

The Relationship between Customer's Personal Characteristics and Usage of Electronic Banking Services*

Aram Heidari

Payam e Noor University of Sanandaj, Iran
aram.heidari@gmail.com

Nadia Asady

SAM Consulting Engineers
shadi_asady@yahoo.com

Abstract: Nowadays, data is the heart of business processes of small and large companies, such as retailers, communication, production, facilities, transportation, insurance, credit cards and banking. Hence, there is a need for a tool that can process stored data and provide users with the information obtained from this process. In this context, data mining is one of many scientific branches that in recent years has experienced a rapid expansion worldwide. Data mining is the process of discovering knowledge (that is) hidden in data, which by describing, explaining, predicting and controlling various phenomena, has a wide application in various fields. The main idea of data mining is based on the fact that old data contain information that can be used in the future and will be valuable. The purpose of data mining is to find those patterns in existing data that make the needs, preferences and desires of the business clearer. Therefore, extracting the symptoms from useless things, i.e. identifying underlying patterns within seemingly random variables, is one of the important roles of data mining. The aim of this paper is to analyze the relationship between individual characteristics of the client and the use of electronic banking services (using data mining methods).

Keywords: Data mining, clustering, e-banking

1. Introduction

In information explosion era, organizations will be producing and collecting huge amounts of data daily. Extracting useful information from databases and converting that information to applicable results is a challenge organizations are faced. Considering nation's progress in information technology and government's special attention towards e-government, penetration of computer systems in industry, and development of large databases by government organization, the need for applying data mining in banks and private sectors is felt deeply. The advent of information technology has affected many industries including banking. Electronic service are usually delivered to customers through electronic channels such as Automated Teller Machines (ATM), phones, the Internet and more recently cell phones.[1] Nowadays IT is delivering many services online and many customers have access to online services. Moreover, E-banking is defined as providing banking services through the Internet and the capabilities that it provides. The Internet has a lot of

advantages both for banks and customers as well. Using this technology, customers not only are able to do their banking tasks anywhere and anytime, but it also has resulted in cost reduction and efficiency increase for banks. However, not all organizations and firms have been equally and adequately successful in employing E-commerce. Thus, it seems that necessary studies should be conducted in order to develop favorable policy for deployment and using e-commerce, and also determining those factors that enhance its increasing trend. Considering the increasing expansion of internet networks in the country and provision of a framework for organization to use it for delivering their services to customers, studying the key factors that lead to better provision of electronic services is essential. Such studies are especially more critical and beneficial for financial-service institutions such as banks which in recent years have made a lot of effort in delivering electronic and online services.[2] Therefore, the aim of this study is to propose a model for predicting the level of success in delivering such services, by determining the key factors effecting successful online delivery of banking services.

2. Weka

There have been many commercial and educational data mining applications in different domains presented to science and technology world so far. Each of those applications is focused on specific algorithms according to the main type of data they handle. Precise scientific comparison of these tools should be performed at different aspects such as diverse input data types and formats, possible volumes for data processing, implemented algorithms, methods for assessing the results, visualization methods, data preprocessing methods, user friendly interfaces, compatible platforms, software price and availability. Among them, WEKA is introduced here due to its much extended facilities, ability to compare output from different methods simultaneously, good documentation, efficient user interface, its compatibility with other Windows applications, and most importantly the existence of a comprehensive book about it.[5]

*This paper was presented at the Third International Scientific Conference of University of Human Development (April, 2016)

3. Data Clustering

Data clustering purpose which is sometimes called Cluster Analysis is to discover natural grouping among a set of patterns, objects or points. An operational definition of clustering can be as follows: given n objects, find k groups (clusters) according to a similarity criteria in such a way that similarity criteria is high among members of a same group, but low between members from different groups.[3]

