Customer Clustering Using a Combination of Fuzzy C-Means and Genetic Algorithms

2 authors, including:

Azarnoush Ansari
University of Isfahan
57 PUBLICATIONS **269** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project online purchase View project

Project mobile banking View project

Customer Clustering Using a Combination of Fuzzy C-Means and Genetic Algorithms

Azarnoush Ansari[1] & Arash Riasi[2]

[1] Department of Management, University of Isfahan, Isfahan, Iran

[2] Institute for Financial Services Analytics, University of Delaware, Newark, DE, USA

Correspondence: Arash Riasi, Alfred Lerner College of Business & Economics, Amstel Ave., Newark, DE, 19716, USA. Tel: 1-(302)-898-6249. E-mail: riasi@udel.edu

Abstract

This study intends to combine the fuzzy c-means clustering and genetic algorithms to cluster the customers of steel industry. The customers were divided into two clusters by using the variables of the LRFM (length, recency, frequency, monetary value) model. Results indicated that customers belonging to the first cluster had a higher length of the relationship, recency of trade, and frequency of trade but lower monetary value compared to the average values of these criteria for all customers. The results also showed that customers belonging to the second cluster had a higher recency of trade and monetary value but lower length of the relationship and frequency of trade compared to the average values of these criteria for all customers. It was also found that the combined algorithm (i.e., fuzzy c-means clustering and genetic algorithm) used in this study had a lower mean squared error (MSE) compared to fuzzy c-means clustering.

Keywords: customer clusterin, fuzzy c-means clustering, genetic algorithms, LRFM model, steel industry, customer relationship management

1. Introduction

Identifying customers and customers' needs is an important factor which helps service providers and producers to gain competitive advantage in providing their products/services to the customers. Managers should prioritize their customers and focus on key customers and they should also try to gain a better intuition regarding the costs associated with losing their customers on a daily basis. This is important because when the customers discontinue doing business with us and move to our competitors, various negative consequences such as losing current revenues due to discontinued business relations or losing good reputation and credibility are expected (Tarokh & Sharifian, 2010). Our current customers will probably share their negative experiences with prospective customers. This loss of credibility can lead to a loss of current and potential customers' trust in our products/services. Experts suggest firms to design policies and competitive strategies, therefore organizations cannot disregard their fundamental goals including the achievement of competitive advantage. Identifying different groups of customers and their needs can lead to customer satisfaction, which in turn contributes to customer loyalty. In the long term, it is more beneficial to identify key customers and retain them rather than acquiring new customers to fill the empty place of those who have decided to discontinue doing business with the organization. This is mainly because the cost of new key customer acquisition is five times more than the cost of current customer retention (Griffin & Lowenstein, 2001). Companies have a higher probability of success in selling their products/services to their current customers compared to their prospective customers; the probability of successfully selling a product/service to current active customers is roughly 60-70 percent, while this probability is only 5-20 percent for prospective customers (Tarokh & Sharifian, 2010).

Classification is one of the most important topics in customer relationship management (CRM). In customer classification, the entire population of customers is divided into smaller groups, such that customers in the same group have similar characteristics. Ideally, organizations should have a good understanding about all of their customers, but this is not feasible in real world. Customer classification and clustering enable the firms to group similar customers together and help managers to better understand the customers' needs; because it is much easier to identify and analyze the characteristics of groups of customers rather than studying each customer individually. Data mining has various tools for customer classification. One of the most famous clustering

techniques is fuzzy c-means clustering which has problems with high dimensional data sets and a large number of prototypes (Winkler et al., 2012). Additionally, the performance of fuzzy c-means algorithm is strongly affected by the selection of the initial centroid clusters (Arnaldo & Bedregal, 2013). Therefore, in order to obtain better clusters, this study combines genetic algorithms and fuzzy c-means clustering. The main goal of this study is to use this combined clustering method to divide customers of the steel industry into two different clusters.

