2528-9705

Örgütsel Davranış Araştırmaları Dergisi
Journal Of Organizational Behavior Research
Cilt / Vol.: 3, Sayı / Is.: S2, Yıl/Year: 2018, Kod/ID: 81S229

AN ANALYSIS OF THE BEHAVIOR OF TEJARAT BANK E-BANKING CUSTOMERS USING DATA MINING BASED ON RFM MODEL

Hossein ASADI ASADABAD[1], Fardideddin ALLAMEH HAYERI[2]*

[1] Department of Management, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran,
[2] Department of Management, Nehjir Mobarekeh Branch, Islamic Azad University, Mobarekeh, Iran.

***Corresponding Author:**
Email: f. haeri@mau.ac.ir

ABSTRACT

Electronic banking plays a central role in the field of electronic payment. The aim of this study is to analyze the behavior of Tejarat Bank e-banking customers using data mining based on RFM model. This study has analyzed and examined e-banking customers in 8 clusters through the analysis of 1700 data and determining the optimal number of clusters by two-step algorithm and determining 8 optimal clusters by the algorithm and using k-mean algorithm. The results show that the cluster which has lower R (exchange recency), lower F (the number of exchange frequency) and higher M (Monetary value) has more loyal customers. The cluster which has lower R is placed in the first cluster rank. In spite of RFM variables, gender, age and education has been found to be effective on clustering. Results of clustering contains 8 clusters: 1) men with the age range of 30 to 40 years old with Bachelor's degree; 2) men with the age range of 30 to 40 years old with a bachelor's degree; 3) women under 30 years old with high school diploma; 4) men with the age range of 40 to 50 years old with high school diploma; 5) men with the age range of 40 to 50 years old with bachelor's degree; 6) men over 50 years old with high school diploma; 7) women under the age of 30 years old with higher education; 8) men under 30 years old with Bachelor's degree.

Keywords: *Electronic Banking, Consumer Behavior Analysis, Data Mining, K-Means Algorithm, RFM Model.*

INTRODUCTION

The transformation of the world through use of information technology, the Internet, and consequently e-government and e-commerce, which in turn is derived from information technology, has created a deep and profound transformation in communication and information transfer processes (Hasanzadeh et al. 2012). In information and communication technology, the issue of interaction is of great importance due to time and cost saving and the growing importance of data sharing (Abedi Jafari et al. 2010; Haghighi Nasab & Khosravi 2011; Jamali, and Hashemi 2012; Zahedi, 2011). The Internet has expanded horizons for businesses worldwide, especially e-banking (Babaei & Ahadi 2010; Alsajjan & Dennis 2010; Gikandi and Bloor, 2010). With the rapid growth of information and communication technology, e-banking plays a central role in electronic payment (Jayawardhena and Foley, 2000; Liao and Cheung, 2001; Sohail Sadiq and Shanmugham, 2003) which provides online transaction to support many e-commerce applications, including e-shopping, electronic auctions, Internet stock trading and many more (Moghaddasi, 2010). In fact, virtual space has changed the role of consumers in the world of commerce. The undeniable advantages of information technology in increasing the

Geliş tarihi/Recieved: 10.11.2017 – Kabul tarihi/Accepted: 09.02.2018 – Yayın tarihi/Published: 21.08.2018

accuracy and speed of the flow of affairs, increasing global quality, reducing costs and more customer satisfaction has led organizations to quickly deploy and use information systems. One of the most important ways to gain competitive advantage for today's banks is the use of information technology to provide banking services known as e-banking services. In addition, the entry of information technology and specifically the Internet in the banking industry has changed the competitive environment of this industry. Due to the widespread changes in global markets and the growing intense competition, the experience of customer interaction in global level and online has become a distinct strategy. In fact, electronic banking is a way to reduce costs and stay competitive in comparison with traditional banking. This research aims to use information technology in order to gain competitive advantage for banks in the field of electronic banking. The use of information technology in electronic banking has led to the creation of a large amount of customer data. Three approaches to design include data mining in banking, data mining in marketing and marketing in banking. In these approaches, data mining can be used as an information technology in electronic banking to explore knowledge in a large amount of customer data. Banks can identify, attract, maintain and develop their customers by analyzing their behavior. In this regard, one of the most important actions of banks is clustering customers into different behavioral clusters in which, it is possible to formulate marketing strategies suitable to each cluster. Customer behavior analysis was performed through K-mean data mining in different industries (Coates and Ng, 2012; Macqueen, 1967; Vattani 2011). Nevertheless, in other researches, demographic variables were not considered as the main variables. This analysis has not been done using the RFM model in Iran, specifically for banking customers. This research is trying to cluster customers together considering the RFM variables and demographic variables.

RESEARCH THEORETICAL FUNDAMENTALS

Since the use of information technology in e-banking has resulted in large amount of data, as mentioned, data mining in banking is the approach to be designed. Analysis of customers' behavior in a variety of industries, such as banks, which face with a large number of customers with different natures and behaviors, is the purpose of marketers and managers of the banking industry. Using clustering approach as a technique for data mining in these industries in order to divide heterogeneous customers into homogeneous groups with similar transaction behaviors (Deshpande and Thakare, 2010; Hand et al. 2001), will help better understanding of customer behavior. Subsequently, developing an appropriate marketing strategy and effective communication with each customer cluster is useful in satisfying and retaining customers, attracting new customers and finally, surviving the organization.

