Integrating AHP and Data Mining for Effective Retailer Segmentation Based on Retailer Lifetime Value

Amin Parvaneh[a], Hossein Abbasimehr[b], Mohammad Jafar Tarokh[c,*]

[a]Post graduate student, IT Group - Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

[b]Post graduate student, IT Group - Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran

[c]Associate professor, IT Group - Faculty of Industrial Engineering , K. N. Toosi University of Technology, Tehran, Iran

Received 23 April, 2011; Revised 14 August, 2011; Accepted 22 February, 2012

Abstract

Data mining techniques have been used widely in the area of customer relationship management (CRM). In this study, we have applied data mining techniques to address a problem in the business-to-business (B2B) setting. In order to continue its business in a manufacturer-retailer-consumer chain, a manufacturer should improve its relationship with retailers. In this regard, segmentation is a useful tool for identifying groups of similar retailers in order to improve retailer loyalty by developing and implementing segment-specific marketing strategies. Therefore, this study proposes a methodology for retailer segmentation based on the value-based segmentation and the analytical hierarchy process (AHP). The proposed methodology is implemented by using the data of a firm in the hygienic industry in Iran. As a result, having found six groups of retailers and labelled them according to their performance, we provide some possible measures that can be taken in order to improve the relationship between the firm and its retailers.
Keywords: Data mining; Value-based segmentation; Clustering; AHP.

1. Introduction

Data mining is the task of extracting valuable information and knowledge from huge amounts of data (Han and Kamber, 2006). Data mining tasks can be classified into two categories: descriptive and predictive. Clustering is an instance of descriptive methods, and classification is an example of predictive methods (Han and Kamber, 2006). Data mining techniques have been used widely in many different areas such as customer relationship management (CRM) (Ngai et al., 2009) and in particular, market basket analysis (Berry and Linoff, 2004) and customer churn prediction (Coussement and Van den Poel, 2008). Customer segmentation is another important application of data mining, especially clustering in CRM. It involves partitioning the customer base into smaller customer segments according to their similarity. Most of the researches in the CRM area belong to the B2C (business-to-consumer) setting. However, in this study we propose a new methodology for customer segmentation in the B2B (business-to-business) setting. In fact, we address a problem in a manufacturer-retailer-consumer chain.

In a manufacturer-retailer-consumer chain, the role of retailers can be crucial in persuading consumers to purchase products of a typical manufacturer. Product homogeneity in any product increases the number of choices for consumers and thus complicates the decision-making process for them. In this situation, any recommendation from the retailer regarding a particular brand or product may influence consumers' decisions. As a result, by improving its relationship with retailers, a manufacturer can gain great benefits. In this regard, retailer segmentation enables manufacturers to better understand retailer behaviors in order to adopt right and segment-specific marketing strategies for them. Consequently, executing segment-specific marketing programs may lead to a profitable long-term relationship between a manufacturer and retailers.

In this study, we propose a methodology for retailer segmentation based on the value-based segmentation and the analytical hierarchy process (AHP). The main difference between this study and the previous researches in the field of CRM is that this research focus on the B2B setting rather than the B2C setting. This is also the main contribution of this work. Then we implement the developed methodology by using the data of a firm in the hygienic industry.

* Corresponding author E-mail address: mjtarokh@kntu.ac.ir

The rest of this paper is organized as follows. Section 2 provides a brief background on concepts such as value-based segmentation, clustering, and analytical hierarchy process (AHP). In section 3, the methodology is explained. Section 4 describes the empirical study and its results. Finally, the conclusion is presented in section 5.

2. Background

2.1 Business-to-Business market segmentation

Business-to-Business (B2B) organizations don't sell their products or services to end consumers directly. They do that via intermediaries. For example, the manufacturer of a typical product distributes its products to some retailers, and then they sell those products to the end consumer.

It is clear that the success of a B2B organization depends on its intermediaries. For instance, in the manufacturer-retailer-consumer chain, a manufacturer needs to rely on cooperating with retailers in order to sell a large volume of products to make profit. Therefore, identifying the high-value and profitable retailers can be an essential task for the manufacturer. Segmentation tools can help manufacturers with this task of identifying different groups of retailers.