2. Research Question and Problem Statement

Effective relationship marketing is an art in today's business environment. Constructing relations with clients is an important skill which is required in order to retain the key customers. Classical marketing theories mostly focused on trade and did not consider customer retention as an important aspect of firms' marketing strategies. But as the markets became more competitive and saturated, and as population combinations of different areas changed over time, companies realized that they are no longer facing a growing economic system with growing markets. In order to survive in these competitive markets, firms need to gain competitive advantage (Amiri Aghdaie et al., 2012; Riasi, 2015a, 2015b). One of the most important determinants of competitive advantage is demand condition (Porter, 1990; Porter 1991) meaning that firms should focus on retaining their customers in order to ensure that their business is economically viable (Riasi &Amiri Aghdaie, 2013; Riasi & Pourmiri, 2015). Nowadays, every single customer has a special value because firms are struggling to obtain a larger share of their fixed or diminishing markets. As a result of this, the cost of new customer acquisition has been significantly increased. Companies have now realized that losing a customer does not simply mean that they are unable to sell a single item to that customer at a specific time, rather it means that they will lose all future purchases of that customer during his/her entire life (Kotler, 1994). Krasnikov et al. (2009) examined the impact of CRM on cost efficiency and profit efficiency and found that although CRM implementation leads to a decline in cost efficiency, it is associated with an increase in profit efficiency. The impact of CRM on long term profitability of the firms is very important, because the main objective of most firms is to maximize the value delivered to the shareholders. In order to maximize the shareholder value it is required to measure the value which is generated from different customers during a specific period and to identify the customers or groups of customers that contribute the most to the firm's value. After identifying these types of customers the firm should motivate them to establish long term relations. As a result of these actions, the customer loyalty will be enhanced, customer life-cycle will be optimized, and eventually the firm will become more profitable. In addition to these benefits, customer satisfaction and relationship commitment have a positive impact on firm's brand loyalty and brand awareness (Kim et al., 2008). It is clear that a firm's relation with its customers can affect its prosperity and it should be considered as an important issue when designing customer retention strategies. By dividing their customers into different clusters, firms can better decide how to effectively allocate their limited resources to different groups of customers based on their value. Also, by using customer clustering techniques, firms can effectively design their customer retention strategies and maximize their overall profitability.

Organizations collect tons of data about their customers, suppliers, and their business partners, but their inability to discover the latent knowledge which is inherent in these valuable data can make the whole data collection process useless. Many business owners are willing to discover this untapped knowledge from their data sets in order to increase their profitability. Data mining techniques including clustering methods enable the firms to exploit this hidden knowledge. Using clustering techniques, customers are divided into homogenous clusters in which customers with similar needs and characteristics are grouped together (Ghazanfari et al., 2010). After identifying the needs and values of its customers, a company must provide valuable products and services to its clients in order to increase their satisfaction and to ensure that they will remain loyal to the firm (Moslehi et al., 2012). But the main question is how to identify the key clients of the organization and how to analyze their behavioral attributes?

3. Literature Review

Various studies have been performed regarding customer segmentation and classification. Ghazanfari et al. (2010) studied the concept of customer segmentation in apparel industry using clustering algorithms. The goal of their study was to classify the countries based on the value of Iran's apparel exports during a 14-year period (1992-2006). To analyze the clusters and to measure their value, and also in order to design an appropriate strategy for each segment, they used RFM (recency, frequency, and monetary) model. RFM is a renowned model for customer value analysis and it has been used by many researchers to perform customer segmentation (Cheng & Chen, 2009; Coussement et al., 2014; Kafashpour et al., 2012; Khajvand & Tarokh, 2011; McCarty and Hastak, 2007; Michopoulou & Buhalis, 2008; Rastegar, 2009; Tao & Yeh, 2003; Wu et al.; 2009). RFM analyzes the behavior of the customers and therefore it is considered as a behavior-based model (Wei et al, 2012; Yeh et al., 2008). Hanafizadeh and Rastkhiz Paydar (2011) used neural networks to classify the customers of auto

insurance into 4 different groups based on their risk. The goal of their study was to use these clusters to identify the appropriate insurance premium for each group of customers. Wei et al. (2012) adopted self-organizing maps (SOM) technique to develop an extended RFM model, called LRFM (length, recency, frequency, and monetary) model for a children's dental clinic in Taiwan to segment its dental patients. They recommended 12 clusters and labeled the 3 clusters having the above average LRFM values as core patients. SOM is considered as a neural network technique and can be used for clustering problems, visualization, and market screening (Fish & Ruby, 2009; Hanafizadeh & Mirzazadeh, 2011; Hsu et al., 2009; Hung & Tsai, 2008; Kiang & Fisher, 2008; Mostafa & El-Masry, 2013; Wang, 2001; Wei et al., 2012). The advantage of SOM is that it can automatically detect strong features in large data sets (Wei et al., 2012). Gholamian and Niknam (2012) proposed an adjusted model for clustering customers in retail banking industry using RFM model. They tested their model on checking accounts of customers of a private bank and compared their results with the benchmark model. Chang and Tsay (2004) used LRFM model to increase the length of communication with customers. After extracting the data, they classified the customers into 16 groups. They believe that adding additional factors to the model, helps to identify loyal customers that provide the firm with higher profits. Li et al. (2011) used LRFM model to classify customers of textile industry into different clusters. Their results indicated that factories with domestic sales had more long-term and loyal customers, even though monetary value and frequency were not high. RFM model is unable to divide customers into different segments based on the length of the relationship with the company (Reinartz & Kumar, 2000). Additionally, the customer loyalty is dependent on both the type of the relationship that exists between the customer and the company (Chow & Holden, 1997) and the length of this relationship (Chang & Tsay, 2004). Therefore, in the current study, the LRFM model which is the extended version of RFM model is used to take the length of the relationship into account.