Clustering is one of the most important tools for data discovery that is used in random discoveries. Today, knowledge extraction is viewed as a major bottleneck in knowledge engineering process. Machine learning and data mining algorithms with the purpose of extracting knowledge from data, are considered as a method to overcome this problem. A common approach in this field is data clustering which, for decision making or clustering or classification, can assign some symbolic decisions on new samples by using existing samples. Clustering is in fact an unsupervised operation. This operation is applied whenever we want to find groups of similar data without having a prior prediction about the existing similarities. It is usually used to determine groups of customers that are not known before. Clustering is a task during which groups of data or items exist, in such a way that each item is assigned to one cluster and members in a cluster should have an intrinsic similarity, and also criterion for similarity must be completely defined and be measurable for each pair of items. Therefore, in each cluster there exists a self-similarity between items in the cluster. In clustering method there are no groups at first, and in fact variables are not divided into independent and dependent variables, but rather we are looking for groups of similar data and by discovering these similarities, behaviors and be better identified and based on that behaviors we can act in such a way that better results are gained.[4] Clustering methods normally consist of two methods: hierarchical and partitioning in hierarchical clustering clusters are formed gradually and final clustering is done after several stages. Partitioning methods, on the other hand, tend to define all clusters at once. The method that is used in this paper is k-means clustering. This method first defines k clusters and determines their centroids. It does not matter what should be considered as clusters' centroids to initiate the task, thus any k points can be considered as initial centers. Then k-mean algorithm performs the following three phases repeatedly:

- 1- Determine coordinates of the centroids;
- 2- Calculate the distance of each point from the centroids;

- 3- Assign each datum (point) to the cluster center of which has the shortest distance from the point.

These steps are repeated until no point moves between the clusters. The final result of clustering is that the distance between each point from the center of its cluster is less than its distance from the centers of all other clusters. K-means method can be defined in different ways in which definition of the cluster centroid and distance are different.[6]

Advantages of clustering are as follows:

- 1- Clustering is an indirect method. The power of clustering is in being an indirect method which means that it can be applied even when there is no prior information about internal structure of database. It can be used to discover hidden patterns and also improving the performance of direct methods.
- 2- Clustering can be used for different types of data. By choosing different distance measures properly clustering can be applied to most data types.
- 3- It is simple to use. In this method it is not required that some fields be considered as input and some fields as output, and in most of clustering methods weighing the data is less necessary.

Also, disadvantages of this method are as follows:

- 1- Choosing the precise measures for distances and weights is not a simple task.
- 2- This method is sensitive to initial parameters such as number of clusters, minimum of closeness and initial clusters.
- 3- Interpreting the results can be difficult and usually needs analysis by experts in the field of commerce.[7]

4. Research Model

Considering the importance of the problem that nowadays most organizations are dealing with a large volume of raw and unprocessed data and records and optimized use of these data by old methods of analysis is very difficult and time consuming, using new knowledge of data mining to solve this problem can be effective. With the help of data mining technics like clustering, the data in data bases are processed and transformed into applicable and useful data, for data mining is a process of knowledge extraction from data which is performed through discovering existing patterns in the data.

In this study, after collecting required data and information about personal characteristics of customers (age, gender, marital status, educational level, occupation, monthly income level, level of using electronic banking services) through questionnaires, all data were integrated and prepared for analysis. In this study, data about 500 customers of one of the branches of Bank Mellat in Sanandaj, in Shahrivar, 1394 (August 2015) was analyzed. Sample selection method was considered to be random and done in different times of different week days. Then data analysis was performed on collected data using Weka software and k-means clustering.

5. Data Analysis

After entering data in Weka, results were obtained as follows where not only centroid of each cluster is determined but percentage of data assigned to each cluster is observable. Here, data is divided into four clusters, and according to the results each cluster can be analyzed and the relationship between customer personal characteristic and usage of electronic banking services can be explored.

Cluster 0: About 28 percent of all sample is in this cluster, and from the information about the cluster centroid it can be concluded that most of people in this cluster, which are females under 20, single, with high school diploma and unemployed which have a monthly income less than 200,000 tomans, have not used any of electronic banking services.

Cluster 1: About 29% of the total sample are assigned to this cluster and it can be seen from the information about the centroid of this cluster that most of people in the cluster are males above 40, married, with high school diploma and self-employed with monthly income more than one million tomans, which among electronic banking services, have used mobile based services more.

Cluster 2: About 28% of the sample fall in this cluster and according to the cluster center information, it can be seen that most of the people in this cluster are men between 20 to 30 years of age, married, with higher education and employed (civil servants) which have a monthly income between 500000 to one million tomans, and among electronic banking services, mostly used internet and mobile services.