Today, there is a huge amount of information from the customers' transaction including customer's profile, transaction frequency, and transaction type and volume in the bank database. In addition, it has been proved to marketers of these large economic firms that they can benefit from these data and information in all aspects of customer relationship management, CRM (Buttle, 2004). Today, CRM have an enviable status in the process of the management of banking institutions (Rouholamini and Venkatesh 2011). Due to the large amount of available transaction data from customers in the database of the bank to be studied and lack of proper use of these data by managers and banking experts, the researchers decided to use this database to identify and analyze customers behavior in order to provide special services or specific facilities to different customer groups and divide them into homogeneous groups that are measured according to similar variables and criteria per customer. Since the purpose of this research is data mining of electronic banking data, it is therefore important to examine its concepts as well as its related indicators. In this chapter, the concepts related to research are described at first, and in the end, studies and researches carried out in this field will be examined.

Electronic banking

With the emergence of the Internet, the concept of electronic banking was formed in 1991, meaning that customers, without having to attend branch, can do their banking affairs by referring to the electronic environment (Hashemian, 2012). E-banking technology represents a variety of different services (Kolodinsky et al., 2004; Lee 2009). It has brought many benefits, such as lack of temporal and spatial limitation, easy access to information, reduced cost of service, and customers time saving, which has led to the rapid growth of the use of electronic banking services (Divandari, 2013). The use of electronic banking services is one of the solutions to gain competitive advantage for banks and has led to the emergence of close competition in this regard; in this case, the level of customer expectations for such services has also increased (Rasouli and Manian, 2012). Electronic banking includes important benefits, such as focusing on new distribution channels, providing reformed services to customers and the use of e-commerce strategies (Azizi Sarkhoni, 2008).

E-services have increased the ability of companies to provide better service and they are divided into four sub-categories according to the customer's view toward them: Internet channels, card channels (ATMs, sales terminals), telephone channels (call centers, phone bank, mobile bank) and automatic channels, which are virtual bank communication channels. Meanwhile, the use of Internet banking is increased rapidly due to the expansion of Internet penetration. Using the Internet, the time and geographical limitations are eliminated. Customers can access their accounts 24 hours a day and 7 days a week (Karjaluto et al, 2002).

In Iran, there are six channels for providing e-banking services that are as follows. The first three channels are based on the card and for any transaction, the physical presence of the card is inevitable.

1-ATMs 2- Sales terminals (POS) 3- Branches terminals (Pni- Pad) 4- Phone Bank 5. Mobile Bank 6-Internet Banking or Internet Bank.

Using these channels, you can use various banking services in the form of e-banking, such as transferring funds, checking account balance, paying bills, and so on.

Analysis of customer behavior

Due to the intensive competition in the market and the existence of a variety of products and services available to customers, the correct recognition of customer behavior is the most important dimension of customer relationship management (Berson et al., 2004). Analyzing customer behavior and choosing the appropriate marketing method based on these analyses is considered as a very important factor for the survival of the firms (Caceres 2007). Different researches in the field of marketing have shown that in the marketing of services, there is no need to serve all customers similarly. The rapid growth of information technology in various business segments and the accumulation of a large amount of customer data, have made the correct recognition of customers, understanding their behavioral patterns and responding to their needs more and more difficult and even impossible. Therefore, the need for methods of customer behavior analysis based on the existing data and the development of appropriate approaches for marketing and communication with customers to satisfy, preserve and attract them is felt more than ever.

Customer behavior analysis became commonplace in the early 1980s and was considered in four dimensions: customer identification, customer attraction, customer retention, and customer development (Seyyed Hosseini and Gholamian, 2010). Customer behavior analysis is a structure that expands the value of the customer, and provides a powerful tool for retaining valuable customers (Bose & Sugumaran, 2003). In order to obtain this, analyzing past and current customer behavior will help understand the characteristics of the current customers of the organization. By employing data mining approaches in customer data, organizations can understand hidden information in the data, behavioral patterns, and customer needs, and accordingly, they can better use their resources to meet customer needs, provide new services, and accept customers

RFM model

This model was first introduced by Hughes (1994). He used the RFM to analyze past customer behavior that is easily tracked and accessed. This model uses three dimensions related to customer transaction data to analyze their behavior. Indexes of this model are defined as follows (Chen & Cheng, 2009):

1) Recency: This index refers to the time interval between the last purchase made by the customer to the end of the particular period (end of the time period examined). The lower distance indicates the high value of this index in the model.
2) Frequency: This index represents the number of exchanges a customer has made in a particular period. The higher the number of exchanges indicates the higher value of this index in the model.
3) Monetary: This index shows the amount of money a customer has spent in a particular period for exchanges. The greater the amount of money spent indicates the high value of this index in the model.

In RFM model, the value of each customer's life cycle is obtained from the sum of the values derived from RFM index (Razmi and Ghanbari 2009). So in this model, it is assumed that customers with high value for each index of the model are the best customers, as long as they behave in the future as they behaved in the past (Fader et al., 2005).

The higher F and R, the more likely it is to make a new transaction with the customer, and if M is also higher, the likelihood of customer's return to purchase is higher. In RFM model, it is assumed that customers which have higher value in each variable of the model are the best customers, as long as they will behave in the future as they behaved the past, in this case, it is believed that these customers have higher profits than others for the firm (Wei et al. 2012). Therefore, the basic assumption of the model is that future patterns of exchange and customer purchase are similar to past and present patterns, and it is emphasized on the ease of calculation and understandability and the ability of RFM to predict customer's future behavior and loyalty. In addition, RFM model is used as a way to measure the value of customer longevity (Baradaran, 2013).

RFM was originally used by marketers in B2C markets, especially in industries such as telecommunication banking and so on. There are different approaches for ranking or scoring in the model, in the following three common approaches are referred to (Ebadi, 2015):

Initial Ranking - Cluster Ranking - Ranking using weighting. Each approach has fans who prefer it to other approaches. It is important to choose the best approach according to the type of organization and its customers.