4. Methodology

This study can be classified as an applied research based on its objective and as a descriptive study based on its data collection method. The statistical society of this research included the managers and experts of steel industry. Since the size of the statistical society was known and all the measures were qualitative, the sample size was calculated to be 120 using the Cochran's formula.

4.1 Genetic Algorithms

Since clustering is an NP-hard problem, heuristic algorithms are used for data clustering. The most notable type of these algorithms is genetic algorithm, which has a strong capability in parallel implementations and has a low probability of being trapped in local minimums. Genetic algorithms became popular through the work of Holland (1975), and they are considered as well-developed search heuristics that mimic the natural selection mechanism. Genetic algorithms are iterative processes based on evolutions which initiate from a population of randomly generated individuals with the ultimate goal of finding comprehensive optimized solutions (Holland, 1975). The flowchart of genetic algorithm (Shyr, 2010) is depicted in Figure 1.
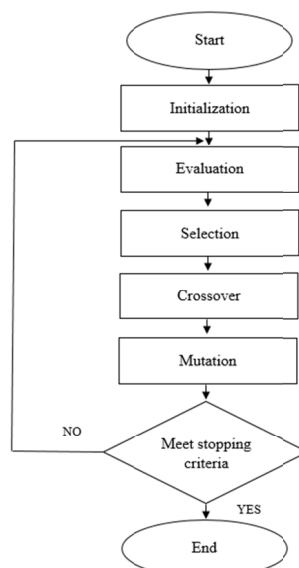

Figure 1. Genetic algorithm flowchart

4.2 Fuzzy C-Means Clustering

In classical clustering each input data point belongs to exactly one cluster and cannot be part of two or more clusters; in other words, the clusters do not overlap each other. Imagine that there is a data point which has characteristics similar to two or more clusters, in classical clustering we have to decide which cluster is more suitable for this data point and allocate it to only a single cluster. The main difference between classical and fuzzy clustering is that in fuzzy clustering a data point can belong to more than one cluster. Similar to classical c-means algorithm, in the fuzzy c-means algorithm the number of clusters is also known beforehand. The objective function of this algorithm is as follows:

$$J = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{m} d_{ik}^{2} = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{m} \left\| x_k - v_i \right\|^2 \tag{1}$$

Where m is a real number larger than 1, x_k is the kth data point, v_i is the centroid of the ith cluster, u_{ik} is the degree to which data point k belongs to cluster i, and $\|x_k - v_i\|$ is the Euclidean distance between k^{th} data point and i^{th} cluster center. Using u_{ik} a matrix U with c rows and n columns can be defined, where each element of the matrix can have a value between 0 and 1. If all the elements of matrix U are either 0 or 1 then the algorithm is similar to classical c-means. Although the elements of matrix U can have any value between 0 and 1, but the elements in each column should sum to 1. In other words:

$$\sum_{i=1}^{c} u_{ik} = 1 \quad , \quad \forall k = 1, \dots, n \tag{2}$$

This indicates that the sum of the proportions that each data point belongs to each of the c different clusters should be equal to 1. Using the above condition and by minimizing the objective function we will have:

$$V_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} , u_{ik} = \frac{1}{\sum_{j=1}^{c} (\frac{d_{ik}}{d_{jk}})^{2/(m-1)}} \tag{3}$$

5. Data Analysis and Results

In order to perform the clustering, the data from 120 customers were collected and normalized. The data included four different variables, namely, length of the relationship, recency of trade, frequency of trade, and monetary value. In order to perform the fuzzy clustering, GA-Fuzzy Clustering software was used. This software performs the clustering by combining fuzzy c-means clustering and genetic algorithm. The parameters of these two algorithms are displayed in Table 1.

