Presentation of customer foreclosure model of e-banking services using decision tree and data mining method

Zahra moradzadeh , Behnaz khodayari

Department of business management, Islamic Azad University, firoozkooh branch, firoozkooh, iran

Abstract

In this research, we use the methods of data mining to discover the pattern of customer reversal of banking services. The statistical community of the Ayandeh bank customers who have been active in the provision of electronic services, among which 149 people were selected by sampling method in West Tehran's branches. In this research, three tree decision algorithms for decision tree C & R TREE, QUEST TREE and CHILD TREE are used to predict the deviations, reversal patterns and the most important attributes that affect it. Based on the results of decision tree algorithms, these algorithms with acceptable accuracy (above 80%) can predict customer behavior. Based on the results, the C & R Tree decision tree algorithm, better than other algorithms, can turn customers away Predict Also, based on decision trees, and considering the percentage of deviations in each of the nodes, it is possible to discover the laws that lead to customer rejection. In this regard, it is recommended that these rules be applied in marketing and customer retention guidelines. Also, based on the results of this study, five important attributes for predicting customer rejection are, respectively, occupation, branch grade, education, inventory grade and type of investment that banks should pay particular attention to.

Key words: customer satisfaction, customer rejection, data mining, customer relationship management.

1. Introduction

With the increasing expansion of data mining tools in the stock industry, financial and accounting forecasting, e-commerce and medical forecasting is essential for this tool to play a role in managing customer relationships between organizations and industries, including in industries that are predicted by the market , Logistics, customer, finance, risk, etc. The use of statistical methods and reporting alone does not provide managers with the necessary information to make the best decision. In a world full of modern changes in customer retention or in other words, marketing activities are one of the most important activities of each company, it is necessary to use new tools for discovering knowledge of customer relationship management in order to analyze customer life cycle. It can be said that customer retention depends on the correct use of the information at the right time and the provision of a suitable offer to individuals. (Shahrabi and Shojaee, 2009: 45) By allocating a budget for the modernization of knowledge and information discovery, significant return on investment can be achieved by maintaining the company's loyal and

satisfied customers. Today, the blindfold competition of prices or the increase and improvement of services for customers is not enough. They should recognize the customers and anticipate the next moves based on their behavioral characteristics and offer acceptable offers to the customers. Today, the survival of organizations depends on their intelligence, and data mining is one of the tools for gaining this intelligence in a constantly changing world. Understanding these factors and foreclosing clients is important for the bank, so that the bank can stop those factors by knowing the rules of switching away from electronic services. Therefore, in order to predict customer behavior, they must collect the appropriate information and use them through science and technology properly analyzed. In this research, the data mining and its ability to predict customer behavior by using customer relationship management tools and data is analyzed in this paper. Based on this, by creating the proper database of the former clients of the Ayandeh bank (given our placeholders, age, gender, job, number of bank accounts, type of services provided from the bank side, etc.) and their current status in three stages, the discovery and proper pattern of data are extracted. In this research, after defining the problem and classifying the former customers based on the current status of the relationship with the bank, is determined using cross-validation and database validation algorithms. Also, the best classification algorithm is determined to predict the depreciation of each customer from banking services in the next few years, and the most important attributes affecting the behavior of customers that the bank can focus on on them can maintain the customer is determined and ultimately the rules will lead to a rollback Customers are determined using tree decision algorithms C & R TREE, QUEST TREE and CHILD TREE.

2. Theoretical Basics of Research

2.1. Data mining

In order to systematically guide data mining analysis, a general process is usually followed. There are a number of standard processes, two of which are described here. One is Crisp: an industry standard process that includes steps to study data mining. Although every step of data mining in any analysis is not required, the steps give it the right steps for data mining. The steps that began with the discovery of data include data collection, data processing, analysis, and deduction and implementation (Alizadeh et al., 2008)

Crisp-Dm process:A cross-industry standard industrial process for data mining that is widely used in the industry.

