Investigating Factors Affecting the Cost of Money in Iranian Banks Based on Artificial Intelligence and Using Data Mining

Ali Haghi Nojehdeh[1], Mansour Esmaeilpour[2]*, Behrooz Bayat[3], Alireza Isfandyari Moghaddam[3]

***Abstract*— One of the potential goals of companies, including banks, is to earn profit and cover current expenses. Iranian banks have also been affected. Calculating and understanding the cost of money is very important because, by calculating and analyzing it, you can estimate and implement the amount and price of paying bank facilities as well as interest on deposits. Considering that the cost of money represents the correct management of the resources and costs of a bank, the investigation of ways to reduce the cost of money in state-owned banks can be an indicator of the efficiency of managers and the performance of a bank.The best way to calculate the optimal cost is to use data mining techniques. In this research, decision tree models, Bayesian rule, neural networks, and the RoughSet model have been analyzed using data mining methods through Weka, Rosetta, and Excel software.The accuracy criterion was using decision tree J48 (0.919), Bayesian theory (0.843), neural networks (0.274), and rough set model (0.0952), which was in the form of a law by genetic algorithm, Johnson, Holt was presented. These rules enable bank managers to adopt policies based on the discovered models to better understand their resources and costs and to balance finances in their branches to achieve better value for money.**

Index Terms*— *Cost of money, Artificial intelligence, Data Mining, Neural Network, Decision Tree, Bank.

I. INTRODUCTION

One of the most important parameters and variables in every economy is the interest rate. Government officials and lawmakers change interest rates for various purposes: controlling liquidity, inflation, and prices; economic growth and development; lending, etc. So, it is important to set the interest rate correctly. If you can predict the interest rate correctly, you can earn and gain profit by investing in various sectors (Shahvaroughi Farahani, 2021). The most important financial resource of banks is individual deposits. Therefore, collecting and acquiring deposits and allocating them to meet the financial needs of various economic activities is one of the most important banking operations. Therefore, banks should know the cost of money (resources), i.e., the costs that banks have to equip and manage resources, such as Qardh al-Hasan savings deposit reward, interest on short-term and long-term deposits, non-operating costs, bad debt, and depreciation, to continue the activity and preserve the resources of depositors and make the right decision.

In this research, the author has done extensive analysis and study on banking data using data mining techniques to obtain the factors affecting the cost of money of Iranian state-owned banks. Here it is necessary to provide a brief description of data mining. The concept of data mining includes algorithms and methods that extract information from data (Cleary, 2006). Nevertheless, collecting and storing large amounts of data can be a waste unless the data is used profitably and becomes a source of funding for the organization. Many organizations have turned to data mining to turn this potential value into strategic information (Yousofi, 2016), as it will be possible to discover hidden relationships, trends, and models between data and gain new knowledge about the overt and covert challenges of the organization through data mining.

This study is based on 501 branches of the National Bank of Iran and other branches of state-owned banks in Iran to predict the factors affecting the cost of money using the decision tree, Bayesian rule, and neural network model as a data mining technique.

One of the essential factors that affect the increase in the banking services price and payment facilities by banks is the cost of money in the banking system. The increase and decrease of this price also depend on many reasons. According to banking experts and activists, one reason for the high cost of banking facilities is the rising cost of money in banks. Customers and economic activists have always criticized the banking system for expensive banking facilities and their other problems (Fallah Shams, Mahdavi Rad, 2012). Determining the factors affecting the prime cost depends on the type of organization and its activities.

The cost of money reflects the proper management of the bank's resources, consumptions, and expenses, and banks and financial institutions, like any other firm, seek profitability and increase the wealth of owners, in addition to providing services. The nature of the banks' activities deals with concepts such as credit, payment systems, and various rates, making such institutions more serious in the proper management of deposits, which are the most critical resources of the bank. Banks should manage and maintain deposits to meet the needs of depositors at different maturities, meet the needs of loan customers, and deliver the required funds to borrowers at a reasonable rate (Attaran, Divandari, Adinov, 2012). The most crucial financial resource of banks and financial institutions is the deposits of individuals. Therefore, collecting and acquiring deposits and allocating them to meet the financial needs of various economic activities is one of the most important banking operations.

Banks need to be aware of the cost of money (resources),

1. Department of Management, Hamedan Branch, Islamic Azad University, Hamedan, Iran.
2. Department of Computer Engineering, Hamedan Branch, Islamic Azad University, Hamedan, Iran.
Corresponding author: esmaeilpour@iauh.ac.ir
3. Department of Knowledge and Information Science, Islamic Azad University, Hamadan Branch, Hamadan, Iran.

i.e., the costs that banks have to equip and manage resources, such as Qardh al-hasan savings deposit reward, interest on short-term and long-term deposits, non-operating costs, bad debt, and depreciation to continue the activity and preserve the resources of depositors and make the right decision (Hasanvand, 2016).