Table 1. Parameters of fuzzy c-means clustering and genetic algorithm

Algorithm	Parameter	Value
Fuzzy c-means clustering	Number of clusters	2
	Stop parameter	0.1
Genetic algorithm	Size of initial population	10
	Mutation frequency	0.5
	Crossover frequency	1
	Crossover percentage	0.5

After running the software, the customers were divided into two clusters. The cluster centers for each criterion were calculated as shown in Table 2. According to Table 2, customers belonging to the first cluster had a higher length of the relationship, recency of trade, and frequency of trade but lower monetary value compared to the average values of these criteria for all customers. Customers belonging to the second cluster had a higher recency of trade and monetary value but lower length of the relationship and frequency of trade compared to the average values of these criteria for all customers.

Table 2. Cluster centers for each criterion

Cluster	Length of the relationship (L)	Recency of trade (R)	Frequency of trade (F)	Monetary value (M)
First cluster	↑ 0.8347	↑ 0.7945	↑ 0.6531	↓ 0.0212
Second cluster	↓ 0.1352	↑ 0.4634	↓ 0.0319	↑ 0.7049

In order to compare the efficiency of the combined algorithm (i.e., fuzzy c-means clustering and genetic algorithm) with fuzzy c-means clustering, mean squared error (MSE) and run time criteria were used. Table 3 shows the MSE and run time for each of these two algorithms. According to table 3, the combined algorithm that was used in our study had a lower MSE but a higher run time compared to fuzzy c-means clustering algorithm. However, it is clear that in performing customer clustering, the MSE criteria is relatively more important than run time, because customer clustering is usually a time consuming process which needs to be done with high accuracy.

Table 3. MSE and run time comparison

Algorithm	MSE	Run time (s)
Fuzzy c-means clustering	0.0875	0.0546
Combined algorithm (fuzzy c-means clustering and genetic algorithm)	0.0152	0.8474

After running the combined algorithm, the customers were divided into two different clusters as shown in figure 2.
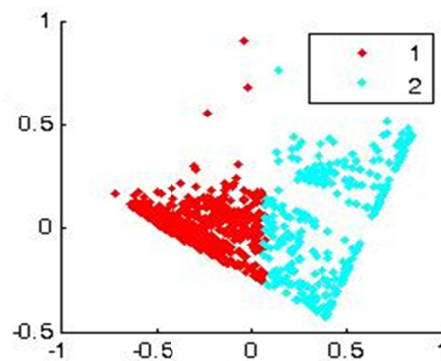

Figure 2. Clusters obtained from the combined algorithm

6. Discussion and Conclusions

As the importance of customer satisfaction became more evident in business environment, many organizations started to focus on customer identification, customer loyalty, and profitability analysis in order to increase their market share and to enhance the level of customer satisfaction. Appropriate customer relationship management is considered as a competitive advantage for organizations, and it is obvious that clustering is one of the most important issues in CRM. Therefore, the goal of this study was to use an appropriate model for clustering the customers of steel industry based on length of the relationship, recency of trade, frequency of trade, and monetary value criteria. As mentioned earlier, various algorithms can be used for data clustering including but not limited to fuzzy c-means algorithm. Although fuzzy c-means clustering is relatively easy to implement and has a low run time, its efficiency is not very high due to its sensitivity to initial centroid values and the possibility of being trapped into local optimum. The known deficiencies of classical clustering techniques, accentuates the need for an analytical clustering method which is both capable of providing optimal clusters and is expandable to complicated systems modelling. In this study, a combined algorithm which was a combination of fuzzy c-means clustering and genetic algorithms was used in order to analyze the data and to perform the

customer clustering. This combined algorithm was used because it had a higher efficiency and it did not have the deficiencies associated with classical clustering algorithms.