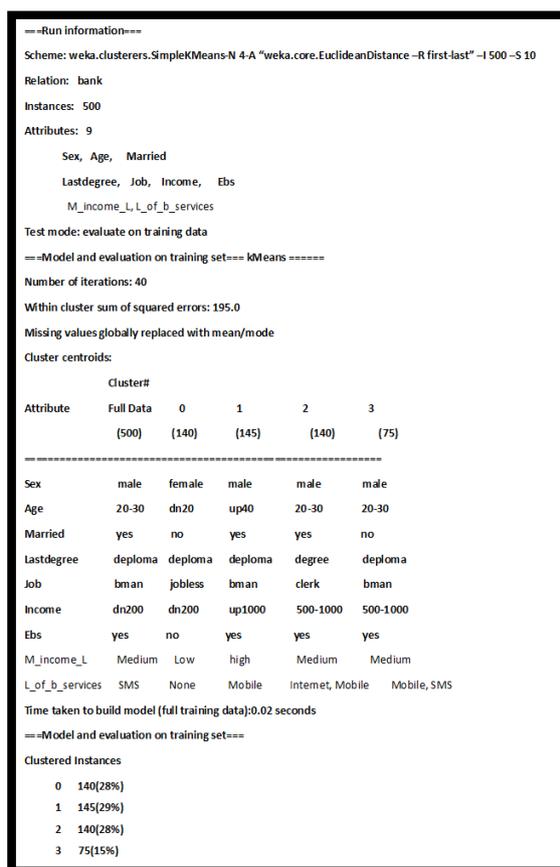


Figure 1 Weka's Result

Cluster 3: About 15 percent of the total sample are in this cluster and from the information about cluster center, it can be concluded that most of the people in this cluster, are males from 20 to 30, single, with high school diploma, self-employed with a monthly income between 500000 and one million tomans, and they have used mobile and SMS based banking services.

Based on the results from cluster analysis it can be concluded that among four clusters, three have used e-banking services. Furthermore, by summarizing these three clusters it can be concluded that most of the customers who are using e-banking services of the studied bank have the following characteristics: male, 20-30 years of age, married, high school graduates, self-employed, with monthly income between 500000 and one million tomans.

6. Conclusion

Today, many organizations are using data mining to assist managing all phases of customer relations which include acquiring new customers, increasing profit from existing customers, and maintaining good customers. By exploring customers' characteristics and reasons for usage of some of their products and services by customers and also reasons why some of

their products and services are not well received by some customers, these organizations can identify their products' strengths and weaknesses and put efforts in increasing the quality of their products. In this manner, banks like any other organization are no exception and need analysis and discovery of different information from their accumulated data. Some of the applications of data mining in banking are as follows:

- Customer behavior analysis and marketing
- Validating a credit customer
- Management of different bank risks
- Lowering transfer costs, simplification of budgeting, planning and profit analysis.
- And most interesting application of data mining is detection of money laundering.

Thus bank can gain considerable achievements in any of the mentioned areas by improving its data bases, bank's information and communication links, and increasing staff's cautiousness in entering initial information and ..., and also by relying on data mining methods.[8,9]

References:

- [1] Salehi Sedghiani, J. and Sarvarnezhad, S., "determine new strategies for banking in the adoption of mobile banking service users in private banks" , Strategic Management Studies, pp. 39-57, No. 6, Summer 2011
- [2] Akbarpour Shirazy, M. and Toopchy, Hussein, "Data Mining: Concepts, Methods and Applications", University of K-N,2012
- [3] Eman M. Ali, Ahmed F.Seddik, and Mohammad Haggag, "data mining techniques", Journal of Biomedical Engineering No. 131, of 11 February 2012,
- [4] Bahrami Zennor, Maryam, "Data Mining: discover hidden knowledge data", Persian date Farvardin, 2010
- [5] Jalali talab, Atoosa. Bozorgzadeh, Sanaz, "a data mining application using weka", University of Science and Technology, the summer of 2010
- [6] Kazemi, M., Kurd, Bagher, and Mehrvarzy, Muhammad, "Key Factors for successful internet services and providing a predictive model using decision tree", Research Management, pp. 29-45, Issue I, winter 2010
- [7] Shahrabi, Jamal, "data mining book," Research Institute of Amir Kabir University of Gita and data processing, printing, 2007
- [8] Soltani, K., "Data Mining with weka", Network Magazine, Issue 122, Persian date May 2011
- [9] Shahraki, Alireza and Pourghassem, Alireza, "Data mining, a powerful tool in improving organizational intelligence", the first national conference on business intelligence, 29 and 30 Persian date October 2010.