

Project Management

Elixir Project Mgmt. 107C (2017) 47026-47032

Elixir
ISSN: 2229-712X

Facilitating the Process of Customer Relationship Management through Data Mining Technique; Evidence from a Private Bank in Iran

Ahmad Reza Arabi¹, Hossein Vazifehdust² and Mohammad Reza Nazari Hashemi²

¹M. A in Executive Management, Islamic Azad University, Science and Research Branch, Khuzestan.

²Faculty Member, Islamic Azad University, Science and Research Branch, Khuzestan.

ARTICLE INFO

Article history:

Received: 9 May 2017;

Received in revised form:

29 May 2017;

Accepted: 8 June 2017;

Keywords

Data Mining,
Clustering,
Customer Relationship
Management,
Transaction.

ABSTRACT

Data mining is a technique that utilizes a range of data to provide valuable information. Customer relationship management through data mining can be a powerful tool for attracting, retaining, and promoting customers in today businesses. Banks, financial institutions, and insurance institutions are of those business that collect and storage many data from their customers on a daily basis. This research on one of private banks in Khomein city was accomplished in September 2012 in which 5939 customers with 7199 transactions were studied. These transactions were modeled using four demographic characteristics including age, gender, education, and occupation. At first, data was preprocessed which was the most time-consuming part of the study. After that, using visualization technique, graphic display of customers' demography was obtained. And then, the customers' transactions were predicted based on their transaction volume and probability, and it was checked whether the customers accept the bank's new service (i. e. instant card issuance) by using decision tree algorithm. Finally, the customers were clustered by conducting the two step algorithm and the result was that the cluster of 2543 transactions (cluster 1) is comprised of target customers. The characteristics of customers in target cluster are the age of 25~40, often having a bachelor degree, and be self-employed.

© 2017 Elixir All rights reserved.

Introduction

In the recent decade, customer orientation has become an important issue for organizations. Planning to response the needs of the customers, or CRM (Customer relationship management) is an inevitable subject that can directly influence changing a customer to a permanent customer and also benefiting the company. On the other hand, merging IT with CRM is essential to inform about customers' needs and their satisfaction. CRM begins mainly with customer's behavior analysis. Organizations that wish to relate with their customers, should collect all the data related to their customers, satisfy them, and finally, predict customers' behavior using IT infrastructures (Al-Mudimigh et al., 2009: 1).

However, the important problem is that the data increases during the time and the old technologies will not be able to help us. For instance, ATMs or Automatic trading machines are among electronic tools that help us satisfy customers' needs. Customers have transactions with ATMs and this has made it essential to increase the number of ATMs. Number of ATMs in Iran was 27403 by 2012 and 184 million electronic cards were issued during the same time.

These huge volume of data is a beneficial resource for extracting data by which we can describe and predict customers' behavior. But how to use these data and changing them into valuable and understandable information is not an easy job. As a result, the case study bank was eager to use this information in order to improve CRM process.

We have to use Data Mining in order to be able to find the answer to this question.

The main question of this study is that: what information is stored in financial transactions of under-study bank, and how can we facilitate CRM using this information?

2. Literature review

2.1. Data mining

Customers' data are recognized as the strategic capital of the organizations. Ability to extract knowledge patterns of this data has become a competitive advantage nowadays. Applying a computer-based methodology including new techniques for knowledge discovery of data is called data mining. Data mining is a kind of cooperation between human and the computer (Kantardzic, 2011: 1). MIT journal has chosen Data Mining as one of the 10 new technologies of the world that will be able to change the world dramatically (Larose, 2006: 2). Fayyad et al. (1996) defined data mining as a part of knowledge discovery process. KDD or knowledge discovery in databases is a process which its purpose is to find valid, beneficial, and understandable patterns in data. Berry and Linoff (2000) defined data mining as the process of analysis of a huge rate of data by automatic and semi-automatic tools in order to extract significant patterns and procedures (Hsu, 2009: 4185). The most frequent definition of data mining is: "extracting knowledge and information, and discovery of the hidden pattern in complex and huge databases". Telecommunication companies, banks, marketing and advertising companies, and all companies with huge databases can benefit from data mining as a strong tool for their advance in businesses.

