0

The value of each cluster is obtained in accordance with the following formula [9, 17, and 9]:

Equation 1) $CLV_{ci}=NR_{ci}* WR_{ci}+NF_{ci}*WF_{ci}+NM_{ci}* WM_{ci}$

In above equation the values of r, f and m are the values of each cluster which are obtained by Kohonen neural network. N in the above equation indicates the presence of a normalized parameter; because the type of present parameters of in this equation is not the same and its and big or small of a parameter strongly affect the equation. To normalize the data, normal values by the following formula are used (khajvand, 2009).

Equation 2) $v=\frac{v-Min_A}{Max_A-Min_A}$

Conclusion and data analysis:

According to the subjects mentioned in the previous section, the numbers of cluster nodes that are in the first stage of two-step method are a 3cluster. By entering 3 optimized number into Kohonen nodes and their implementation optimized cluster will be achieved which are 3. Then the achieved clusters will be evaluated WRFM model. The below table provide the results of the first phase:

Table 2: clusters and normality of parameters

cluster	normal r	normal f	Normal m	CLV value	rating cluster
1	0 / 316	0 / 00039	0 / 00073	0 / 02,811	3
2	0 / 56	0 / 003	0 / 003	0 / 0506	2
3	0 / 92	0 / 02	0 / 02	0 / 092	1

According to the above results the third clusters was known as more valuable cluster and in the first phase the 10,986 customer was recognized as valued customer and other customers in the other clusters were identified as less valuable customers.

As we look for customers' value and creating an accurate classification and understanding; by the use of a two-step model and Kohonen clustering we cluster obtained customers from cluster 3. To this end the present table node software environment for identification and separator of customers in cluster 3 are used to produce usable data for new clustering.

In the second phase, after sorting the data and making them usable, new clustering performed. Repeat the previous steps. As it is seen two-step node, determine 3 clusters to entry to Kohonen node. Kohonen node provides 3 sets of records. Following table represent the results of the phase 2.

Table 3: clusters and normality of parameters

cluster	Normal r	Normal f	Normal m	value CLV	rating cluster
1	78/0	0062/0	00014/0	067/0	3
2	917/0	0098/0	0025/0	025/0	2
3	98/0	28/0	012/0	102/0	1

As it can be seen the third cluster is the optimized cluster, and possesses higher value of CLV compared to other clusters.

As 33 percent of customers were chosen in the first phase of the cluster and in the second phase 50% of the costumers were chosen, therefore, 16.5% of customers determined as golden customer, and this criteria is in line with required technical demands of the banks which wanted to consider 10 to 20 percent of customers and apply marketing plan on that group.

IJCSI
www.IJCSI.org

As the data in each cluster are so close to each other and in re-clustering they don't tend to separate and this study reached into local optimum point.

Topics and applications:

This study is useful in determining marketing strategies and formulation of CRM programs based on group separation, strategies which can be offered based on the results of this study include:

Design multiple products and services, and a variety of different node for customers, especially "relatively" high-valued customers "that make up the population of the large number of customers.

Providing maintenance programs for "golden clients" that make up approximately 16 percent of customers population, but they provide a large amount of resources for the bank.

Exclusive marketing specialization and separation of retail banking to designing private banking programs, to manage customers' asset to achieve greatest returns for customers in the cause-for the "golden customers" group to enhance their satisfaction.

Use the results in lending to the facilities applicant based on the rank value of each customer group, to be considered as one of inputs in measurement of customer facilities credit.

The results can be valuable in assessing the branches in attracting customers and each branch share from each group's customers.

Other research achievements can be found in the following:

Solve the problem of selected refining feature based on content using initial stages of joint marketing analysis and map preferences on the characterization of the group of services (yang, 2005).

Providing a new model-based survival clustering and triple clients ranking, and achieve better results compared to similar models.

Conclusion

In this study, to identify and clusters customers a new structure was used. In this structure; using data mining and analysis customers' value to an organization were determined. First, data to be used in clementine environment were purified and entered. The number of optimized cluster nodes to log on the network by two-step model was 3 then this number entered to nodes of Kohonen which is based on neural network architecture and after that implementation 3 node clusters

obtained. The value of each cluster by model CLV is determined, but since it is not the local optimum point, re cluster again and repeats the above steps. Thus, 16.5 percent of the clients to the number 5493 are determined as the most valuable customer and the most valuable classes.