Banking services

One of the essential tools for the realization and expansion of e-commerce is the existence of an electronic banking system that, in keeping with global financial and monetary systems, facilitates e-business activities. In fact, the implementation of e-commerce requires the realization of e-banking. For this reason, the use of electronic systems in the world's financial and credit institutions has grown rapidly and the number of users of e-banking services is increasing. (Pourzandi and Najafi, 2012)

Over the last few decades, there have been tremendous developments in the banking system; these can be subdivided into four periods, and to some extent computers and software are replaced by humans and paper. In other words, new technology and e-banking will allow them to increase the speed, quality, accuracy, cost and variety of services. In the following four stages of the evolution of the banking system are presented:

Automation of the back door,- Front-end automation, Connect customers to their accounts, integrating the systems and connecting the customer to all banking operations. (Sanayee & Salehnia, 2008)

3. Research background

Mirmohammadi and others. 2017 Customer Deviation Forecasting Using Decision Tree in the Banking Electronics Industry In this research, Crisp algorithm and decision tree method at the modeling stage are used. The results of modeling the characteristics of customer diverting from e-banking services are shown.

Ashoori 2017 A hybrid algorithm for predicting customer rejection The results indicate the importance of variables among customers and their Internet service providers should be considered for Ayandeh plans. Also, the inclusion of advertising variables in the decision tree does not indicate a significant effect on customer rejection and only affects the attraction of new customers.

Emami and others in 2016 provided a model for the removal of customers from the National Bank using structured equations based on the results obtained from this research. It was determined that management factors as the most effective factors including internal factors and ethnic considerations and tribalism as the most effective factor in external factors. Environmental) and the exploration issue of this research. This means that the studied bank should pay attention to the reasons mentioned above.

Mousavi 2016 Modeling Dealing Behavior of Real Customers at Mellat Bank Using Machine Learning Techniques The results show that customers are less likely to be members of the club than other members of the Club, have no accounts, did not use bank facilities, and did not have any other account. In the second stage, a random forest technique that is suitable for behavioral modeling is used to classify and extract the rules. This technique averaged 96% of the correct prediction of class members based on the characteristics of the local profile. To evaluate the results of the study, the comparison with the accuracy of the prediction of the fuzzy inference system was considered. Random forest with better prediction accuracy, faster runtime and more comprehensive is a more appropriate technique for predicting customer rejection.

Amiri and others. 2014 Forecasting the behavior of bank customers in order to identify the appropriate method for offering the proposed services. In this research, the results of predicting the Ayandeh trend of customer transactions (positive or negative trend or behavioral pattern change) by expressing perceptron neural networks in identifying The time is right to suggest the services of interest. These services have been identified based on customer preferences and results from behavioral model prediction.

Sepehri and others 1011 The discovery of customer rejection reasons for banking services by combining data mining techniques and survey research. In this paper, decision tree techniques with the target variables

of the offset label are used. Validation of the results is obtained by testing it on the validation data and calculating the lifting rate and overall error. The results indicate a tendency to decline in a large segment of the Bank's agricultural customers, the most important of which is its impact, especially in higher-income groups.

Sayed Hosseini and others 2011Designing an RF-based Methodology to Measure Customer Loyalty Using Data Mining Techniques This paper presents a new model based on the development of an RFM model and the determination of weights of parameters by pairwise comparison after clustering of products with The number of optimized clusters by the Davis Boulder Index has been designed to determine the degree of customer loyalty in a winning and winning concept.

Tavakoli et al. 2010 Using Data Mining Process to Prevent Customer Deviations in Insurance This article aims to explain the capabilities of data mining in customer rejection management using the standardized data mining method of CrispDem to explore the databases of one of the companies The public interest insurance company has been involved in the field of fire. The results show that the customer recruitment channel is the main predictor of customer rejection or customer survivability in the company, and in the following cases, the purchase history and the use of the insured place are considered as predictive factors of the switch. Get it

Mehrabi and others in 2010 provided an integrated model for implementing the concept of customer relationship management at Mellat Bank. In the Mellat Bank in this study, the analysis was performed using statistics. To analyze the data, frequency distribution, central indexes and distribution, Pearson correlation analysis and inferential SPSS and software usage and Friedman test have been used for ranking variables. According to the data analysis, the results of the research showed that: t changes in organizational culture, technological changes and changes in organizational structure have a significant and positive effect on successful implementation and implementation of customer relationship management.

Calgat and Stewart 2017. Identifying and analyzing the factors affecting customers' rejection in their research showed that high income and younger (young) customers are more likely to turn away and adopting an appropriate strategy to reduce the rate of escape to customers is required.

Yo et al., 2017 provide a model for predicting customer departures from their banking services, suggesting the method of extended vector vectors compared to the three methods mentioned. This algorithm has the advantage of examining unbalanced and nonlinear samples.