Data mining

The history of discovery of knowledge from the databases is not related to long ago, and today it is known as data mining. Data mining is a research trend that is rapidly growing and many researches have been done on it (Fayyad and Uthurusamy, 1996; Sami Zadeh, 2007; Taqavi Fard et al., 2007; Vercellis, 2009). Researchers and programmers from different fields are collaborators in the art of data mining, so it is difficult to provide a comprehensive overview of data mining techniques. The term "knowledge discovery" was first introduced in the 1990s and attracted researchers' attention to data mining algorithms. The aim of data mining is to discover new, valid and consistent knowledge using intelligence and statistical tools in a large amount of data (Marban et al., 2009). Data mining is the extraction or adaptation of knowledge from a data set and is referred to as the process that extracts knowledge from data, and this knowledge is expressed in terms of patterns and models (Taghavi Fard, 2012). It is worth mentioning that data mining comes from a combination of several disciplines. Statistics, machine learning, optimization methods, pattern identification and recognition methods, database, visualization, neural networks, mathematical models, information retrieval, genetic algorithm and artificial intelligence techniques that data mining uses (Mahmudi, 2013).

The life cycle of a data mining project involves six steps: understanding business issues, understanding data, preparing data, modeling, evaluating results, and applying a model (Azar et al., 2010).

Different techniques of data mining can be divided into "predictive" and "explanatory" types based on the type of operation they perform. Predictive techniques, by constructing a database model, are responsible for predicting unknown cases. Since the explanatory techniques discover understandable patterns of data for human beings. The purpose of the classification is to identify the characteristics by which different classes can be distinguished from each other. Classification in the data mining takes place in two steps. First, different classes are identified from old data, and then the incurrence of new data in existing classes are predicted. Classification is a learning technique with observer, because it classifies new data using a training dataset (as a guide). This method is also considered as a predictive method (Turbun, 2011). Grouping is considered as an explanatory method. This method, with the thought of dividing and solving, divides the data into a large system and divides them into smaller components. A grouping is appropriate when the data objects within each group are very similar to each other and differ from objects in other groups. The criterion of similarity and difference between data objects are determined by a distance function. (Momeni, 2011).

Clustering

Clustering, can be considered as the most important issue in uncontrolled learning. It has become a common method for the marketing research (Punj and Stewart, 1983). Clustering is finding a structure inside a collection of unlabeled data (Forcht and Cochran, 1999). The cluster is referred to as a set of data that is similar to each other. In clustering, it is tried to divide the data into clusters that maximize the similarity between the data within each cluster and minimize the similarity between the data within different clusters (Momeni, 2011).

Clustering by K-mean:

This method is a basic method for many other clustering methods. It is a monolithic and flat method. In a simple kind of this method, first, according to the number of clusters needed, some points are randomly selected. Then, according to the degree of proximity or similarity, they are attributed to one of these clusters and thus new clusters are obtained. By repeating the same procedure, it is possible to calculate new centers by averaging the data in each iteration and reassign the data to the new clusters. This process continues until there is no longer any change in the data (Amorim, 2012). The algorithm below is the basic algorithm for this method:

1. At first, the K points are selected as the points of the cluster centers.
2. Each sample is attributed to the cluster whose center has the shortest distance to that data.
3. After assigning all the data to one of the clusters, for each cluster, a new point is considered as a calculation center. (Average points belonging to each cluster).

Steps 2 and 3 are repeated until there is no change in the cluster centers (Celebi, 2013).

In the K-mean method, by specifying cluster centers, the number of clusters are defined previously, it is attempted to allocate observations to each cluster with the aim of minimizing the distance of each observation from the cluster center. The repetitions of this algorithm continue until there are no changes in cluster centers in successive repetitions. In this technique, patterns and the hidden relationships between data sets are determined by minimizing the intervals within the cluster (Momeni, 2011).

Two-step clustering:

This type of clustering, as well as its other types tries to divide data sets in separate groups, while these groups and their characteristics are not clear from the beginning. That is, the target field is not specified. Instead of trying to predict the output, this clustering model tries to find the existing patterns of input data sets. This model is a two-step method. In the first step, a general survey is performed on data, in which raw input data is divided into sub-clusters. In the second step, a hierarchical clustering method is applied to the data that merges sub- clusters to reach larger clusters without needing to re-scan it (Christodoulakis, 2009). The advantage of hierarchical clustering is that at first there is no need to specify the number of clusters. Among the clustering algorithms, a two-step algorithm is more widely considered due to its simplicity, precision and speed.

Literature review

- ***Interior background***

Farrokhi and Teimurpour (2016), in a research titled identifying and classifying Iranian banking system customers in terms of perceived expectations and value of banking services, divided customers into three clusters using data mining techniques, using K-mean. The first cluster involves customers for whom the factors of understanding, knowing, and communication are of great importance; the second cluster involves customers for whom diversity and differentiation are desirable, and the third cluster includes customers that accessibility, efficiency, and cost factors are important for them.

Ebadi and Alizadeh (2015), in a study titled analyzing customer behavior in purchase and sending online group SMS using data mining based on the RFM model, identified six clusters of customers with a case study on 51534 records of a company's data that is active in Online group SMS industry, and provided solutions for customer retention and recruitment. It was a case study of 51534 records from data of an active company in online group SMS. The results indicate that the company has faced serious problems in process of sending SMS during the period of study and has gradually lost its old and most lucrative customers. Although it has succeeded in the SMS acquisition process and has attracted a large number of customers.