The customers of steel industry were divided into two clusters according to the variables of the LRFM model. The results revealed that customers belonging to the first cluster had a higher length of the relationship, recency of trade, and frequency of trade but lower monetary value compared to the average values of these criteria for all customers. On the other hand, the results indicated that customers belonging to the second cluster had a higher recency of trade and monetary value but lower length of the relationship and frequency of trade compared to the average values of these criteria for all customers. Therefore, based on the findings of Chang and Tsay (2004), the customers in the first cluster can be considered as loyal customers according to loyalty measures and they can be considered as uncertain customers based on value measures. Also the customers in the second cluster can be considered as newcomers based on the loyalty matrix and they can be considered as uncertain customers based on the value matrix. Additionally, the findings revealed that the combined clustering algorithm had a lower MSE and a higher run time compared to the fuzzy c-means algorithm. Since the accuracy is more important than speed in customer clustering, the authors suggest that future studies use a combination of fuzzy c-means clustering and genetic algorithm in order to obtain the most accurate clusters. This study clearly contributes to the literature in this field by showing that a combination of fuzzy c-means clustering and genetic algorithms is an efficient way to perform customer clustering in steel industry.

References

Amiri Aghdaie, S. F., Seidi, M., & Riasi, A. (2012). Identifying the Barriers to Iran's Saffron Exportby Using Porter's Diamond Model. *International Journal of Marketing Studies*, *4*(5), 129-138. http://dx.doi.org/10.5539/ijms.v4n5p129

Arnaldo, H. A., & Bedregal, B. R. C. (2013). A new way to obtain the initial centroid clusters in Fuzzy C-Means algorithm. *2013 2nd Workshop-School on Theoretical Computer Science (WEIT) - IEEE,* 139-144.

Chang, H., & Tsay, S. (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.

Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert Systems with Applications*, *36*(3), 4176-4184.

Chow, S., & Holden, R. (1997). Toward an understanding of loyalty: The moderating role of trust. *Journal of Managerial Issues*, *9*(3), 275-298.

Coussement, K., Van den Bossche, F. A., & De Bock, K. W. (2014). Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees. *Journal of Business Research*, *67*(1), 2751-2758. http://dx.doi.org/10.1016/j.jbusres.2012.09.024

Fish, K., & Ruby, P. (2009). An artificial intelligence foreign market screening method for small businesses. *International Journal of Entrepreneurship*, *13*, 65-81.

Ghazanfari, M., Malek Mohammadi, S., & Alizadeh, S. (2010). Customer Segmentation in Clothing Exports Based on Clustering Algorithm. *Iranian Business Studies Quarterly*, *14*(56), 59-86.

Gholamian, M. R., & Niknam, Z. (2012). Providing an adjusted model for clustering bank customers based on their life cycle value. *Iranian Journal of Executive Management*, *4*(7), 59-76.

Griffin, J., & Lowenstein, M. W. (2001). *Customer winback: How to recapture lost customers and keep them loyal*. San Francisco, CA: Jossey-Bass.

Hanafizadeh, P., & Mirzazadeh, M. (2011). Visualizing market segmentation using self-organizing maps and Fuzzy Delphi method–ADSL market of a telecommunication company. *Expert Systems with Applications*, *38*(1), 198-205.

Hanafizadeh, P., & Rastkhiz Paydar, N. (2011). A model for risk-based clustering of auto insurance customers based on their risk using data mining techniques (Case study: auto insurance in an insurance company). *Iranian Insurance Research (Insurance Industry)*, *26*(2), 55-81.

Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press.

Hsu, S. H., Hsieh, J. P. A., Chih, T. C., & Hsu, K. C. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications*, *36*(4), 7947-7951.

Hung, C., & Tsai, C. F. (2008). Market segmentation based on hierarchical self-organizing map for markets of

multimedia on demand. *Expert Systems with Applications*, *34*(1), 780-787.

Kafashpour, A., Tavakoli, A., & Alizadeh, A. (2012). Clustering customers based on the life cycle value using data mining techniques and RFM model. *Iranian Journal of Public Management Research*, *5*(15), 63-84.

Khajvand, M., & Tarokh, M. J. (2011). Estimating customer future value of different customer segments based on adapted RFM model in retail banking context. *Procedia Computer Science*, *3*, 1327-1332.

Kiang, M. Y., & Fisher, D. M. (2008). Selecting the right MBA schools–An application of self-organizing map networks. *Expert Systems with Applications*, *35*(3), 946-955.

Kim, K. H., Kim, K. S., Kim, D. Y., Kim, J. H., & Kang, S. H. (2008). Brand equity in hospital marketing. *Journal of Business Research*, *61*(1), 75-82.

Kotler, P. (1994). *Marketing Management: Analysis, Planning, Implementation, and Control*. Upper Saddle River, NJ: Prentice Hall.

Krasnikov, A., Jayachandran, S., & Kumar, V. (2009). The impact of customer relationship management implementation on cost and profit efficiencies: evidence from the US commercial banking industry. *Journal of Marketing*, *73*(6), 61-76.