Nai et al., 2016 show that the customer rejection model using two logistic regression logic algorithms and Decision Decision tree show that demographic information has a smaller share in customer rejection and the performance of the regression method is slightly better than the decision tree.

ABDO and DIYGAR 2016 Customer Credit Risk Estimation and Continued Customer Relationship Using Decision Tree Algorithms Based on the results obtained from this research, the CART decision tree algorithm has the highest accuracy in predicting customer behavior and their relationship with bank.

Karmety et al., 2016. Predicting the behavior of bank customers using decision tree algorithms in their e-services department. In this paper, decision tree algorithms were used. Based on the results obtained from the decision tree tree algorithms, they are able to: The prediction of customer rejection model using Artificial Neural Network Technique shows the results of Artificial Neural Network Technique for finding the best model of customer stored data for the prediction of deviations Better suggestion. Father and Eliade 2015 An effective data mining framework for predicting customer behavior In this paper, a model is developed to predict customer behavior in order to enhance decision-making processes for maintaining valuable customers. Also, in this paper, an efficient data mining framework and two proposed models Naio Bays and Neural Networks have been studied. The results show that the accuracy of the neural network is relatively better. Wireline and MirrorSchedel 2013 The use of data mining techniques to improve customer relationship management is the target of identifying potential customers who are willing to buy and service The data collection of this research is the actual data of the companies The applied records and data sets were first pre-processed after the preprocessing stages using CAID and C5.0 decision trees for classifying customers and helping the decision making of the organization concerned. The results of the two tree decision methods say that the more important characteristics Azziova et al. 2012 Predicting Customer Behavior in a Multi-Service Finance Organization: Profitability Perspective In this paper, an adaptive segmentation approach that involves identifying a "neighborhood" using the measure of similarity and definition A variable space is provided. Future revenues for each customer are projected using the same probability distribution forecast based on the exhibitor's clients' behavioral characteristics similar to those of previous periods. The model for development and implementation for retail banks in the UK shows a good comparison with other benchmark models. Maggoes et al. 2012 Customers Data Mining for Life Style Segmentation This article is based on the lifestyle of vendors and with extracted information support. From a large trading database, it offers a method for segmenting the retail market. Using a variable clustering algorithm, a set of common shopping baskets was extracted from the database and used to find out the lifestyle of customers. Customers were assigned to a lifestyle section based on their purchase history. This study was conducted in collaboration with a European retailer. Anjie et al. 2009 The use of data mining techniques in customer relationship management The focus of this paper is to examine and categorize data mining processes and techniques. Results show that research Criteria and customer relationship modeling and data mining are a rescue tool to guide future research and facilitate knowledge accumulation in the use of data mining. Keymones and Gao 2008 Consumer Awareness And different customer behaviors This article is devoted to how to get a more accurate understanding of how How online reviews are associated with auction and product sales. Recent applied studies suggest that consumers use different ways in purchasing experiences. And the variety of consumer purchasing behaviors during the various online shopping experiences is the best predictor of the success of different online products, depending on what consumers are buying, what they are buying and how they buy.

Research method

The method of this research is applied in terms of its purpose, since it contains a set of valid and reliable rules and tools to investigate the facts and discover the unknowns and find solutions to problems that can be used by one or more organizations. Since this research examines the status quo, it is descriptive in the field of research, and as it discusses the prospect of a customer turning away from banking services, it is exploratory.

4. Information Collection Method

In this research, information is gathered through library studies that includes documents in libraries, research, papers, and dissertations. Also, field methods are used to collect the necessary information.

5. Data and information gathering tools

Since this research is an applied research, the required information should be collected fieldwise. In collecting field information in this research, two interviewing tools and a database in the form of bank customer information that includes an account opening form and the use of electronic services are used. The interview used to determine the basic criteria affecting the prediction of customer behavior. And the database is used to collect the information needed by customers.

6. Society and statistical sample

In this research, the statistical society includes all customers who have been active in the future bank in the years 1394 to 1396. Also, a sample of these customers is selected using available sampling method, which is the number the sample is 149 people.