A decision tree is a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root" that has no incoming edges. All other nodes have exactly one incoming edge. The tree's complexity has a crucial effect on its accuracy. Decision tree induction is closely related to rule induction. Each path from the root of a decision tree to one of its leaves can be transformed into a rule simply by conjoining the tests along the path to form the antecedent part and taking the leaf's class prediction as the class value (Rokach & Maimon, 2005).

The resulting rule set can then be simplified to improve its comprehensibility to a human user and possibly its accuracy (Quinlan, 1987).

Usually, in the previous methods, the decision tree and related algorithms of card, child, and c5 were used, as in the articles of (Zhu et al., 2018) and (Moslehi, Haeri & Gholamian, 2019) and (Rahi, Ebrahimnejad & Motameni, 2024) and (Ayinla & Akinola, 2020), but in this research, the j48 decision tree, which is related to the c4.5 algorithm, was used along with other algorithms to examine and discuss different angles.

The reason for using the j48 decision tree method in this research according to the articles of (Yang & Zhang, 2020), (Aljawarneh, Yassein & Aljundi, 2019), (Cruz & Tumibay, 2019), and (Ahishakiye, Taremwa, Omulo & Niyonzima, 2017) who have used the j48 decision tree method and its results show that compared to other algorithms, this method increased the segmentation criteria by pruning and optimization up to 90% and produced a smaller and more efficient decision tree. For this reason, it has made it more interpretable. It is also emphasized in the article (Kapoor & Rani, 2015) that decision trees, regardless of advantages such as the ability to explain the selection method and low computational costs, usually find relatively good results in evaluation with other machine formulas.

In this paper, the definitions and concepts of cost, factors of cost of services, types of cost in the new system, and then calculating the cost in banks and financial and credit institutions will be briefly described.

* Two standard methods are applied to calculate the price of money in banks.

A. Calculation of direct cost of money

At this stage, the direct cost (interest) of money must be calculated. To this end, it is necessary to examine and calculate each resource's cost at the bank's disposal. The procedure is as follows:

1. The average balance of the bank's resources is calculated during the year. (Resources are all accounts to the left of the balance sheet except internal accounts.)
2. The legal deposit rate is multiplied by each average resource subject to legal deposit. The result is deducted from the average of each of the sources. Thus, the average resources are obtained after subtracting the legal deposit.
3. The liquidity ratio is multiplied by the resources included in it. Resources included in the liquidity ratio include current Qardh al-Hasan accounts, savings, short-term and long-term deposits, as well as cash deposits of guarantees. The liquidity ratio must be multiplied by the above headings and then deducted from each relevant resource.
4. The average of each resource is now available after deducting the legal deposit and deducting the liquidity ratio called net resources or free resources. The average free resources for each topic are now available.
5. Now the percentage of free resources in each heading is obtained from the sum of free resources.
6. In this stage, the cost percentage of each resource is multiplied by its corresponding heading (for example, the percentage cost of current accounts is zero and the percentage of savings accounts is 3%, as well as the percentage of short-term and long-term deposits and other resources). The percentage cost of each heading should be multiplied by the average balance of each heading, not the average free resources of each heading. The percentage cost of each heading should be multiplied by the average balance of each heading, not the average free resources of each heading.
7. The Central Bank of the Islamic Republic of Iran pays interest on the legal deposit that banks deposit based on the balance of deposited items. Therefore, in this stage, it is necessary to obtain the interest on the legal deposit belonging to each heading to adjust the cost of each heading of resources. To this end, the resources multiplied by the legal deposit rate should be considered, and the results obtained by multiplying the legal deposit rate in them by a factor of 2% should be obtained.
8. Now the adjusted cost of each heading should be obtained. To this end, the cost percentage of each heading is multiplied by the average of each heading. It is named A, and then the interest on the legal deposit of each heading is deducted from A. The product is divided by the free resources of each heading and multiplied by 100. In this way, the percentage of the adjusted cost of each heading is obtained.
9. Finally, the percentage of the adjusted cost of each heading is multiplied by the percentage of free resources of each heading obtained in paragraph "E." Hence, the percentage of the direct cost of money for each heading is obtained. The direct percentage of the total money for the bank is obtained by adding the percentage of all headings.