Mohammadi et al (2014), by analyzing the problems of branches of Ayandeh Bank all over the country using the data mining method and presenting the RFS model, have clustered all the branches based on the similarity of the factors of R (the recency of the problem), F (frequency or the number of problems), and S (branches satisfaction from the contact center), and found the relationship between the factors with the type of problems declared. The branches were distributed according to their behavioral pattern in four optimal clusters, the results were analyzed and at the end, suggestions were made to improve the contact center's performance.

Radfar et al (2014) classified Internet bank customers using data mining algorithms. They did this by using the decision tree. One of the methods of data mining is decision tree (Kiss, 2003), and if the decision tree is made appropriately, customers can be classified optimally. In this research, a suitable model for customer classification based on the use of Internet banking services is presented. This model is based on the CRISP-DM standard (Chapman et al., 2000) and the required data is extracted from the Sina Internet Bank's database. Among other decision trees, the final decision tree is based on the criteria of efficiency and accuracy, and according to customer classification in three levels of high, medium and low new customers who are applying for Internet bank are predicted.

Khan-Babaei and Zain al-Abedini (2013) examined electronic banking services by the model of using data mining techniques in identifying, segmenting and analyzing customer behavior, and they divided and analyzed the customers into six groups with two approaches of customer segmentation and customer value determination using a demographic dataset and customers transactions. They collected data and expectations of customers by distributing questionnaires among customers and analyzed and segmented these data in order to provide better services.

Fouladifar et al. (2016) in a research entitled market segmentation for Internet banking customers using RFM and K-mean in a private bank in France, clustered customers using a 1478-sample of electronic banking customers and presented marketing strategies. The researchers, performed clustering using demographic and RFM variables, and with respect to demographic variables, they presented strategies for segmentation of the target market.

Hu et al. (2014), in a study entitled discovery of frequent patterns in customer RFM analysis, identified the valuable customers of the company to develop marketing strategies and used the integration of RFM and presented a model for segmentation. This pattern, using frequent patterns in customer analysis, identified valuable customers and, by clustering them, presented a strategy for marketing.

Hua et al (2013), in a research attempted to identify the valuable customers of the company using the RFM model. They used the RFM model to discover pattern in customers data by weighting. They discovered a model that identified the valuable customers of the company and described strategies for maintaining them. Moreover, for the rest of the customers they developed strategies to prove their loyalty to the company.

Wei et al (2013), in a study on customer relationship management in the hairdressing industry, using data mining techniques and K-mean attempted to cluster in RFM model, and finally, they identified four groups of loyal customers, potential customers, new customers and lost customers.

Namvar et al. (2010) in a research using a two-step method to segment the bank customers, extracted 3 clusters using the RFM model in the first step. In the second step, by using demographic variables of age, education, and place of residence they performed internal segmentation of each of the first step clusters and ultimately calculated the value of the lifetime of the customers for the 9 extracted clusters.

Christodoulakis and Aggelis (2009) in their study using the data of Greek banks and RFM variables in customer data, and using the concept of the customer value pyramid, divided the customers into five categories based on clustering, K-mean and two-step algorithms: main customers, regular customers, retailer customers, and inactive customers. They offered strategies for maintaining and retaining the loyalty of high value customers. They retained their main customers by providing appropriate services for them, and presented appropriate strategies for regular and inactive customers.

MATERIALS AND METHODS

Research conceptual model

In RFM model, the value of each customer's life cycle is obtained from the sum of the values derived from the RFM indexes. Demographic variables are also considered for customers. Considering the above issues, the model of this research is as follows:

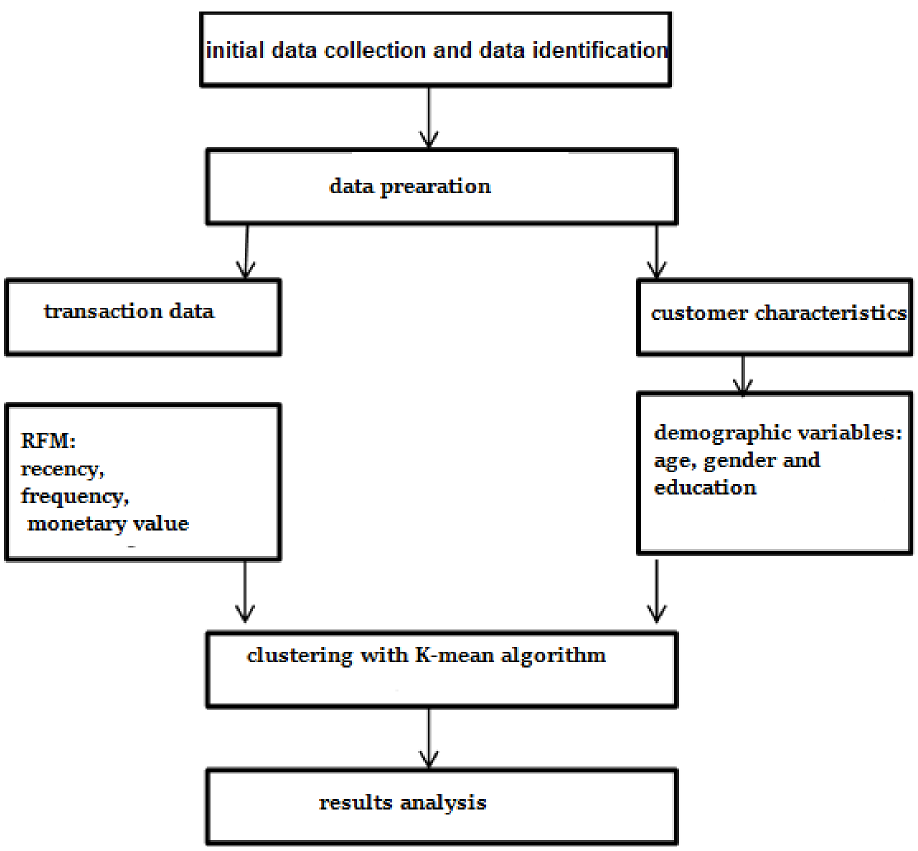

Figure 1: Conceptual model of the research.