-7Information Analysis Method

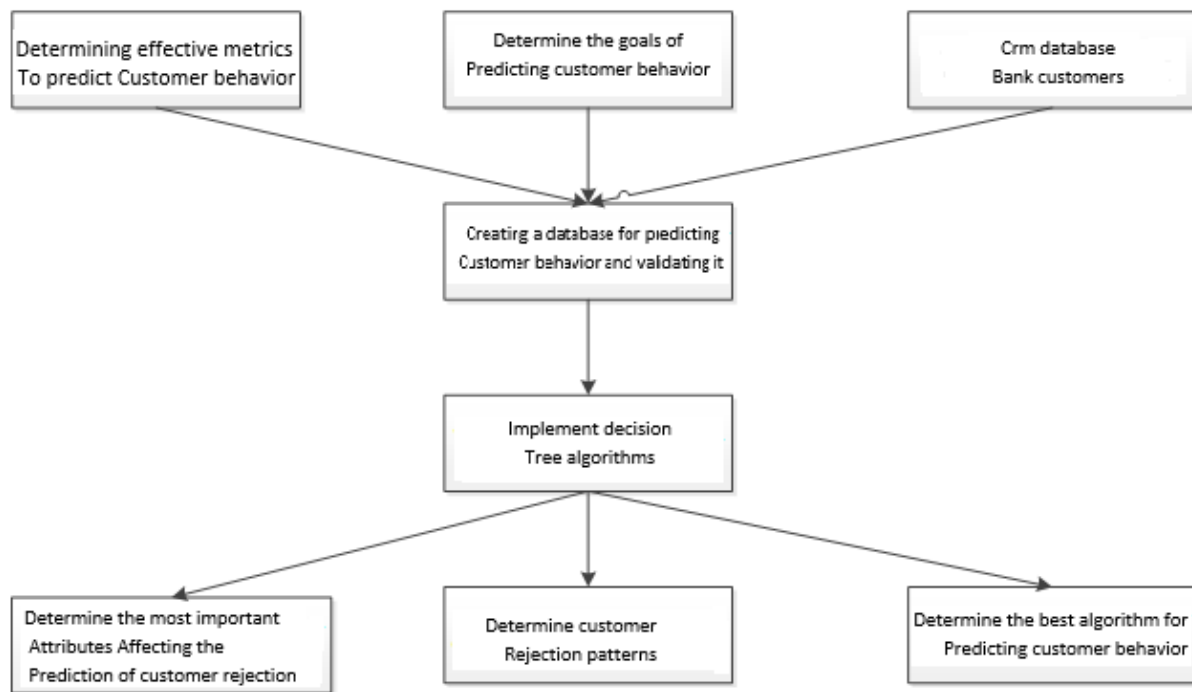The method for analyzing information in this study is in accordance with Fig. 2.

Figure 2 Information Analysis Method

8. Findings

The research is carried out in three main phases, in which we will conduct a research methodology according to these phases.

8.1. Creating a Database

1. Determine the scope of the problem

At this stage, the main issues are determined and, according to the goals of the problem, the information should be extracted. The problem in this research is to predict the real customers' reversal of the future bank of West Tehran Branches between 1394 and 1396, when available customers were selected.

2. Determine records

At this stage, should actual customers that want a reason to abandon them explore and predict we determine that the information 149 customers real randomly chosen from their information will be used and the validation of 10 can be about The number of data adequacy has been commented.

3. Specify the attributes

At this point, the basic attributes of each customer, including quantitative and qualitative items, are defined as follows:

1. Branch degree

2. Gender

3. Age

4. Education

5. Occupation

6. Type of investment

7. Amount of investment

8. Period of investment period

9. Rate of inclination

10. Return on investment

11. Duration of opening account

12. Account history

13. Number of accounts

14. Balance of inventory

15. Housing situation

4. Specify the class (s) in the database

At this stage the most important measure of reluctance that banks are willing to predict about each customer is true is determined that the tendency to do or unwillingness to invest client that the client into one of three categories on both databases put:

Customers who are likely to sideline electronic activity:

These customers have registered the rate of positive electronic activity in the Ayandeh bank. (During the four months of this year, they had at least 10 electronic activities per month.(

Customers with a moderate degree of electronic activity:

These customers have registered the rate of neutral electronic activity in the Ayandeh bank. (During the 4 months of this year, at least 3 and up to 9 electronic activities per month.(

Customers who are likely to turn away a lot of electronic activity:

These customers have registered the rate of negative electronic activity in the Ayandeh bank. (During the 4 months of this year, they had a maximum of 2 electronic activities(.