2. 2. Customer relationship management

Today, customers are more knowledgeable and influential compared to 60s and 70s, and so, it is now harder to understand them. There are also more complex relationships with customers nowadays. According to Thearling (2000), we should find new and quick methods for achieving customer orientation in companies (Turk, 2007: 14). CRMs make infrastructures to facilitate long relationship with customers (Hendricks et al., 2007: 68). Brown and Coopers (1999) defined CRM as: "CRM is a key competitive strategy you need to focus on customers' needs and achieve customer orientation in your organization". Chatterjee (2000) claimed that CRM is a system focusing on automating and developing business processes along with customers' relation management in areas of sale, management, services and supporting the customers. According to Feinberg and Kadam (2002) when customer retention increases 5%, benefits increase 25-85% (Shahin & Nikneshan, 2008: 69). CRM is a new management mechanism used for improving the relationship between organizations and customers (Hua & Zhenxing, 2009: 252).

3. Research framework and questions

CRM is essential for every organization because they can recognize their customers well. A successful organization needs to increase cooperation with customers so CRM has become a success factor in recent years. One of the most important concepts of CRM is having integrated information about the customers and one of the resources of this information is customers' data storage. But changing data into beneficial and valuable information has stages that must be evolved carefully. Our applied framework for extracting the information from bank database is divided into three main parts: "pre-process of the collected data and changing them into the favorable format, data-mining with proper techniques and finally, evaluation and measurement of the extracted models with trial data".

As the name "data mining" conveys, data mining is a process on data. The first part is the theoretical framework of data pre-processing and changing them into the favorable format. In this part, using SPSS and EXCEL software will help us pre-processing of the data and coding data with the favorable format. The second part of the process is modeling the data including three operations of prediction, clustering, and imaging. And finally, the third part is to evaluate and inspect the models by entering trial data.

Questions of which response we are seeking here are:

Q1: what are the demographic features of the bank's target customers (i. e. those customers who have high financial transactions)?

Q2: is there a same financial pattern (in terms of financial transactions) in demographic features of the bank's target customers?

Q3: what are the demographic features of the bank's special customers (i. e. those who have used new services of the bank)?

Q4: is there a same pattern (in terms of response to new service) in demographic characteristics of the bank's special customers?

Q5: What are the clusters of the bank's customers of ATM service by using the data mining technique?

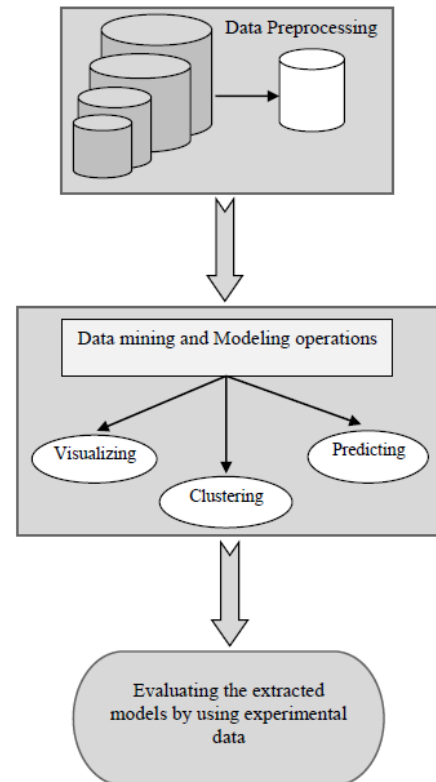


Fig 1. Conceptual model.

4. Research method

The aim of this cross-sectional applied research is to describe the properties of financial transactions and express the demographic features of ATM customers plus finding ways to facilitate the process of relationship with customers. Data of research are obtained from the central server of the case study bank of which financial transactions are stored in ORACLE database (all data are kept confidential for the security of the bank). Statistical society of the research includes all the customers who have a credit card of the case study bank and had transactions during September 2012 in Khomein city; Iran. The software of Clementine 11. 1 was used for the analysis of the data because it has a variety of algorithms and graphical charts.