In the business-to-customer (B2C) context, customer segmentation approaches are divided into customer need-based, characteristics-based segmentation (Greengrove, 2002), and customer value-based segmentation (Kim et al., 2006), through which many customer segmentation researches have been performed. In these studies, several data mining techniques have been used to group customers into different businesses and industries such as hardware retailing (Liu and Shih, 2005), retail industry (Ho Ha, 2007), textile manufacturing (Li et al., 2011), electric utility (Lopez et al., 2011) and so on. Most of these studies have used the value-based segmentation.

Contrary to the B2C context in which customer segmentation has received considerable attention, the B2B context suffers from lack of enough studies in the segmentation area. Yet, in the B2B setting there are some studies focused on customer loyalty. For instance, Lam et al. (2004) proposed and analyzed a conceptual framework for identifying factors affecting customer loyalty in a B2B context, including customer perceived value, customer satisfaction, and switching costs. Davis-Sramek et al. (2009) also investigated factors influencing retailer loyalty in the supply chain for consumer durable products.

In this study, we use some concepts belonging to the B2C context in order to group retailers (customer in our case). More specifically, we aim to apply the value-based segmentation in a manufacturer-retailer supply chain to identify different groups of retailers.

2.2. Value based segmentation

Customer segmentation based on customer value is a common approach used to identify profitable customers in order to develop strategies to target them. Customer value is often known as LTV (Life Time Value), CLV (customer Lifetime Value), CE (customer equity) and customer profitability (Kim et al., 2006). According to Kottler (1974), CLV is "the present value of the future profit stream expected over a given time horizon of transacting with the customer.

There are many models developed for measuring CLV (Gupta et al., 2006), among which the RFM (Recency, Frequency and Monetary) model developed by Hughes (1994) is an important model for estimating customer lifetime value. The RFM model was then extended by Chang and Tasy (2004) by adding the customer relation length (L) to it, therefore the LRFM model was developed.

2.3. Weighted LRFM

The RFM model has three dimensions: (1) Recency: is the time interval between the last purchase and a present time reference; the shorter the time interval is, the bigger R is, (2) Frequency: is the number of customer's purchases in a particular period; a higher frequency is more valuable, (3) Monetary value: the total amount of money consumed by the customer over a particular time period; the higher the monetary value is, the bigger the contribution to business is.

Although RFM and its successor LRFM make it possible to assess CLV, there are some challenges to use them. The major challenge relates to the importance of four variables of L, R, F, and M and determining their weights. Experts have differing views on this issue. For instance, regarding the RFM model, Hughes (1994) showed that the importance (weight) of the three variables is equal while Stone (1995) considered different weights for the RFM variables. The weight of each RFM variable depends on the characteristics of the industry. In addition, there are some researches that have used the weighted RFM model (e.g. Liu and Shih, 2005; Seyed Hosseini et al., 2010).

It is important to note that in the studies that used the RFM and LRFM models (e.g. Liu and Shih, 2005; Seyed Hosseini et al., 2010), no relationship was found between the variables. As a result, the LRFM variables are considered independent; for example, the high frequency does not affect the high monetary and vice versa. In this study, we determine the weights (relative importance) of each LRFM variable through conducting a survey of the AHP, which is a simple solution for this problem.

2.4. Clustering

Clustering, which is a subset of unsupervised learning techniques, is the process of grouping a set of objects into classes of similar objects. There are many clustering methods including partitioning methods, hierarchical methods, density-based methods, grid-based methods, and model-based methods (Han and Kamber, 2006). K-means from the category of partitioning methods is the most widely used clustering algorithm in CRM and marketing. This algorithm introduced by MacQueen (1967) can process large amounts of data quickly. The operation of K-means is as follows: (1) selecting K initial centroids; (2) assigning each object to its closest centroid; (3) updating the centriod of each cluster to the mean of its constituent instances; and (5) repeating steps 2 and 3 until centroids stop changing.