8.2. Database Validation

In this study, the initial selection character by experts nominated bank and so should the validity of the attributes tested in this study database consider using validation cross towards validation feature on the database Action

Table 1: determines the validity of the database with 10-point crossover validation

algorithm	Correct prediction	Wrong prediction	accuracy
C&R Tree	133	16	89
Child Tree	125	24	83
Quest Tree	129	20	86

8.3. Implement decision tree algorithms on the database

At this stage, we will predict the probability of disruption by implementing each decision tree algorithm .using the Clementine software

C & R Tree tree algorithm

By implementing the C & R Tree algorithm, the most important attributes are shown in Figure 3.
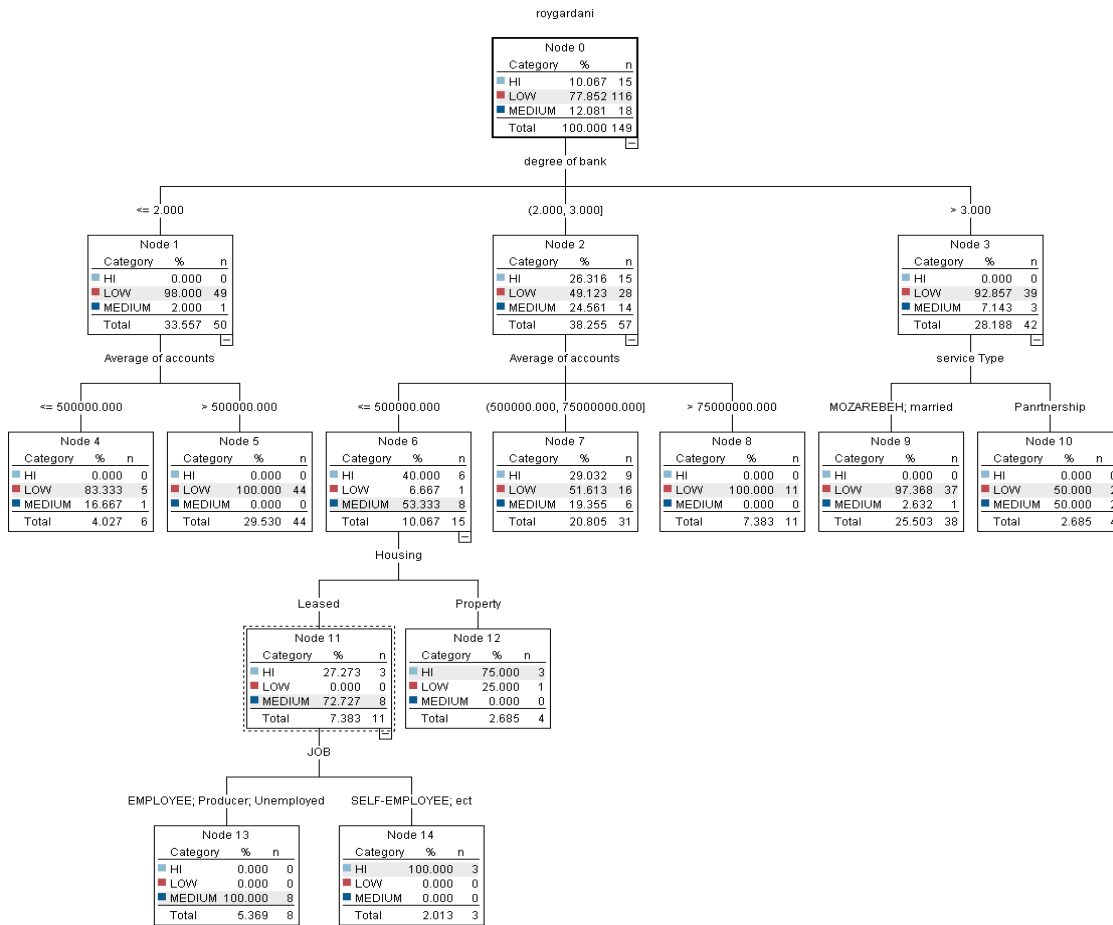
International Journal of Research

Available at https://pen2print.org/index.php/ijr/

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 21
October 2018

Figure 3: The most important attributes using the C & R Tree algorithm

Figure 4 :C & R Tree tree with a depth of 6

Child Tree tree algorithm

By implementing the Child Tree algorithm, the most important attributes are shown in Figure 5.