5. Research findings

5. 1. Explaining the customers characteristics

As it is shown in figure 2, the proportion of good transactions is higher for males and proportion of moderate and bad transactions are equal for males and females.

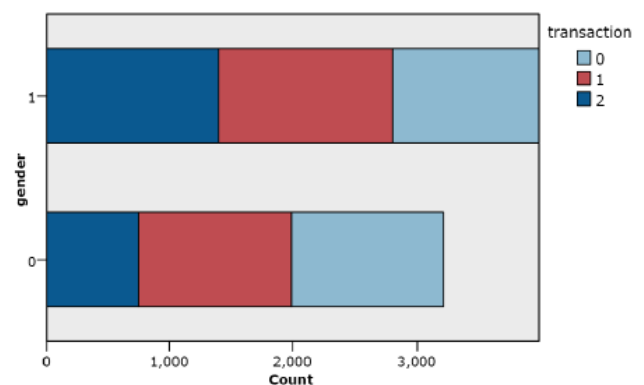


Fig 2. Distribution of transactions' amounts by gender.

As it is shown in figure 3, the proportion of good transactions is higher for the age range of 25~40 and 40~55.

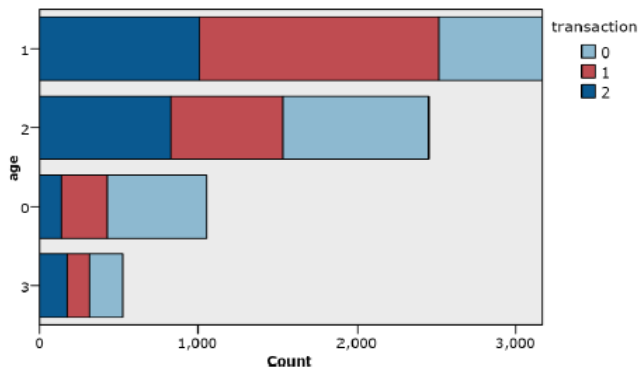


Fig 3. Distribution of transactions' amounts by age.

As it is shown in figure 4, the most of transactions are good for the MA level and the proportion of good transactions to moderate and bad ones is higher for BA than diploma.

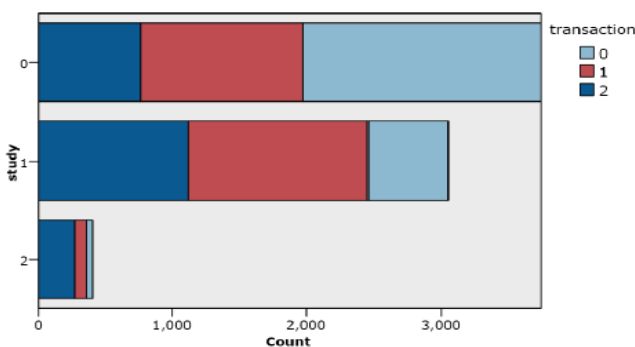


Fig4. Distribution of transactions' amounts by education.

As it is shown in figure 5, the proportion of good transaction to bad ones is higher for informal than formal employed people.

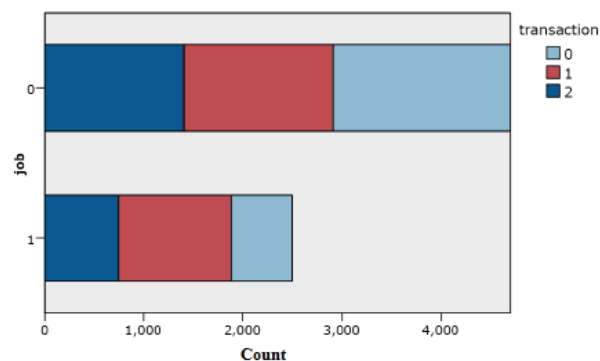


Fig5. Distribution of transactions' amounts by job.

As it is shown in figure 6, the proportion of good transactions is higher for those people who have used the bank's new service.