Research conceptual model (Fouladifar et al, 2016)
In the first step, the initial data collection is performed. This step involves collecting the initial data, describing the data, discovering the data, and changing the quality of the data. According to the research model and the precise understanding of the research, the background for understanding the data and the background for collecting data from the central database was provided. Initial data was provided to the researcher as the raw data of the customer's account flow. At this point, there is also a series of demographic information.
The second step is data preparation. Data preparation is one of the most important and often time-consuming aspects of data mining projects and includes data selection, data cleansing, new data structuring, and merging. At this step, the raw data that has been taken into account is prepared, that is descriptive statistics are performed on it and the background for modeling is provided. The data is normalized and prepared to enter the modeling step. That is, the RFM variables are formed using processing, and the demographic variables are separately analyzed and placed in the spectrum to be prepared as inputs for the clustering algorithm.

Research hypotheses

The main objective: the analysis of the behavior of e-banking customers of Tejarat Bank using data mining (k-mean) based on RFM model.
Secondary objectives:

1. Calculation and efficacy of R, F, M parameters for electronic banking customers using data mining and clustering tools
2. Determining the characteristics of each cluster based on demographic variables
3. Using general clustering results to evaluate customer behavior

Research methodology

Researches based on the objective are divided into fundamental and applied researches. The researches are divided into two groups based on the data collection method: descriptive research and experimental research. Descriptive or non-experimental research consists of five categories: surveying, correlation, post-event, action research, case study. The experimental research is divided into two categories: full-experimental research and semi-experimental research.
This research is applied in respect to the objective. The research method is descriptive and data mining method is used. This technique explores for patterns containing information in the existing data. The K-Mean Algorithm and SPSS MODELER software are the techniques of this research.
The statistical population of this study is the users of Tejarat Internet Banking services in a period of 15 months from Farvardin 2016 to Tir 1396. In this research, Internet Banking services, Internet Bank data, mobile bank and telephone bank services are used. The data is extracted from the Bank database in the form of an Excel file and includes demographic information such as gender, age, and education level, as well as their account information, including the account number, transaction amount, and the number of transactions in the dates during these fifteen months. In fact, customer account flow information has been extracted using Internet banking. The number of these customers was about 1,700, and by deleting invalid data and empty data, a final number of 1197 valid data was considered.
In general, the methods of data collection in this research can be divided into two categories: library and field studies. Library methods were used to collect data about the subject literature and the history of the research. Field and documentation researches were used to extract data

from the customers of Tejarat Bank database. At first, raw data includes values related to customer transactions. To analyze the data, a descriptive statistic for data is initially considered. Then raw data must be processed to be prepared for the modeling. That is, the information must be obtained by processing raw data so that the field for the main analysis is performed. Also, the data is normalized so that the different units in the numerical range from zero to one normal can be normalized. Now the background for the data mining field is ready. The purpose of clustering is to divide the existing data into several groups, so that the data of the different groups should be as varied as possible, and the data in one group should be very similar.

RESULTS

Normalized data should be clustered in the SPSS MODELER software. But what is important is the number of optimum clusters for clustering by the K-mean algorithm. The number of clusters in the K-mean algorithm should be between 2 and 9. However, the exact number should be found by relationships and indicators. To determine the optimal number of clusters, a two-step algorithm is used. Therefore, first, clustering is done by a two-step algorithm to obtain the optimal number of clusters, and then the final clustering is done using the K-mean algorithm. Normalized data was opened in the software, and a two-step algorithm was applied to them. Due to two types of data, nominal variables (gender, age, education) and relative and distance variables (RFM), variables of both types in two-step clustering and K-mean are used. At this point, the quality of clustering is displayed with the cluster quality option. The quality of clustering is very close to its good quality. It is also observed in the results that the number of input variables in the input section is equal to 6 and the number of clusters considered in the cluster section is 8.

Given that the number of clusters was 8, the data were given to the K-mean algorithm and the number of 8 clusters entered. Demographic variables and RMF variables were both selected as inputs to the algorithm. Results can be seen in several tables. The important table formed is the mean formed for the cluster in this section. This average shows which data is located at which center of the cluster. This is the table used in the interpretation of the results, and according to this table, each cluster should be analyzed.

Table 1. The center of cluster in each field

	cluster							
	1	2	3	4	5	6	7	8
recency	.095	.093	.093	.092	.092	.096	.091	.101
frequency	.066	.061	.061	.062	.054	.053	.059	.066
Monetary value	.009	.009	.009	.008	.010	.008	.007	.010
gender	1	1	1	2	1	1	1	2
age	3	2	4	1	3	0	2	1
education	3	3	1	1	1	2	2	3

The next table shows the number of customers in each cluster. This table is also used to grade each cluster. Because it measures the density of the accumulation of each cluster and shows the number of users in each cluster.

The number of people in each cluster

Table 2. The number of customers in each cluster.

cluster	1	131.000
	2	223.000
	3	85.000
	4	160.000
	5	141.000
	6	35.000
	7	327.000
	8	95.000
valid		1197.000
invalid		.000

Summary and conclusion

Due to the data analysis and the results of the research, it is possible to discuss the research objectives.