Figure 6 Child Tree tree with depth 5

Quest Tree tree algorithm

With the implementation of the Quest Tree algorithm, the most important attributes are shown in Figure 7.
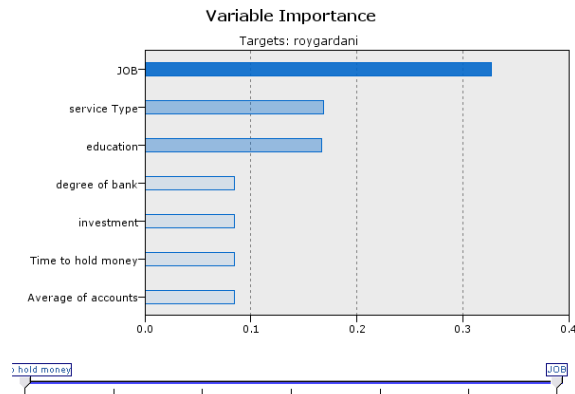
International Journal of Research

Available at https://pen2print.org/index.php/ijr/

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 21
October 2018

Figure 7 The most important features using the Quest Tree algorithm
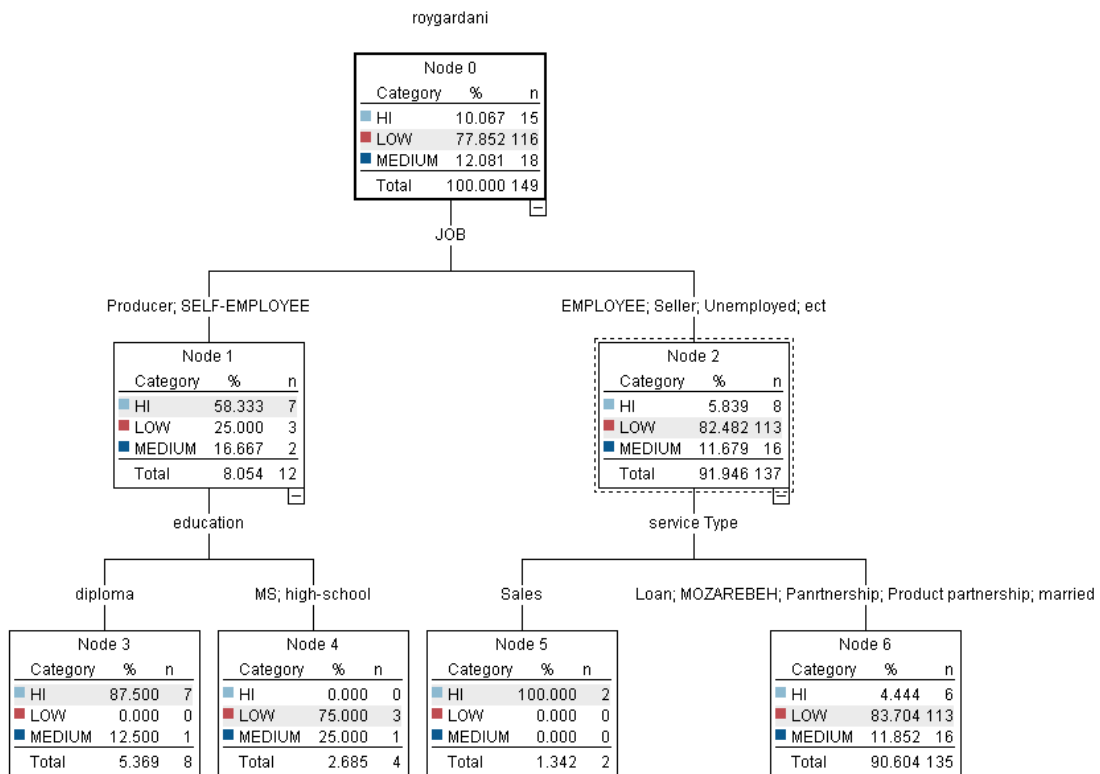