2.5. Analytical Hierarchy process (AHP)

The Analytic Hierarchy Process (AHP) developed by Saaty (1980) in the 1970s is a method for multi-criteria decision-making. It is useful for integrated and fuzzy problems based on the human brain assessment. The AHP uses paired comparison judgments from a fundamental scale of absolute numbers approached by decision-makers to prioritize alternatives for a problem in an architectural structure (Saaty, 2003). Decision-makers must assign a number from 1 to 9 to each comparison (Table 1). This method also measures the degree of inconsistency between judgments. If the inconsistency degree exceeds 0.1, judgments must be revised.

3. Methodology

Our proposed methodology for retailer segmentation is shown in Figure 1. In our methodology, we use the LRFM model which, as mentioned earlier, is the extended version of the RFM model and considers the customer relationship length (L) in determining the value of each customer. Throughout the rest of this paper, we use the word 'customer' instead of 'retailer' since we focus on the relationship between a manufacturer and its retailers, and retailers ask for what they want from the manufacturer. Therefore, retailers are customers for the manufacturer.

Table 1
Numbers for representing paired comparison judgments

Comparison importance	Description
1	Equal
2	Intermediate between equal and moderately dominant
3	Moderately dominant
4	Intermediate between moderately and strongly dominant
5	Strongly dominant
6	Intermediate between strongly and very strongly dominant
7	Very Strongly dominant
8	Intermediate between very strongly and extremely dominant
9	Extremely dominant

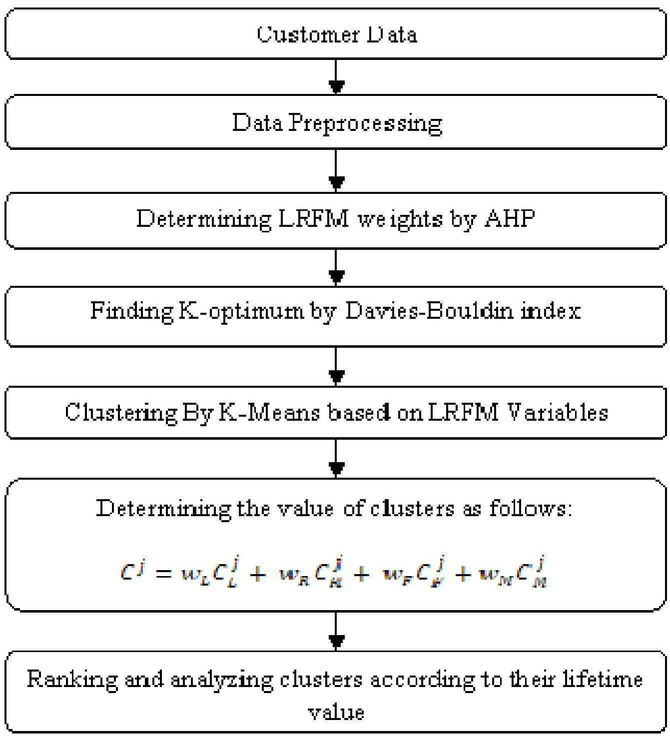

Fig. 1. The Proposed Methodology

4. The Empirical Study

4.1. Customer data

The case studied in this research is a hygienic manufacturer with over than 100 products. This manufacturer is one of the biggest and oldest producers of hygienic and cosmetic products in Iran. It distributes its products to his customers that range from small retailers to chain and big retailers across the country. The decision-makers of this corporation must target some customer groups and develop different marketing strategies for each group of them based on their value. Therefore, customer segmentation has an important role in determining these strategies.

There are two datasets of the corporation used in this study: a dataset of 18-month customers' transactions with about 1000000 records, and a dataset of customers’ profiles including about 72000 records.

4.2. Data Preprocessing

Data preprocessing is an important step in the data mining process because it improves the accuracy and efficiency of subsequent modeling (Han and Kamber, 2006; Tan et al., 2005). In this paper, some data preprocessing techniques such as data cleaning, data transformation, data integration and data reduction are used to improve the quality of data for clustering. Customers who did not make any purchase during one last year are removed from the dataset. After performing this step, we reach to a dataset with 63599 customers. From the integrated dataset, the L, R, F and M variables are extracted for each customer. It is important to notice that in this research the L value (customer relationship length) is computed 6 monthly because of the large distance between customers' relationship length and bad effect of this long distance on clustering when we use it by days. In addition, we are interested to take into account the new customers.