By identifying valuable customers, other customers marketing programs including strategy, marketing and increasing profits and allocation of resources and other objectives related to economic issues and marketing will be effective and efficient and 20% of the most valuable customers bring 80% of profits.

Suggestions

In this paper, the field of customers' retail banking based on value and division design was selected. The field of electronic banking, institutional banking and private banking as well; regarding clients and their area of activities can be selected and implemented as target customers.

Category of customers based on customer value has been the main objective of this study and research on sectors and identify factors affecting the customer's credit rating and the implementation model for the assessment of customer credit can be investigated.

References

[1] Amiri, Fahimeh (2010). Consider Of Customers Treatment by Use of Banking Tools Based On Data Mining Approach, M.B Thesis, Alzahra University (In Persian)

[2] Blattberg, R. And Deighton J, (1996). Managing Marketing By The Customer Equity Test: Harvard business , Review

[3] Freeman, J. A. & Skapura. D. M. (1992). Neural Network Algorithms, Applications, and Programming Techniques. Ma: Addison-Wesley Publishing Company.

[4] Ghazanfari, Mehdi (2008). Data Minig And Knowledge Exploration. Elm O Sanat Uni Publishing,P 39-47 (In Persian)

[5] Gummesson, Evert. (1978). Toward A Theory Of Professional Service Marketing. Industrial Marketing Management, Volume 7, Issue 2, 89-95.

[6] Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Et Al, (2006). Modeling Customer life-Time Value. Journal of Service Research

[7] Imielinski, T., and Virmani, A. (1999). Msql – Query Language for Data Mining Applications. Datamining and Knowledge Discovery Journal

[8] Jain, D., & Singh, S. S. (2002), Customer Lifetime Value Research In Marketing: A Review and Future Directions. Journal of Interactive Marketing, 16(2), 34–45.

[9] Khajvand,Mahbubeh, (2010), Customer Segmentation Based On CLV Model , M.B Thesis, Khaje Nasir Uni (In Persian)

[10] Kumar, V.(2006) , "CLV: A Path To Higher Profitability," Working Paper, University Of Connecticut, Storrs

[11] Kumar, V., Lemon, K. & Parasuraman, A. (2006). Managing Customers for Value: An Overview and Research Agenda. Journal of Service Research, 9(2), 87-94.

[12]Kwang, Yun-Chiang. (2006)Artificial Neural Networks: A New Methodology for Industrial Market Segmentation. Industrial Marketing management, 24, 431-438.

[13] Larose, D. T. (2006).Data Mining Methods and Models. John Wiley & Sons, Inc., Hoboken, New Jersey,

[14] Ngai, E. W. T., Xiu, L., And Chau, D. C. K.(2009). Application of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classification. Expert Systems with Applications, 36, 2592–2602

[15] Pfeifer, P. E. &, F., P. W. (2004). the Elasticity of Customer Value to Retention: The Duration of a Customer Relationship. Journal of Interactive Marketing, 18, 20-31.

[16] Pearson S. (1996) Building Brands Directly: Creating Business Value From Customer Relationships. London: Macmillan Business

[17] Rastegar, Negar, (2010). Customer Segmentation Based On WRFM And LTV Models, Ph.D Thesis, Elm O Sanat University. (In Persian)

[18] Razmi, Jafar , (2010) , Presentation Of New Model Based On Customer Life Time, It Management Publishing, No.1, Page 35-50, (In Persian)

[19] Reinartz, W. And Kumar, V. (2007). The Impact of Customer Relationship Characteristics On Profitable

[20] Roberts M. L. And Berger P. D. (1989). Direct Marketing Management. Prentice-Hall: Englewood Cliffs, Nj

[21] Tan, P. -N., Steinbach, M., And Kumar, V. (2005) Introduction To Data Mining. Pearson Education

[22] Venkatesan, R. And Kumar, V. (2004). A Customer Lifetime Value Framework for Customer Selection And Resource Allocation Strategy, Journal Of Marketing, 68 (4), 106-25

[23] Yang, A. X. (2005). Using Lifetime Value to Gain Long-Term Profitability. Database Marketing and Customer Strategy Management, 12, 142-152.