Figure 8 Quest Tree tree with depth 3

9. Discussion and Conclusion

In this research, the main goal is based on the discovery of customer reversal pattern in the Ayandeh bank of selected branches of West of Tehran, which is carried out using data mining tools. Initially, a database is created, based on which important attributes are assigned and for each customer (149 actual customers) each attribute is set up. Accordingly, the 15 characteristics selected are: Branch-gender-age-education-occupation-type of investment-investment amount-investment period-rate of willingness-return on investment-time of opening an account-account history- Number of Accounts - Inventory Average and Housing Condition. Also, the segmentation of Ayandeh bank customers is also based on their investment volume, which are grouped into three categories, which are: customers who are likely to shed much of their electronic activity, those who are likely to have a moderate degree of electronic activity, and customers who are likely to They have a lot of electronic activity. Based on the results of decision tree algorithms, these algorithms with acceptable accuracy (above 80%) can predict customer behavior. Based on the results, the C & R Tree decision tree algorithm, better than other algorithms, can turn customers away Predict Also, based on decision trees, and considering the percentage of deviations in each of the nodes, it is possible to discover the laws that lead to customer rejection. In this regard, it is recommended that these rules be applied in marketing and customer retention guidelines. Also, based on the results of this study, five important attributes for predicting customer rejection are, respectively, occupation, branch grade, education, inventory grade and type of investment that banks should pay particular attention to.

Based on the results obtained from the implementation of decision tree algorithms, the C & R Tree algorithm has the highest accuracy in predicting customer reversal. In this regard, it is suggested that the bank, with the implementation of this algorithm on its customers, is likely to turn away Identify.

The most important criteria in forecasting the decline of the customers are the job and grade of the bank. In the case of a job, it should be noted that those who have a non-official job are more likely to be deprived of banking services. Investigating decision tree algorithms and databases shows that these people usually receive banking services but are not willing to work with the bank as co-workers.Also, regarding the degree of branches, it is noteworthy that most reversals have been carried out in grades 2 and 3, which should be given special attention to these branches.A review of various database rules suggests that in some cases, customer rejection may reach 100 percent, including:If the job is individual, employee, seller, unemployed, etc., and the type of service is discounted, then the likelihood that the e-banking service will withstand a high probability is 100%.If a person has a job as an employee, unemployed, etc., and the bank's grade of <3,500 then the likelihood that the e-banking service will be turned off is likely to be 100%.If 2 = Bank = 3 and the average of accounts is <75000000.0, then the likelihood that the e-banking services will be turned off with a low probability is 100%.If 2 = Bank = 3, and the average of accounts is 500,000.0, and their housing status is rented and the job is employee, producer or unemployed, then the probability that the e-

banking service is likely to be turned off is 100%.If 2 = Bank = 3, and the average of accounts is 500,000.0, and their housing status is rented, and the occupation is either self-employed, or so, then the likelihood of a high probability of turning off the electronic banking service is 100%.Based on these nodes, the approach should be considered so that the customer does not even fall in this situation in order to reduce the likelihood of a reversal, and if the bank can not prevent a reversal, it will provide such facilities, such as a loan, to such a customer by obtaining appropriate documents do it.In order to do this research, other researchers are recommended to take steps towards:

Review each of the five key customer attributes (occupation, branch size, education, inventory grade, and type of investment) and how to deal with them and different customers, so that they are conducted on any practical feature from the bank's side so that the customer does not go away.Implementation of this research in other banks and compare results with the results of this research.Use of other learning algorithms including neural network and compare results with the results of this research.Assessing the rules obtained using various techniques, including data envelopment analysis model.Checking the Ayandeh Bank's Customers' Lifestyle.Forecast the opening time until closing the account in the Ayandeh bank

References:

Abdou, H,2016, Consumer behavior and marketing action, Boston: MA, USA: PWS KENT Publishing Company.

Ahmadi P, Poorashor Yasanaleh, 2011, Customer Satisfaction Path in Commercial Banks, Moderator of Humanities, No. 11, 29-58

Alizadeh S., Timurpour B, Ghazanfari M., 2008, Data mining and knowledge discovery, Iran University of Science and Technology.

Amiri, B., Shakouri, H., Mousavi, M., Shawardi, M., 2011, Predicting the behavior of bank customers to identify the appropriate method for offering proposed services, knowledge and technology, Year 1, No. 4.

Angi, A., Money, A. H. & Berthon, P. R,2009, Service quality and Satisfaction. European Journal of Marketing, 34 (11), pp. 1338-1352.

Asgarzadeh S., Timurpour B, Ghazanfari M., 2009, Data mining and discovery, Tehran, Iran University of Science and Technology.

Ashouri, A, 1396, A hybrid algorithm for predicting customer rejection: Case study of Internet companies, Master thesis, Tarbiat Modarres University.