4.3. Determining LRFM weights by the AHP

In this paper, the AHP method is utilized for calculating the LRFM weights according to the opinions of decision-makers. This is done through 3 steps according to the AHP definition. First, four decision-makers from the three different management layers of the sale department are selected for making paired comparisons. They include one top level manager, two middle level managers, and one operational manager. You can see the results of their judgments in Appendix A. In the second step, the inconsistency index is computed and checked for each decision-maker judgment. At last, the LRFM weights are determined by computing the eigenvalues (Figure 2). The calculated weights for LRFM are 0.238, 0.088, 0.326, and 0.348, respectively.

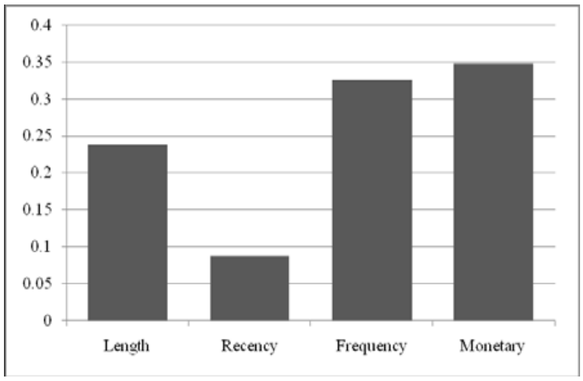

Fig. 2. LRFM weights obtained from the AHP

4.4. Finding the K-optimum by the Davies-Bouldin index

Many clustering algorithms have been introduced; however, there is no best algorithm. In fact, due to the exploratory nature of clustering, seeking the best clustering algorithm is useless. Yet, the K-means algorithm is the most popular partitioning algorithm (Jain, 2010). According to Jain (2010), despite being proposed over 50 years ago, K-means is still one of the most widely used clustering algorithms. The main reasons for the popularity of K-means are its easy implementation, simplicity, efficiency, and practical success (Jain, 2010).

In this study, we use the K-means algorithm for clustering retailers. K-means requires the user to determine the number of clusters. As the improper selection of k as the number of clusters may lead to inaccurate results, there are useful clustering quality indexes that can help with determining the optimal number of clusters. We use the Davies–Bouldin index (Davies and Bouldin, 1979) for this purpose. Identifying sets of clusters that have small intra-cluster distances and large inter-cluster distances is the aim of this index (Davies and Bouldin, 1979).

The Davies –Bouldin index (DB) is defined as:

$$DB = \frac{1}{k}\sum_{i=1}^{k} max_{i\neq j}\left\{\frac{a_i+a_j}{d(C_i,C_j)}\right\} \quad (1)$$

where k is the number of clusters, a_i is the intra-cluster distance of cluster i, and $d(C_i, C_j)$ represents the inter-cluster distance between clusters i and j. The number of clusters that minimizes the DB index is taken as the optimal number of clusters. In this study, the optimal number of clusters based on the Davies-Bouldin index is 6 clusters (Fig 3).

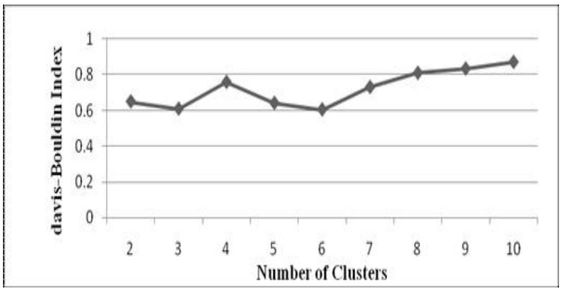

Fig. 3. Cluster Validation

4.5. Clustering by K-Means based on the LRFM variables

In this stage, customers are segmented into similar clusters using K-means and according to their LRFM variables. The number of clusters for the K-means algorithm is set to be 6 according to the results of the previous stage. After performing the clustering, we obtain 6 clusters. Table 2 shows the clustering results.