Li, D. C., Dai, W. L., & 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.

McCarty, J. A., & Hastak, M. (2007). Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of Business Research*, *60*(6), 656-662.

Michopoulou, E., & Buhalis, D. (2008). Performance measures of net-enabled hypercompetitive industries: The case of tourism. *International Journal of Information Management*, *28*(3), 168-180.

Moslehi, N. (2012). *Using LRFM model for clustering customers based on their life cycles in order to improve customer relationship management* (Master's thesis). Ferdowsi University of Mashhad, Mashhad, Iran.

Mostafa, M. M., & El-Masry, A. A. (2013). Citizens as consumers: Profiling e-government services' users in Egypt via data mining techniques. *International Journal of Information Management*, *33*(4), 627-641.

Porter, M. E. (1990). *The Competitive Advantage of Nations*. New York, NY: The Free Press.

Porter, M. E. (1991). *Canada at the crossroads: The reality of a new competitive environment*. Ottawa: Business Council on National Issues and Minister of Supply and Services of the Government of Canada.

Rastegar, N. (2009). *Presenting the developed RFM model for clustering customers of banking services*. Paper presented at the first international conference of banking services marketing, Tehran, Iran. Retrieved January 15, 2016, from http://www.ensani.ir/fa/14822/magazine.aspx

Reinartz, W. J., & Kumar, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *Journal of Marketing*, *64*(4), 17-35.

Riasi, A. (2015a). Barriers to international supply chain management in Iranian flower industry. *Management Science Letters*, *5*(4), 363-368. http://dx.doi.org/10.5267/j.msl.2015.2.005

Riasi, A. (2015b). Competitive Advantages of Shadow Banking Industry: An Analysis Using Porter Diamond Model. *Business Management and Strategy*, *6*(2), 15-27.

Riasi, A., & Amiri Aghdaie, S. F. (2013). Effects of a Hypothetical Iranian Accession to the World Trade Organization on Iran's Flower Industry. *Consilience: The Journal of Sustainable Development, 10*(1), 99-110. DOI: 10.7916/D8HQ3ZK8

Riasi, A., & Pourmiri, S. (2015). Effects of online marketing on Iranian ecotourism industry: Economic, sociological, and cultural aspects. *Management Science Letters*, *5*(10), 915-926.

Shyr, W. J. (2010). Parameters Determination for Optimum Design by Evolutionary Algorithm. In M. Crisan (Ed.), *Convergence and Hybrid Information Technologies*. Intech. Retrieved January 15, 2016, from http://www.intechopen.com/books/convergence-and-hybrid-information-technologies/parameters-determination-for-optimum-design-by-evolutionary-algorithm

Tao, Y. H., & Yeh, C. C. R. (2003). Simple database marketing tools in customer analysis and retention. *International Journal of Information Management*, *23*(4), 291-301.

Tarokh, M., & Sharifian, K. (2010). Applications of data mining in improving customer communication

management. *Iranian Industrial Management Studies Quarterly*, *6*(17), 153-181.

Wang, S. (2001). Cluster analysis using a validated self‐organizing method: Cases of problem identification. *Intelligent Systems in Accounting, Finance and Management*, *10*(2), 127-138.

Wei, J. T., Lin, S. Y., Weng, C. C., & Wu, H. H. (2012). A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Systems with Applications*, *39*(5), 5529-5533.

Winkler, R., Klawonn, F., & Kruse, R. (2012). Problems of fuzzy c-means clustering and similar algorithms with high dimensional data sets. In W. A. Gaul, A. Geyer-Schulz, L. Schmidt-Thieme, & Y. Kunze (Eds.), *Challenges at the Interface of Data Analysis, Computer Science, and Optimization* (pp. 79-87). Heidelberg, Germany: Springer Berlin Heidelberg.

Wu, H. H., Chang, E. C., & Lo, C. F. (2009). Applying RFM model and K-means method in customer value analysis of an outfitter. In S. Y. Chou, A. Trappey, J. Pokojski, & S. Smith (Eds.), *Global Perspective for Competitive Enterprise, Economy and Ecology* (pp. 665-672). London, U.K.: Springer London.

Yeh, I. C., Yang, K. J., & Ting, T. M. (2009). Knowledge discovery on RFM model using Bernoulli sequence. *Expert Systems with Applications*, *36*(3), 5866-5871.

Copyrights