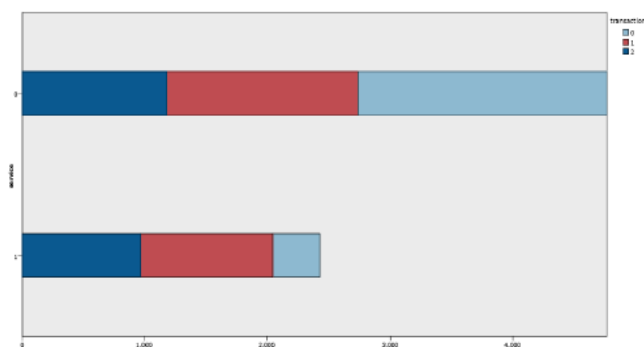


Fig6. Distribution of transactions' amounts by new service.

5. 2. Predicting the customers' behavior using the C&R Tree algorithm

As it can be seen in figure 7, decision tree algorithm shows all the combinations demographic variables based on which the customers' behaviors are predictable. Furthermore, this algorithm shows the probability of an event and concludes based on if-then rules.

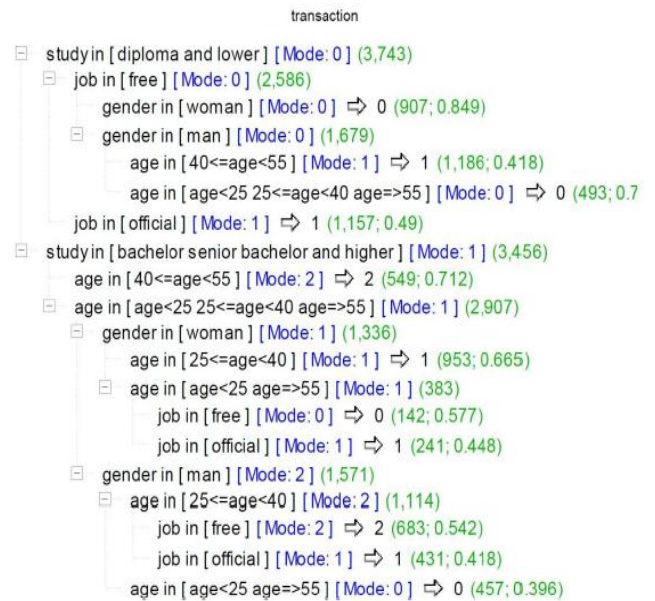


Fig 7. Algorithm for predicting transactions by demographic variables.

As is can be seen in figure 8, tree diagram tells that who (what demographic characteristics) have good, moderate, and bad transactions. For example, one of the extracted rules from this algorithm is as follows: Those people who have MA degree and higher, who are male, who are ranged in 25~40, and who are self-employed will have good transactions with probability of 54. 173 percent.

As is can be seen in figure 9, extracted rules for predicting the people who will use the bank's new service are only based on the two fields of age and education. In fact, the fields of gender and job have no role in predicting the bank's new service usage. For example, some extracted rules are as follows: a) those who have Diploma or lower will negatively answer to the bank's new service with the probability of 92. 359 percent; b) those who have BA and 55 years old or higher will negatively answer to the bank's new service with the probability of 65. 532 percent.

5.3. Clustering the customers using the two step algorithm

To determine the number of desired clusters automatically, SPSS software offers an efficient two-step procedure for analytical clustering. This is an additive clustering algorithm which is suitable for handling very large data pools and works with both the discrete and continuous data. The steps are as follows: 1) primarily clustering of records to a large number of small sub-clusters; and 2) extracting the desired and optimized number of clusters from the first stage (Zhang et al., 1996; Chiu et al., 2001). As it can be seen in the following table, this clustering algorithm categorized our data to three optimized clusters each of which have different characteristics. Table 1 provides the statistics of these clusters.

Figure 11 shows the characteristics of extracted clusters based on existing fields. For example, gender diagram in the cluster 2 is bigger than the other clusters indicating the majority of customers in this cluster are male.