Research questions were defined as follows:

1. How are R, F, M parameters are calculated for e-banking customers and what is its impact on data mining and cluster creation?
2. How are each cluster characteristics determined based on demographic variables?
3. How are the customer behavior in each cluster evaluated using the whole clustering results?

The main objective was to analyze the behavior of e-banking customers of Tejarat Bank using data mining (K-Mean Algorithm) based on the RFM model. It has secondary objectives of calculation and impact of R, F, M parameters for electronic banking customers using the data mining tool, creating clusters and determining the characteristics of each cluster based on demographic variables.

A customer with a lower R value and higher F and M values is a loyal customer to the Internet banking services. As seen, the cluster with lower R value was in the first rank in the cluster. This means that the customer's recency of transaction is very important. But according to the results of the research, it was found that the higher R value alone would not be the criterion for determining the lower-rank cluster. In determining the factors influencing the cluster, it can be stated that, in addition to RFMvariables, demographic variables affect the clustering too.

For example, in clusters 2 and 3, there is no difference in the RFM variables, also, in the case of gender, both clusters include men, but the cluster 2 is in the second rank and the cluster 3 is placed in the seventh rank of clustering. This is due to the age of users and their level of education. In the second cluster, there are young people with university education and in the third cluster there are people with over fifty years old with lower education. The low rank of this cluster is due to these two demographic variables of age and level of education. Therefore, plans should be made in the field of electronic banking education in the elderly. Because this group of people is often unable to use the internet with mobile phones, they do not have enough knowledge in this regard. By investing in training this group, more users can be placed in this cluster.

Regarding the higher value of R variable, the comparison between the two clusters 8 and 6, which have the highest values of R, is worth to be considered, and the cluster 8, despite having higher R than the cluster 6, is placed in higher rank. The reason for this is the higher values of F and M in the cluster 8 and also the demographic variables, which are completely different from each other in both clusters. The cluster 6 is in the last rank because age is not mentioned in its information. Updating customer information can prevent such incidents.

As the results show, in both clusters there were females. Cluster ranking in one of them was at a low level due to the high R value, although the level of education and the F and M values were higher in this cluster. By examining the full details of these women, it was found that women in this cluster used less e-banking services in recent months, and most of their visits were in person. Due to the high amount of money transfer, these women did not trusted using e-banking services. Managers should be able to re-use e-banking services for this level by relying on their trust and using secure internet channels. Moreover, managers should also be creative and efficient in developing and presenting strategies for using Internet banking for this group of people. Providing services that are tailored to the needs and desires of this class is important.

Age is also an important variable in determining clusters. As the results indicate, users who were between the ages of 30 and 40 were ranked first in clustering. Due to the younger age of the group and more use of the Internet, this age range use more Internet banking services. Comprehensive training at all levels and all ages and the culture of Internet banking should be established.

Research suggestions

- ***Practical suggestions:***

Customers are the most valuable assets of any organization. Due to clustering patterns, we need to consider the needs of different clusters. Managers should consider appropriate strategies for users to be encouraged to use Internet banking services. The benefits of the organization ensure users' loyalty by designing appropriate strategies and recognizing customers of each cluster in terms of transaction and demographic variables. Research questions were defined as follows:

1. How are R, F, M parameters are calculated for e-banking customers and what is its impact on data mining and cluster creation?
2. How are each cluster characteristics determined based on demographic variables?
3. How are the customer behavior in each cluster evaluated using the whole clustering results?

According to the results of the research, in order to increase each of the above issues, the following suggestions are made:

1. A proper decision must be made with a comprehensive look at six variables.
2. It is very important to provide services that are tailored to the needs and desires of each class in terms of gender, age and education level.
3. Creating a secure communication platform, without disconnection and error is very important.
4. With the proper management by branch managers potential customers can be identified and they can plan for them.
5. Customers who have a different and irregular financial behavior over a period in the communication chain with a bank and in return, may not have appropriate profitability for receiving various services. It is necessary for the managers and heads of the branches to take appropriate measures to improve the efficiency and profitability of this class of customers.

6. Training the employees of the branches in order to guide customers in using these services according to the customer cluster, increases the level of use of electronic banking services.

Research suggestions for future researches:

1. In this research, the clustering was conducted using the K-mean algorithm. The optimal number of clusters was determined by a two-step algorithm. It is suggested that genetic algorithms and decision trees be used in future research.
2. It is suggested to conduct a similar research on other modern banking networks, as well as ATM data which is not included in this study. It is suggested that appropriate strategies be developed by clustering ATM data.
3. Presenting a variety of model-based data mining algorithms in identifying customer behavior in the modern banking field;

Use of computational intelligence to decide and ranking the six variables studied in the research.

References

A. Hasanzadeh, Hesam M., Ghanbari (2012) "Classification of mobile banking users by data mining approach: Comparison between artificial neural networks and naïve Bayes techniques", Management Researches in Iran, 16 (2): 57-71, (in Persian).

Abedi Jafari, H. Jam Por az Mei, M and Biriya, E. (2010) The Challenge of Human Resource Management in Virtual Organizations - A Study of the Relationship between the Degree of Organization Duty and Organizational Commitment. Journal of Information Technology Management 5 (2) pp. 90-73

Alsajjan, B. & Dennis, C. (2010). Internet Banking Acceptance Model: Cross-Market Examination. Of Business Research Journal, 63 (9/10): 957-963.

Amorim, RC; Mirkin, B. (2012). "Minkowski Metric, Feature Weighting and Anomalous Cluster Initialisation in K-Means Clustering". Recognition Pattern. 45 (3): 1061-1075.