Table 2
Clustering results

Cluster	#Customers	Length	Recency	Frequency	Monetary
1	12956	6.43	67.498	10.472	20131186.45
2	3860	1.596	256.847	3.57	6804235.026
3	3065	7.899	14.702	32.215	156010514.4
4	13859	2.263	41.93	8.685	16838199
5	5060	5.805	264.484	5.805	11130428.05
6	24799	7.869	22.991	14.655	22991847.78

4.6. Calculating the values of clusters

To calculate the value of each retailer segment, we normalize the LRFM variables for centroids by using the Min-Max normalization method. For more information about this method, you may refer to Han and Kamber (2006).

Having normalized the LRFM values, we calculate the CLV of each cluster as follows:

$$C^j = w_L C_L^j + w_R C_R^j + w_F C_F^j + w_M C_M^j \quad (2)$$

Where C^j is the LRFM rating for cluster j, $C_L^j, C_R^j, C_F^j, C_M^j$ are the normalized L, R, F, and M for cluster j, and w_L, w_R, w_F, w_M are the related weights of L, R, F, and M obtained from the AHP.

4.7. Ranking and analyzing the clusters according to their lifetime values

After calculating the CLV for each cluster, we rank the clusters according to their CLV values (see Table 3).

Ranking retailer segments according to their lifetime values can help managers to allocate marketing resources according to the profitability of each segment.

In addition, an in-depth analysis of each segment by knowing their LRFM attributes may inform the firm about the purchasing behavior of each segment. This in turn can help marketing managers to develop effective marketing strategies that can lead to a profitable long-term relationship with retailers.

In order to analyze the clustering results, we use the customer value and customer loyalty matrices. The customer value matrix proposed by Marcus (1998) uses the two parameters of customer buying frequency (F) and monetary value (M) as its two axes. This matrix is illustrated in Figure 4.

Table 3
Clusters information

Cluster	#Customers	Length	Recency	Frequency	Monetary	CLV	CLV Rating
1	12956	0.76693638	0.788631687	0.240949555	0.089318972	0.361563004	3
2	3860	0	0.030574661	0	0	0.00269057	6
3	3065	1	1	1	1	1	1
4	13859	0.105822624	0.890992946	0.178565195	0.067248939	0.185208048	5
5	5060	0.667777249	0	0.078024088	0.028994711	0.194456998	4
6	24799	0.995240362	0.966815063	0.38697853	0.108491498	0.485856974	2

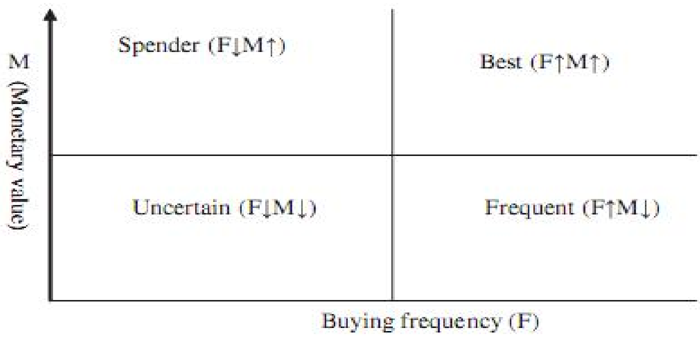

Fig. 4. The customer value matrix (Marcus, 1998)

Two other indicators are customer relationship length (L) and recency (R) that relate to customer loyalty, and we can consider them as the customer loyalty matrix.

According to the opinion of the marketing managers of the firm, new customers are those who have launched their relationship with the firm in the last 1.5 years (three six-month periods). Based on this assumption, we consider the customers with their L (length of relationship) lower than 3 as the new customers, and those with their L higher than 3 as the long life (established) customers.

For the recency indicator, we consider two states of Low and High. If the recent transaction time value of a cluster is smaller than the average total value, it is considered a High Recency value cluster, but if it is larger than the average total value, it is regarded as a Low Recency.

For frequency (F) and monetary (M) dimensions, we also consider two states: Low and High. If the frequency/or monetary value of a cluster is smaller than the median point, it is considered a Low frequency/or monetary value cluster; otherwise, it is considered a High frequency/or monetary value cluster. As Table 2 shows, the median point for all Monetary values is M=18484692.72 and the median point for all frequency values is F=9.58.
Figure 5 illustrates the status of all clusters.