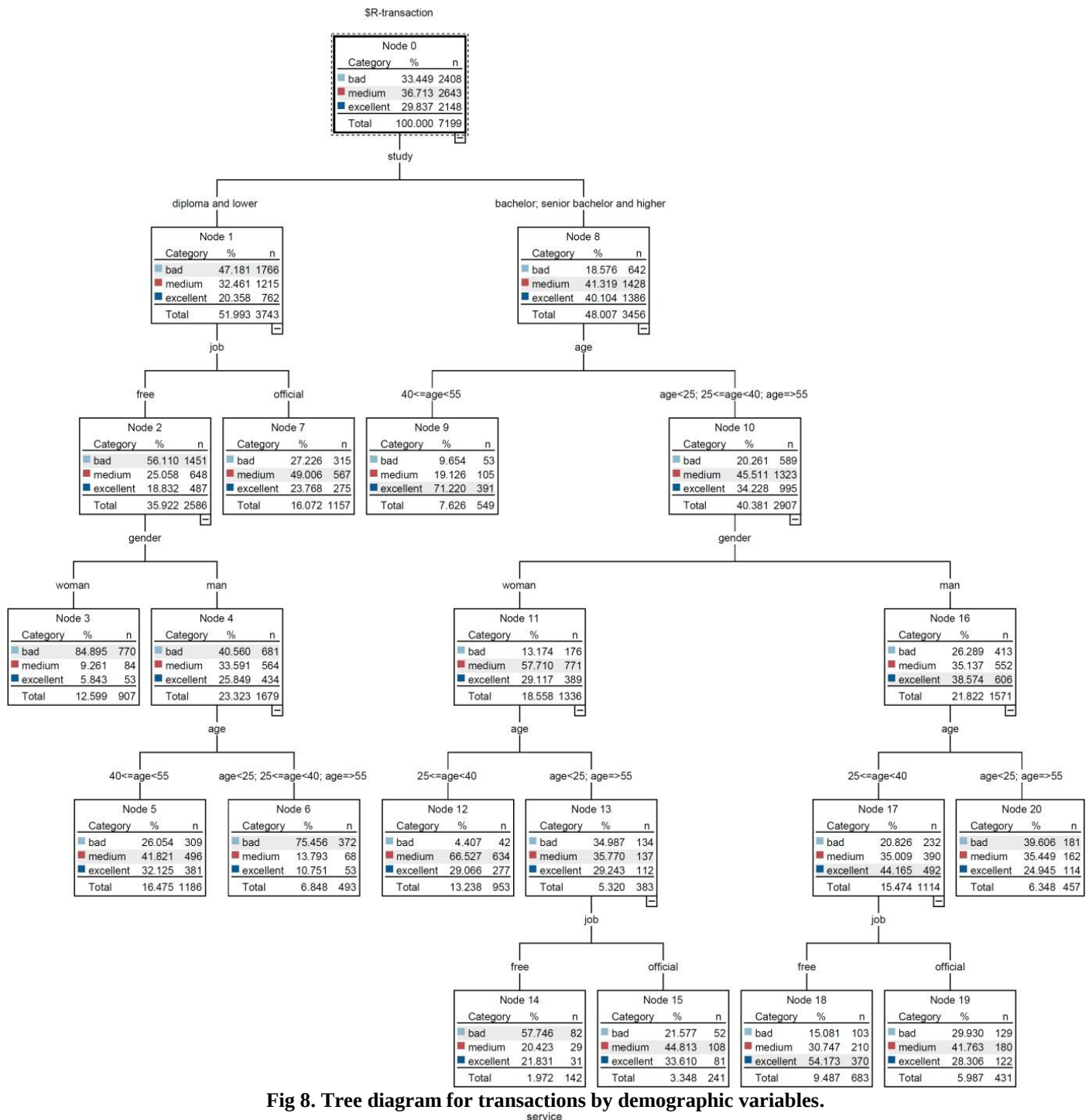


Fig 8. Tree diagram for transactions by demographic variables.