Audzeyeva,A, Summers,B , Schenk-Hoppe,K,2012, Forecasting customer behaviour in a multi-service financial organisation: A profitability perspective, International Journal of Forecasting, Volume 28, Issue 2, April–June 2012, Pages 507–518

Bahari,F.T.,Elayidom,S.M.,2015, An Efficient CRM-Data Mining Framework for the Prediction of Customer Behaviour, Procedia Computer Science 46, 725 – 731

Clemons,E, Gao,G,2008, Consumer informedness and diverse consumer purchasing behaviors, Electronic Commerce Research and Applications, Volume 7, Issue 1, Spring 2008, Pages 3–17.

Colgate, M. & Stewart, K,2017, Customer defection: A study of the student market in Ireland", International Journal of Bank Marketing, 14 (3), pp. 23-29.

Ellsson De, Delane D, 2010, Advanced Data Mining, Kerman, University of Kerman Province, First Edition.

Emami, L, Poorashorf, Y. Tolbabi, Z, 2016, A Model for the Reduction of Customers from the National Bank Using Structured Equations, Marketing Management, Volume 10, Issue 26, 25-46.

Industries, A., Salehnia, M., 2008, Identification and ranking of barriers to the acceptance and development of electronic banking, the second international conference on electronic banking.

Jenabi,G.,Mirroshandel,S.A.,2013, Using Data Mining Techniques for Improving Customer Relationship Management, European Online Journal of Natural and Social Sciences, Vol.2, No.3 Special Issue on Accounting and Management. ISSN 1805-3602

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozzafari, M., & Abbasi, U,2016, Improved churn prediction in telecommunication industry using data mining techniques. Applied Soft Computing, 24, pp. 994-1012.

Mehrabi, J, Babaei Ahari, M, Taati, M, 2010, Introducing a Unified Model for Implementing the Concept of Customer Relationship Management at Mellat Bank, Development Management Magazine, Volume 4, 61-71.

Migueis,V.L.,Camanho,A.s., Falcao e Cunha,J.,2012, Customer data mining for lifestyle segmentation, Expert Systems with Applications, Volume 39, Issue 10, August 2012, Pages 9359–9366

Mirmohammadi, M., Ghanei, H., Keramati, A., 2016, Customer Deviation Forecast Using Decision Tree Method in Electronic Banking Industry (Case Study: Bank Ansar), Quarterly Journal of Monetary and Banking Management, Vol. 3, No. 6, 1 -23

Mousavi,2016, Modeling the Deviant Behavior of Real Customers in Mellat Bank Using Machine Learning Techniques, Master's Thesis, Kharazmi University.

Nie, G., Rowe, W., Zhang, L., Tian, Y., & Shi, Y,2016, Credit card churn forecasting by logistic regression and decision tree. Expert Systems With Applications, 38, pp. 15273-15285.

Pourzandi, A., Najafi, M., 1391, Investigating the Effect of Quality of Electronic Banking Services on Customer Satisfaction, Researcher. Management Quarterly, ninth year, No. 26, 29-41.

Sepehri, M., Nowroozi, A., Timurpour, B., Maghubar, S., 2011, Disclosure of Customer Reversal Reasons for Banking Services by Combining Data Mining and Survey Research Methods, Management Researches in Iran, Volume 15, Number 4, 97-126.

Seyed Hosseini, S., Gholamian, M., Maleki, A., 2011, Designing a RFM Based Methodology to Measure Customer Loyalty Using Data Mining Techniques, International Journal of Industrial Engineering and Production Management, Issue 2, Volume 22, 172-179

Shahrabi, J, Shojaee, AS, 2009, Advanced Data Envelopment Concepts and Algorithms. University Jihad (Amirkabir University of Technology).

Shahrabi, J., Shakurnas, M., 2008, Data mining. University of the Amirkabir University of Technology.

Sharma, A., & Panigrahi, P. K,2016, A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services. International Journal of Computer Applications, 27(11) , pp. 26-31.

Tavakkoli, A. Mortazavi, S., Kahani M., Hosseini, Z., 2010, Application of the data mining process to predict customer rejection patterns in insurance, Business Management Outlook, No. 4. 51-55

Yu, X., Guo, Sh., Guo, J., & Huang, X,2017, An extended support vector machine forecasting framework for customer churn in e-commerce. Expert Systems With Applications, 38, pp. 1425-1430.