B. Indirect cost of money

After obtaining the direct cost of money, it is necessary to calculate the percentage of the indirect cost of money and add it to the percentage of direct cost to get the cost of money as a percentage. The method of calculating indirect costs is as follows:

A. The depreciation cost is obtained at the level of the whole bank.
B. The cost of bad debt receivables is obtained for the whole bank.
C. Administrative and personnel expenses are obtained for the whole bank.

D. The sum of the above costs is obtained, and it is divided by the average of free resources obtained in Part 1-D and multiplied by 100. With this operation, the percentage of the indirect cost of money is obtained.

Finally, this number is added to the result obtained in Part 1, and thus the cost of money is obtained.

II. RESEARCH LITERATURE

Database knowledge discovery has a short history, and it is currently referred to as data mining. When the phrase "knowledge discovery" was first used in the 1990s, data mining methods caught scholars' interest. Using artificial intelligence and statistical methods on large data sets, data mining seeks to uncover fresh, legitimate, and traceable information (Marbán, Segovia, Menasalvas, Fernández-Baizán, 2009). Data mining began to take shape in the late 1980s, achieved notable advancements in this area in the 1990s, and is predicted to continue expanding and changing in the twenty-first century (Smyth & David Hand). Mannila Heikki (2001). It is predicted that in the upcoming decades, it will undergo revolutionary development. Data mining is ranked among the top ten innovations by the Massachusetts Institute of Technology that will have a big impact on how the world develops (G&E, 2011). One of the newest and most popular subjects in recent years is data mining, which investigates the understanding of massive data and examines both permanent models and data linkages (Rahmani, Haji Zein Al-Abedini, 2015).

Therefore, data mining can be applied to discover and identify models between data in this large volume of data (Hayati, Sadeghi Mojarad, Jafari, 2010). Data mining can be considered the discovery of a model in mass data in which statistical techniques, mathematics, artificial intelligence, machine learning, and so on are used (Moradi and Ghasemi). Forecasting is the primary purpose of data mining. The decision tree, as one of the data mining techniques, has proven to be the most accurate prediction among other techniques, including artificial neural networks and regression models (Delen, Walker, & Kadam 2005). Predictive data mining is the most common type of data mining and has the highest commercial applications (Giudici, 2005). Data mining is a bridge between statistics, computer science, artificial intelligence, modeling, machine learning, and visual data representation, and it is possible to extract valid, previously unknown, understandable, and reliable information from large databases in a process framework. In other words, data mining is the science of extracting useful information from databases or datasets (India et al., 2001). The purpose of this research is to obtain factors affecting the cost of money through data mining. The following is a history of the banking industry in terms of the components that affect banks' income, costs, and efficiency:

The banking industry is one of the essential sectors of any economy because banks are considered the intermediaries of monetary resources along with the stock market, and insurance is the main pillar of financial markets. Banking is more critical in the Iranian economy because these banks are responsible for long-term financing due to the lack of necessary capital market development. Therefore, the efficiency of banks will also be of particular importance, given the role of banks in the economy. Profit margin is one of the most important criteria in evaluating the efficiency of banks (Shaeri & Eisazadeh, 2010). A study was conducted on the effect of the cost of money on market inflation based on bank resources from 2007 to 2011. The results showed a significant relationship between the percentage of long-term deposits, short-term deposits, and Gharz al-Hasna Bank deposits with the cost of money (Sadeghi and Azam Sharifi, 2016).

Before starting the calculation of the cost of money, it seems necessary to be familiar with some words:

A. Average resources

A part of bank resources or debts that are on the left side of the balance sheet. People's deposits and savings with banks are in the form of short-term, long-term, foreign currency participation bonds, etc. Another type of resource (debt) is loans and facilities received from other banks. Equity and funds such as the reserve bad debts and the reserve depreciation are also other resources. The average resources are obtained by calculating the average total of these debts from the beginning of the year.

B. Legal deposit

The legal deposit is obtained by multiplying the average resources by certain coefficients, which is deposited with the Central Bank. For example, the rates of Qardh al-Hasan deposits, corporate bonds, and term deposits are 2%, 0%, and 10%, respectively.

C. Liquidity ratio

It is a percentage of the average resources deducted from the legal deposit by the central bank to provide liquidity to bank branches and ATMs, reduce the central bank's resources (legal deposit), and increase the resources available to the bank.

D. Precautionary coefficient

Some of these resources are also available at any time to settle the exchange and respond to depositors. Although this amount of resources is set by the bank, it cannot be used other than what has been described.