Azar, A., Ahmadi, P. & Sabt, M. (2010). Designing a Human Resources Selection Model with a Data Mining Approach (Recruitment of Entrants' Tests for a Business Bank in Iran). Information Technology Management, 4 (2) pp. 30-22

Azizi Sarkhoni, M, Allah Gholizadeh Azari, M., Kardlouei, H. (2008). Examining the existing infrastructure of the Tejarat bank to deploy e-banking. Scientific-extensive journal of management (researcher), (10) 1 pp. 11-5

Babaei, M. & Ahadi, P. (2010). Investigating the Relationship between Personality Characteristics and Internet Users' Purchasing Behavior in Iran. Information Technology (4) 2 pp. 52-43

Baradarn, V. & Farrokhi, Z. (2013), The Segmentation of Customers in the Banking Industry Using the Developed RFMC Model. Brand Management Quarterly Journal. The first year, the second issue, pp. 154-135

Berson, A., Smith S., & Thearling K. (2004). Building Data Mining Applications for CRM. New Dehli, Tata McGraw Hill.

Bose, R. & Sugumaran, V. (2003). Application of Knowledge Management Technology in Customer Relationship Management. Process Management and Knowledge, 10,3-17.

Buttle, F. (2004) Customer Relationship Management: Concepts and Tools, Elsevier Butterworth Heinemann, 2 (3) 256-274

Caceres R. (2007), Service quality, relationship satisfaction, trust, commitment and business-to-business loyalty Service quality. Journal of Marketing European; 41 (7): 836-867.

Celebi, ME, Kingravi, HA, and Vela, PA (2013). "A comparative study of efficient initialization methods for the k-means clustering algorithm". Systems with Applications Expert. 40 (1): 200-210.

Chapman, P. Clinton, J. Kerber, R. Khabaza, T. Reinartz, T. Shearer, C. Wirth, R. (2000). CRISP-DM 1.0 Step by Step Data Mining Guide. CRISP-DM Consortium.

Chen, YL, Kuo, MH, Wu, SY & et al. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customer's purchasing data. Commerce Research and Application Electronic, 8 (5), 241-251,

Cheng, Ching-Hsue and Chen, You-Shyang (2009), Classifying the via RFM model and RS theory of customer value segmentation, Expert Systems with Applications, Vol. 36, pp. 4176-4184.

Christodoulakis, D. Aggelis, V. (2009). Customer Clustering using RFM analysis. Expert System with Applications, 36, 2678-2685.

Coates, Adam; Ng, Andrew Y. (2012). "Learning feature representations with k-means ". In G. Montavon, GB Orr, K.-R. Müller. Neural Networks: Tricks of the Trade. Springer.

Deshpande, SP & Thakare, VM (2010). Data Mining System and Applications: A Review. Journal of Distributed and Parallel Systems International (Ijdps), 1 (1): 32-44.

Divandari, A., Abedi, A., Naserzadeh, M. R (2013). Presentation a Conceptual Model for Explaining the Key Factors Affecting the Quality of Internet Banking Services Provider Systems (a survey around Mellat Bank). Information Technology Management: (1) 5 pp. 36-19

DJ Hand, Mannila H., P. Smyth (2001) " Principles of data mining

Ebadi, M & Alizadeh, S. (2015). Customer behavior analysis in buying and sending online group SMS using data mining based on RFM model. Sharif Industrial Engineering & Management. D. 1-31 Issue 2 pp. 35-27

Fader, PS, Hardie, BGS, & Lee, KL (2005). RFM and CLV: Using ISO-value curves to customer base analysis. Of Marketing Research Journal, Emerald Management Reviwe, 42 (4), 415-430.

Farrokhi, S & Timourpour, B. (2016) Identification and Classification of Customers of Iran's Banking System in terms of perceived expectations and value of banking services using data mining techniques. Journal of Research in New Marketing Research, Year Six, No. 1, No. 2, pp. 25-14

Fayyad, U. Uthurusamy, R. (1996). Data Mining and Knowledge Discovery in Databases. Of the ACM Communications, 39: 24-26.

Fouladifar A, Taghipour E, Hedayati A. (2016). Market Segmentation for Marketing of Banking Industry Products Constructing a Clustering Model for E-Banking Customers Using RFM Technique and K-Means Algorithm. Business Management International .10 (6): 1106-1119

Gikandi, Joyce, Wangui, Bloor Chris. (2010) Adoption and effectiveness of electronic banking in Kenya. Electronic Commerce Research and Applications; 9: 277-282.

Haghighi Nasab, M. & Khosravi, S. (2011). Evaluation of the level of maturity of organizational interactivity of Iran Information and Communication Technology Research Institute. Journal of Information Technology Management 6 (3) Pages 20-1

Hashemian, M. Isaie M. T MiKaeili, F Tabatabaei, M. (2012). Factors Affecting Adoption of Electronic Banking Tools by Customers (Survey on Saman Bank). Information Technology Management: (11) 4 pp. 174-155

Hu Y.-H. and Yeh, T.-W. (2014) Discovering frequent patterns based on RFM analysis valuable information without customer Identi cation, Knowledge-Based Systems, 61, pp. 76-88

Hua, Y.-H., Huangb, TC- K. and Kaoa, Y.-H. (2013) Knowledge discovery of weighted RFM customer sequential patterns from sequence database ", The Journal of Systems and Software, 86, pp. 779-788

Hughes, AM, Strategic Database Marketing, Probus Publishing, Chicago (1994).