After analyzing each segment, we label each cluster according to its status (Table 4). Furthermore, we suggest some possible actions that can be taken in order to improve the relationship between the firm and retailers.

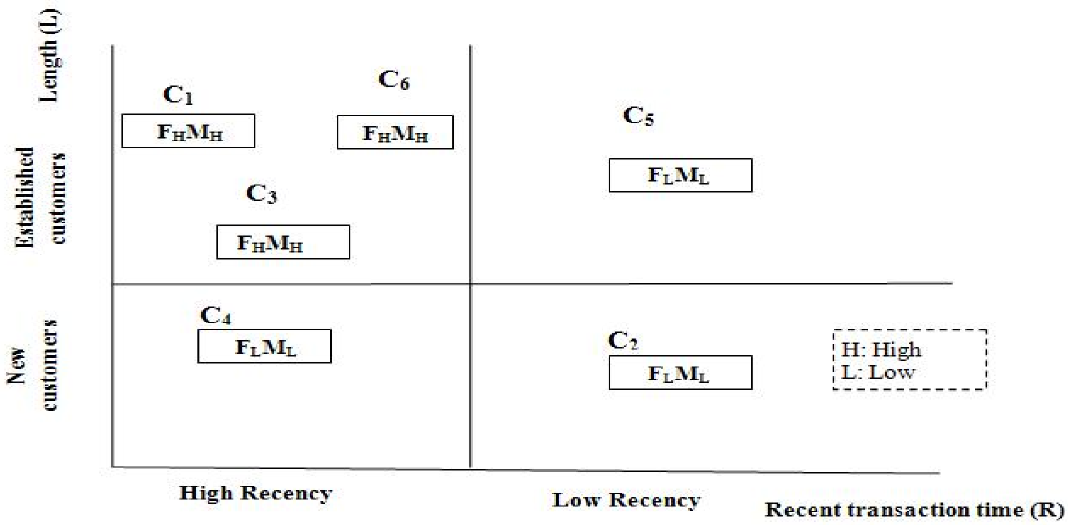

Fig. 5. Clusters status

Table 4
Cluster Labeling

Cluster	Cluster Label	Description	Possible actions
C_3	Superstar Segment	The highest value, the highest frequency, the highest recency, and the highest lifetime.	Special attention should be paid in order to retain retailers of this segment
C_6	Golden segment	The second highest value, highest frequency, highest recency, and a high lifetime.	There are many retailers (0.39 percent of all retailers) belonging to this segment. Strong strategies should be developed in order to maintain relationship between the firm and retailers of this segment.
C_1	Average value segment	This is an average value segment that has low basket size.	Marketing programs are should be developed in order to increase basket size of this segment
C_5	Dormant segment	This segment has low recency, low frequency, and low monetary value.	Although they have a long time relationship with the firm, they exhibited very bad performance. In addition the recency of this segment is very low; this may be a sign of attrition or long hiatus. Strong anti-attrition programs should be developed for this segment.
C_4	New Low value customer	This segment has high recency; this means that they maintain their relationship with the firm. But they have low frequency and low monetary value.	Because the number of retailers belonging to this segment is relatively high, marketing programs for this segment should encourage the retailers to buy more products.
C_2	New Dormant segment	This segment has low recency, low frequency, and low monetary.	The recency value for this segment is very high. This segment is at risk of churn. By considering that the value of this segment is very low, Thorough analysis should be carried out to understand whether the retailers of this segment are worth keeping or not

5. Conclusion and Future Work

Contrary to the B2C setting, the B2B setting suffers from lack of enough studies. Thus, conducting this study in the B2B setting, we propose a methodology for retailer segmentation based on the value-based segmentation and the analytical hierarchy process (AHP). In fact, we addressed a problem from a manufacturer-retailer-consumer chain. Because of their intermediary role in the relationship between a manufacturer and consumers, retailers are important for the survival of the manufacturer. We implemented the proposed methodology by using the data of a firm in the hygienic industry. The results indicated that there are six groups of retailers. After analyzing each segment, we labeled each retailer group according to its performance. Finally, we provided some possible actions that can be taken in order to improve the relationship between the firm and retailers.