The resources that are still available are those that are free for the bank to lend or invest in. These resources remain after the liquidity coefficient is added and the legal deposit and precautionary coefficient are subtracted. The net dividends paid are divided into free resources to determine the adjusted interest rate. This rate reflects the interest paid on all resources, so the portion of the adjusted interest rate that corresponds to each resource's share is determined by multiplying the adjusted interest rate by the percentage of free resources shared, or the "Balanced resource rate."

E. Variables

Operating expenses include all the expenses that banks incur in acquiring deposits, most of which are related to interest payments on customer deposits. Non-operating expenses include administrative and personnel expenses, depreciation costs for movable and immovable property, and the cost of doubtful receivables.

E.a. Dependent variable

The dependent variables of this research are 1- Average of 25 weeks 2- Average of legal deposit of 25 weeks 3- 2% of liquidity 4- Free resources 2- Net interest paid. Each of these

5 components is obtained from the total field digits of 1. Transaction Qardh al-Hasan 2. Transaction Qardh al-Hasan of non-bank credit institutions 3. Temporary creditors 4. Unclaimed balances 5. Banker's order 6. Unspent managed funds 7. Types of sold bank cheques (net) 8. Qardh al-Hasan deposit Rial savings 9. Special housing Qardh al-Hasan Savings 10. Youth Special Qardh al-Hasan Savings 11. Unused special Qardh al-hasan Savings 12. Short-term investment deposit 13. Special short-term investment deposit 14. One-year long-term investment deposit 15. Two-year long-term investment deposit 16. Three-year long-term investment deposit 17. Four-year long-term investment deposit 18. Five-year long-term investment deposit 19. Savings of government employees (employee share) 20. Savings of government employees (government share) 21. A cash deposit of guarantees in Rials (public sector) 22. A cash deposit of guarantees in Rials (private sector).

E.b. Independent variable

The independent variable in this research is overhead costs, which include 1. Administrative staff expense; 2. Depreciation of movable and immovable property expense; and 3. The miscellaneous expense of the branch. These expenses are all considered non-profit costs of the branch and are known in accounting as overhead costs of the branch.
Each of the variables of overhead costs, return on assets, and inflation, as well as the increase in market focus, which means an increase in the monopoly power of the banking system, reduces the efficiency of the banking industry (Isazadeh, Shaeri, 2012).

F. Sample and Population

This research has been conducted on 501 branches of National Bank and other branches of state-owned banks from 2019 to 2024.

III. RESEARCH METHODS

This applied research has employed a documentary study in terms of data collection. A list of the required information was extracted to access the information of bank branches from the database of state-owned banks. In this study, the branch cost data are divided into the first three levels with the symbols A, B, and C. At the beginning of the work, the minimum and maximum cost values of money are extracted from the data, and the above mathematical expression of class distance is obtained:

(1) Max (Cost)-Min (Cost) = 19.99 - 5.18 =14.81

14.81/3=4.93

The range of the first level A is between 5.18 and 10.11.
The range of the first level B is between 10.11 and 15.04.
The range of the first level C is between 15.04 and 19.99.
The research method is data mining. Data mining is a method and technology that has been introduced since 1994 (Trybula, 1997). Data mining is typically an iterative process in which multiple steps must be repeated several times (Qudsi, 2016). Data mining methods and technology turn large volumes of data into meaningful information to effectively support decision-making (Koh, 2011). The decision tree model, the neural network model, and the Bayesian rule-based model were the three primary data mining models whose performances were compared. The model with the lowest cost of ownership was found and assessed. To increase predictive power and decrease expenses and time in predicting the cost of money, this study has employed neural networks, decision trees, the RoughSet approach, and the Bayesian rule.
Artificial intelligence models such as decision trees and RoughSet are strong in predictability and model classification (Abrishami, 2006).

A. RoughSet

The RoughSet theory was developed in 1980 by Zdzislaw Pawlak to express and examine issues in which there is uncertainty and ambiguity and find discrepancies and connections in information. The most important features of this theory are:

1-The optimal algorithm for finding models in data 2- Finding relationships that are not discovered by statistical methods. 3- Possibility of using quantitative and qualitative data. 4- Finding the minimal set of data useful for classification (such as reducing the size and number of information). 5- Evaluating the importance of data 6- Generating decision-making rules from the information.