Jamali, Gh, & Hashemi, M (2012). Measuring Relationship between Factors Affecting Risk of Mellat Bank IT Projects in Bushehr Province Using Fuzzy DEMATEL. Information Technology (9) 3 pp. 35-29

Jayawardhena, C. and Foley, P. (2000). Changes in Banking Sector - the Case of Internet Banking in UK. Internet Research 10 (1): 19-30.

KA Forcht, Cochran K. (1999) "Using data mining and data warehousing techniques", Industrial Management & Data Systems, pp. 189-196.

Karjaluoto, H., Mattila, M. & Pento, T. (2002). Electronic Banking in Finland: Consumer Belief Sand Reactions to a New Delivery Channel. Financial Services Marketing of Journal, 6 (4): 346-361.

Khan Babaei, M. & Zain al-Abedini, F (2013), A Model for Using Data Mining Techniques to Identify Segmentation and Customer Behavior Analysis of Electronic Banking Services. Journal of New Marketing Research, Third Year, No. 188-175

Kiss, F. (2003). credit scoring processes from a knowledge management prospective. Periodica Polytechnica Ser. SOC. MAN. SCI, 11 (1): 95-110.

Kolodinsky J.M, Hogarth J.M., Hilgert M.A. (2004), The adoption of electronic banking technologies by US consumers. The International Journal of Bank Marketing; 22 (4): 238-25.

Lee Ming-Chi. (2009) Factors influencing the adoption of internet banking: An perceived risk and perceived benefit of TAM and TPB with integration. Commerce Research and Applications Electronic; 8: 130-141.

Liao Z., Cheung M.T. (2001) Internet-based banking and consumer attitude. Information & Management; 38: 299-306.

M. Namvar, Gholamian MR, Khakabi S., (2010), A Two Phase Clustering Method for Intelligent Customer Segmentation, International Conference on Intelligent Systems, Modelling and Simulation. Pp. 215-219.

Macqueen, JB (1967) Some methods for classification and analysis of multivariate Observations, Proceeding of 5th Berkeley Symposium on Mathematical Statistics and Probability, 1, pp. 281-297 Available at: www.elsevier.com/locate/dsw

Mahmoudi, K. Rostami, h. Sayebani, M. & Moradi, A. (2013). An overview of data mining and its applications in offshore industries. 5th National Conference of Offshore Industries of Sharif University of Technology. OIC May 31 and June 1, 2013

Marbán, O. Segovia, J. Menasalvas, E. Fernández-Baizán, C. (2009). Toward data mining engineering: a software engineering approach. Systems Information, 34 (1): 87-107.

Moghaddasi, A. (2010) Types of Payment Methods in Electronic Banking. Information Technology Age monthly journal. 58 pp. 75-71.

Mohammadi, Sh, and Alizadeh, S (2014), analyzing the problems of Ayandeh bank branches all over the country using data mining. Information Technology Management, Faculty of Management, University of Tehran, Volume 6, Issue 2, Pages 350-333

Momeni, M. (2011). Data clustering (cluster analysis). Tehran Forouzesh Publisher

Punj, GN and Stewart, DW (1983) Cluster analysis in marketing Redearch: Review and suggestions for application, Journal of Marketing Research, 20, pp. 134-148

Radfar, R. Nezafati, & Yousefi-al Asl, A. (2014). Classification of Internet Bank Customers Using Data Mining Algorithms. Information Technology Management. 6 (1) pp. 90-71

Rasooli, H. Manian, A. (2012). Designing a Fuzzy Inference System for Selecting Electronic Banking Services (Case Study of Sepah Bank). Information Technology Management, (12) 4 pp. 64 -41

Razmi, J. and Ghanbari, A. (2009). a new model for Calculating customer lifetime value ", Journal of Information Technology Management, 1 (2), pp. 35-50

Rouholamini, M. Venkatesh, S. (2011), A Study of Customer Relationship Management in Iranian Industry Bank, International Journal of Information Technology and Knowledge Management, 2. Pp. 723-729.

Sami Zadeh, R. (2007). Data mining and customer relationship management. Tehran: Roshd Andishe Publication.

Seyed Hoseini, M., MA Gholamian, MR (2010). Cluster analysis using data mining approach to develop CRM methodology to asses the customer loyalty. Applications with System Expert, 37, 5259-5264.

Sohail Sadiq, M., Shanmugham. B. (2003) E-banking and customer preferences in malaysia: An empirical investigation. Information Science; 150: 207-217.

Taqavi Fard, M. T., Mansouri, T. Naserzadeh, M. R., Ferasat, A.R (2007). Data Mining and Its Application in Decision Making. Knowledge Management, 20 (79) pp. 17-3

Taqavi Fard, M., & Nadali, A. (2012). Classification of Applicants for Credit Facilities by Using Data Mining and Fuzzy Logic. Industrial Management Studies, (52) 9, pp. 91-52

Turban, E. (2011). Business Intelligence: A Managerial Approach, 2nd ed. Saddle River Upper, NJ: Prentice Hall,

Vattani., A. (2011). "k-means requires exponentially many iterations even in the plane". Computational Geometry and Discrete. 45 (4): 596-616

Vercellis, C. (2009) Business Intelligence: Data Mining and Optimization for Decision Making. Hoboken, NJ: Wiley & Sons.

Wei, J., Lee, M., Chen, H. (2013). Customer relationship management in the hairdressing industry: An application of data mining techniques. Systems with Applications Expert, 40, 7513-7518.

Wei, J., Lin, Sh., Weng, Ch. (2012). A case study of applying LRFM model in market segmentation of a children's dental clinic. Systems with Applications Expert, 39,5529-5533.

Zahedi, Shams al-Sadat (2011). Research on websites of five major Iran State Universities and providing a good model. Journal of Information Technology Management 6 (3) Pages 44-21.