Those who may be interested in doing further research into the topic of this paper may perform in-depth analyses of all segments to gain a better understanding of their behavior in order to see how segment-specific marketing strategies could be made and thereby long-term and profitable relationships with retailers could be built.

6. Acknowledgment

We would like to thank the anonymous Iranian manufacturer of hygienic products for providing us with the data.

7. References

[1] Berry, M. and Linoff, G. (2004). Data mining techniques: For marketing, sales and customer relationship management. New York, NY: John Wiley & Sons.
[2] Chang, H. H., & Tsay, S. F. (2004). Integrating of SOM and K-mean in data mining clustering: An empirical study of CRM and profitability evaluation. Journal of Information Management, 11, 161-203.
[3] Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert Systems with Applications, 34, 313–327.
[4] Davies, D. L. and Bouldin D. W. (1979). A cluster separation measure. IEEE Transactions on Pattern Recognition and Machine Intelligence 1(2), 224-227.
[5] Davis-Sramek, B., Droge, C., Mentzer, J. T., & Myers, M. B. (2009). Creating commitment and loyalty behavior among retailers: what are the roles of service quality and satisfaction?. J. of the Acad. Mark. Sci., 37, 440–454.
[6] Greengrove, K. (2002). Needs-based segmentation: principles and practice. International Journal of Market Research (IJMR) 44(4), 405-421.
[7] Gupta, S., Hanssens, D., Hardie, B., Kahn, W. Kumar, V. and Lin, N. (2006). Modeling Customer Life-Time Value. Journal of Service Research, 9 (2), 139-155.
[8] Han, J. and Kamber, M. (2006). Data Mining Concepts and Techniques, 2nd edition, Morgan Kaufmann
[9] Ho Ha, S. (2007). Applying knowledge engineering techniques to customer analysis in the service industry. Advanced Engineering Informatics, 21,293–301.
[10] Hughes, A. M. (1994). Strategic Database Marketing. Chicago: Probus Publishing.
[11] Jain, A. K. (2010). Data clustering: 50 years beyond K-means, Pattern Recognition Letters 31(8), 651-666.
[12] Kim, S.-Y., Jung, T.-S., Suh, E.-H. and Hwang, H.-S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. Expert Systems with Applications 31,101–107.
[13] Kotler, P. (1974). Marketing during periods of shortage. Journal of Marketing 38(3), 20–29.
[14] Lam, S. Y., Shankar, V., Erramilli, M. K. and Murthy, B. (2004). Customer value, satisfaction, loyalty, and switching costs: An illustration from a business-to- business service context. Journal of the Academy of Marketing Science 32(3), 293.
[15] Li, D. C., Dai, W. L., and Tseng, W. T. (2011). A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. Expert Systems with Applications 38(6), 7186-7191.
[16] Liu, D. R., and Shih, Y. Y. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. Information & Management, 42, 387–400.
[17] López, J. J., Aguado, J. A., Martín, F., Mu noz, F., Rodríguez, A., and Ruiz, J. E. (2011). Hopfield–K -Means clustering algorithm: A proposal for the segmentation of electricity customers. Electric Power Systems Research, 81,716
[18] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume I, Statistics. Edited by Lucien M. Le Cam and Jerzy Neyman. University of California Press.
[19] Marcus, C. (1998). A practical yet meaningful approach to customer segmentation. Journal of Consumer Marketing 15(5), 494–504.
[20] Ngai, E. W. T., Xiu, L., & 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.
[21] Saaty, T.L. (1980).The analytic hierarchy process. New York: McGraw- Hill.
[22] Saaty, T.L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary?. European Journal of Operational Research, 145, 85–91.
[23] Seyed Hosseini, S. M., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. Expert Systems with Applications, 37, 5259–5264.
[24] Stone, B. (1995). Successful Direct Marketing Methods, Lincoln-wood, NTC Business Books, IL.
[25] Tan, P. N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining. Pearson Education.