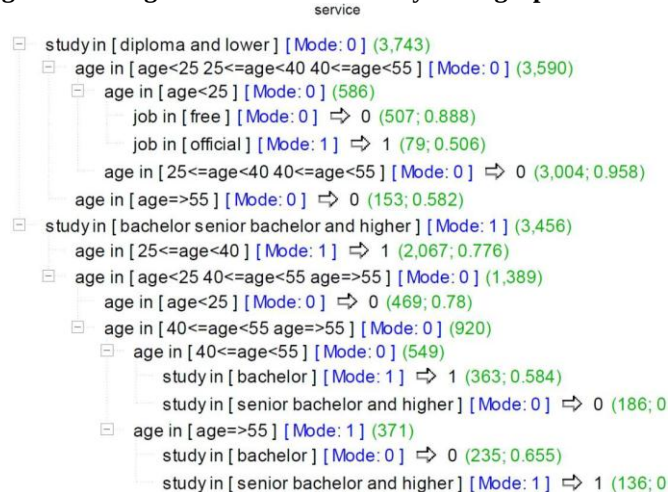


Fig 9. Algorithm for predicting the demographic variables of new service users.

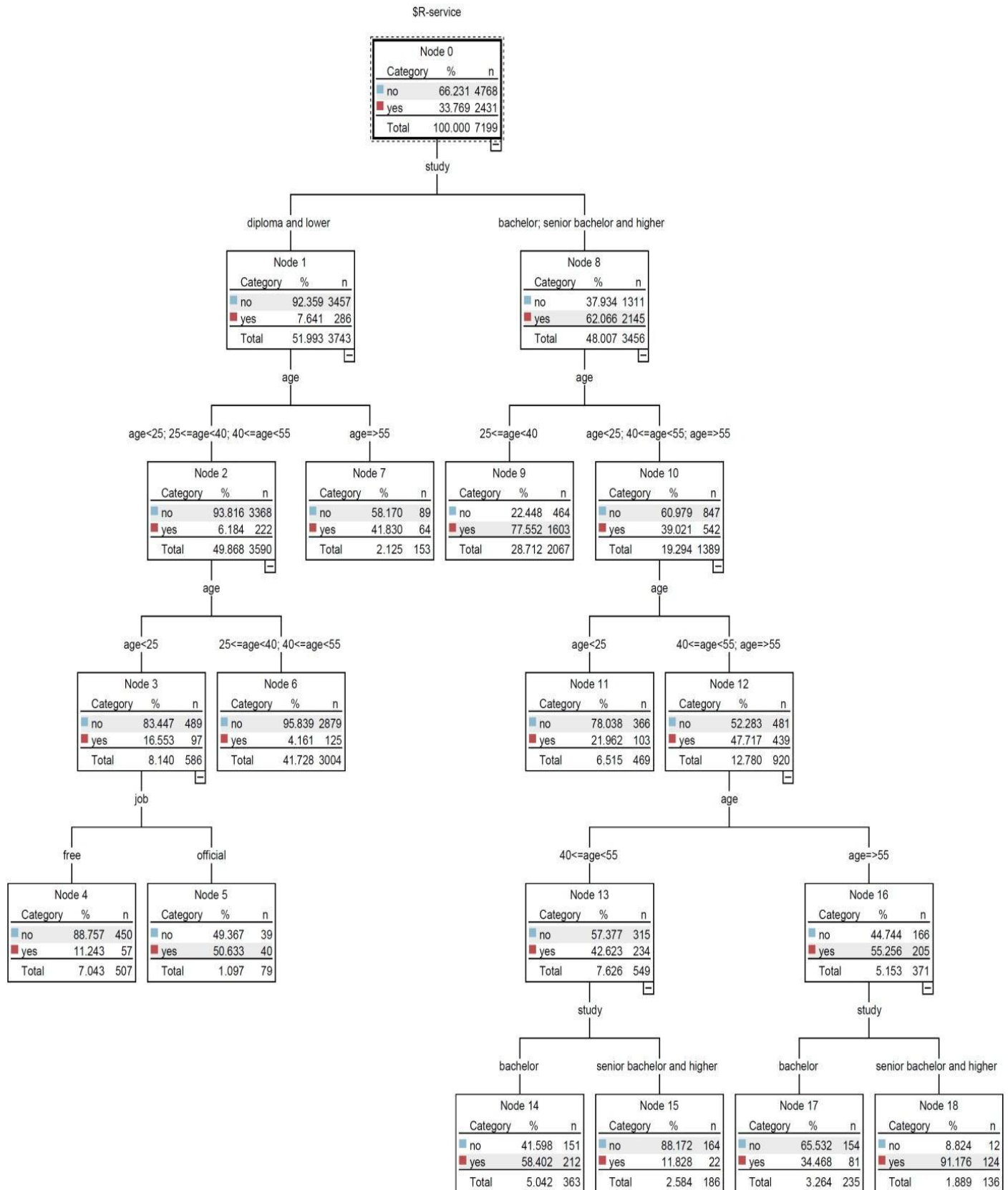


Fig 10. Tree diagram for predicting the demographic variables of new service users.

Table 1. Statistics of extracted clusters.

Clusters		Age	Gender	Education	Job	New service	Transaction
Cluster 1	Mean	1. 285	0. 526	1. 037	0. 355	0. 944	1. 291
	SD	0. 746	0. 499	0. 486	0. 479	0. 23	0. 697
Cluster 2	Mean	1. 42	0. 685	0. 196	0. 000	0. 01	0. 641
	SD	0. 884	0. 464	0. 403	0. 000	0. 1	0. 775
Cluster 3	Mean	1. 271	0. 347	0. 391	1. 000	0. 000	1. 061
	SD	0. 764	0. 476	0. 549	0. 000	0. 000	0. 737

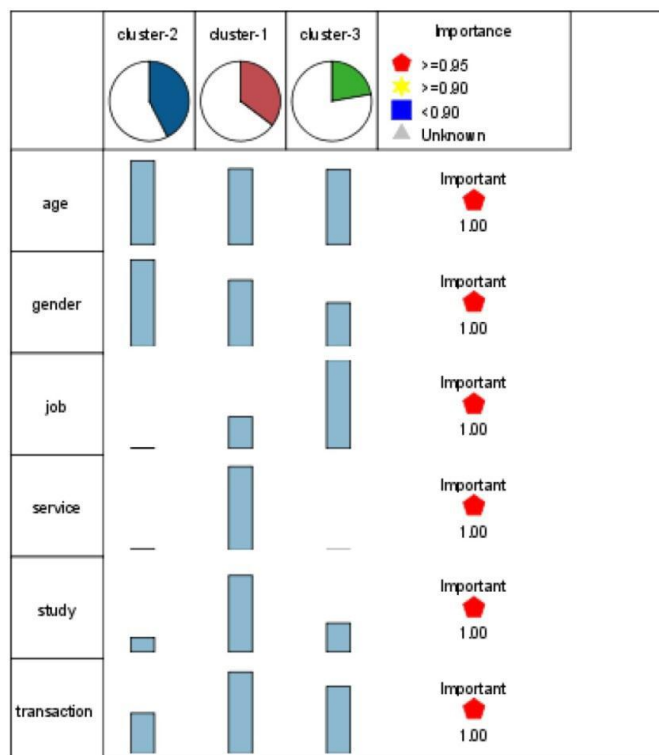


Fig 11. Clustering the customers using two step algorithm.

Conclusion

What are the demographic features of the bank's target customers (i. e. those customers who have high financial transactions)? Cluster 1 with transaction mean of 1. 291 and standard deviation of 0. 697 is the cluster of the bank's target customers. This cluster has 2543 records with the following features: a) Age average of this cluster is 1. 285 and standard deviation of 0. 746 that shows the age range of this cluster is 25-40; b) Gender average of this cluster is 0. 526 and the standard deviation is 0. 499 that shows this cluster contains female and male samples; c) Job average of the customers is 0. 355 and the standard deviation is 0. 479 that shows most of the customers have jobs in business.

Is there a same financial pattern (in terms of financial transactions) in demographic features of the bank's target customers? In order to find the response to this question, we used decision-making tree to determine the demographic patterns of the customers and predict the future possible customers and their transaction rate. The results are as follows: a) Those whose education degree is BA or higher and their age range is 40-55 have in 71% of the cases, the highest rate of transactions; b) Those male cases who work in business jobs, and their age range is 25-40, with BA education and higher, have a high rate of financial transaction in 54. 173% of the cases.

What are the demographic features of the bank's special customers (i. e. those who have used new services of the bank)? In order to identify special customers of the bank, we used the clustering method. Cluster 1 with new service average of 0. 944 and standard deviation 0. 23, is the cluster with this type of customers which includes 2543 records with the following characteristics: a) Age average of this cluster is 0. 526, standard deviation 0. 499 that shows this cluster includes both male and female cases; b) Education average is 1. 037 with a standard deviation of 0. 486 that shows most of the members of this cluster have BA; c) Job average of this cluster is 0. 355 with a standard deviation of 0. 479 that shows most of them work in business jobs; d) Average of customers who use new services, is 0. 944 with a standard

deviation of 0. 23 that shows most of them use new services; e) Financial transactions average is 1. 291 with a standard deviation of 0. 697 that shows most of them had an average and good transaction rate.

Is there a same pattern (in terms of response to new service) in demographic characteristics of the bank's special customers? Results of tree technique are as follows: a) Those with BA degree and higher in the age range of 25-40 answered positively to the new services of the bank in 77. 55% of cases; b) Those with MA degree and higher, in age range > 55 answered positively to the new services of the bank in 91. 176 of cases; c) Those with BA degree and age range of 40-55 answered positively to the new services of the bank in 58% of cases.

What are the clusters of the bank's customers of ATM service by using the data mining technique?

Results are divided into three clusters. Cluster 1 with 2543 transactions (35% of the total transactions) includes both target and special customers of the bank. Cluster 2 with 3061 transactions (42% of the total transactions) includes customers with middle to bad transactions who didn't use new services of the bank.

Cluster 3 with 1595 transactions (22% of the total transactions) includes customers with middle transactions who didn't use new services of the bank.

7. References

1. Al-Mudimigh, A.S., Saleem, F., Ullah, Z., & Al-Aboud, F. N. (2009, August). Implementation of Data Mining Engine on CRM-improve customer satisfaction. In Information and Communication Technologies, 2009. ICICT'09. International Conference on (pp. 193-197). IEEE.
2. Berry, M., & Linoff, G. (2000). Data mining Techniques. John Wesley and Sons, Inc. USA: In Bicen, Pelin and Oktay Firat, SU (2003). "Knowledge Discovery in databases (KDD) and Data Mining: An application of customer segmentation analysis in banking sector". In Bulletin of the International Statistical Institute, 54th Session (Vol. 60, p. 136).
3. Brown, S.A., & Coopers, P.W. (1999). Customer relationship management: A strategic imperative in the world of e-business. John Wiley & Sons, Inc.
4. Chatterjee, J. (2000). Managing customer relationships in the e-business economy.
5. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), 37.
6. Feinberg, R., & Kadam, R. (2002). E-CRM Web service attributes as determinants of customer satisfaction with retail Web sites. International Journal of Service Industry Management, 13(5), 432-451.
7. Hendricks, K.B., Singhal, V.R., & Stratman, J.K. (2007). The impact of enterprise systems on corporate performance: A study of ERP, SCM, and CRM system implementations. Journal of operations management, 25(1), 65-82.
8. Hsu, C.H. (2009). Data mining to improve industrial standards and enhance production and marketing: An empirical study in apparel industry. Expert Systems with Applications, 36(3), 4185-4191.
9. Hua, J., & Zhenxing, C. (2009, May). Optimization of Data Mining in CRM Based on Rough Set Theory. In Information Technology and Applications, 2009. IFITA'09. International Forum on (Vol. 3, pp. 252-255). IEEE.
10. Kantardzic, M. (2011). Data mining: concepts, models, methods, and algorithms. John Wiley & Sons.
11. Larose, D.T. (2006). Data mining methods & models. John Wiley & Sons.

12. Shahin, A., & Nikneshan, P. (2008). Integration of CRM and QFD: A novel model for enhancing customer participation in design and delivery. *The TQM Journal*, 20(1), 68-86.
13. Thearling, K. (2000). Data mining and customer relationships. Building data mining applications for CRM.

- New York, NY: McGraw Hill. Available: <http://www3.shore.net/~kht/text/whexcerpt.htm> (April 15, 2002).
14. Turk, M.A., supervisor: Soilen, K.S. (2007), Data Mining and Mobile CRM: An Update, A thesis submitted in partial fulfillment of the requirements for the degree of MBA (Master of Business Administration), MBA DISSERTAT.