There was no need to normalize the data in the implementation of the RoughSet model of artificial intelligence. Initially, the data were classified as qualitative data because the RoughSet model is for qualitative data and gives the final result and analysis. In this model, the data are divided into two parts: 80% training and 20% testing. The RoughSet model can receive large volumes of inputs and outputs to show the rules derived from the models with high accuracy. This model cannot show forecast numbers and graphs and is a law-based model that only extracts rules. The RoughSet model is performed by Rosetta software. Three algorithms of genetics, Johnson, and Holt are used in the RoughSet model, which is explained in order of rules and results after the implementation of each algorithm as follows:

A.a. Genetic algorithm

Using a genetic algorithm, 1560 laws are generated, and finally, the most influential laws are presented.
Rule 1: If the amount of interest expense is 6.84 or lower and the amount of non-interest expense is 1.78 or lower, the branch is at the first level of the cost of money.
Rule 2: If the amount of interest expense is 14.04 or higher and the amount of non-interest expense is 1.78 or lower, the branch is at the third level of the cost of money.
Rule 3: If the amount of overhead costs is between 13716 and 53341 and the depreciation cost of movable and immovable property is 623 or lower, the branch is at the first level of the cost of money.
Rule 4: If the average amount of resources is 639684 to 924946 and the amount of non-interest expenses is 2.35 to 2.37, the branch will be at the first level of the cost of money.
Rule 5: If the branch's average legal deposit is between 67583 and 84486 and the non-interest expenses of the branch are between 2.35 to 2.37, this branch will be at the first level of the cost of money.
Rule 6: If the average free resource of the branch is 571093 and 160085 and the amount of non-interest expenses is between 2.35 to 2.37, this branch will be at the first level of the cost of money.

A.b. Johnson algorithm

Johnson algorithm generates 3 rules as follows:

Rule 1: If the cost of the branch is lower than 10.03, the branch is in the first level in terms of cost division. A total of 15 branches are in this level out of 501 branches studied in this study. Therefore, this level is the best in terms of branch ranking, and their position should be maintained.

Rule 2: If the branch cost is between 10.03 and 51.04, the branch is in the second level in the cost division. A total of 52 branches are in this level out of 501 branches studied in this study. This level is average in branch ranking and can grow and flourish and climb to the first level.

Rule 3: If the branch cost is between 15.04 and above, the branch is in the third level in the cost division. A total of 17 branches are in this level out of 501 branches studied in this study. Therefore, this level is the worst in branch ranking, and changes must be made in those branches to migrate to the first or second level.

A.c. Holt algorithm

Holt's algorithm generates 419 laws, and finally, the most effective laws are presented.

Rule 1: If the overhead costs of the branch are from 13716 to 14335, then that branch is at the first level of the cost of money.

Rule 2: If the average resource value is between 1188140 and 1244430, then that branch is at the first level of the cost of money.

Rule 3: If the average amount of free resources is between 1331880 and 1395790, then that branch is at the first level of the cost of money.

Rule 4: If the contingency reserve amount is from 1410 to 1901, then that branch is at the first or second level of the cost of money.

Rule 5: If the depreciation costs of movable and immovable property are 745 to 1140, then that branch is at the second level of the cost of money.

Rule 6: If the amount of interest expenses is between 8015 and 10.49, then that branch is in the second level of the cost of money.

According to the mentioned rules, the accuracy criterion using the RoughSet model in the test of the Holt algorithm is as much as 0.095238, which is shown in the following figure in Rosetta software.

B. Decision Tree J48

In this research, data sets are classified using J48 decision trees, and a higher detection percentage can be achieved by using classification and training of the detection system and evaluating performance based on correctly and incorrectly classified items using the J48 algorithm and minimizing false positives.

From the subset of trees, option J48 and validation-Cross 20 are selected, and the program is executed.

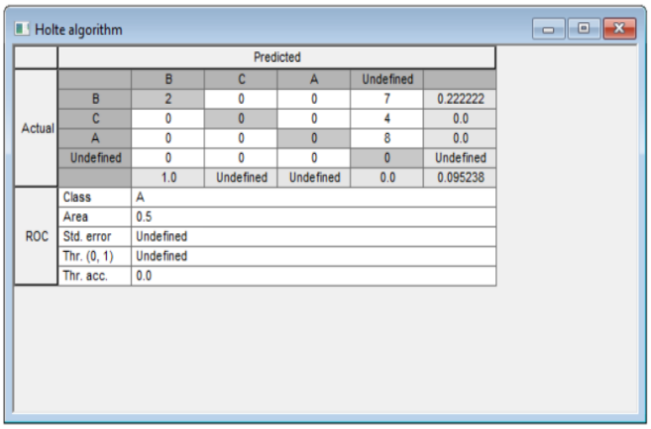

Fig. 1. Holt's algorithm - RoughSet model

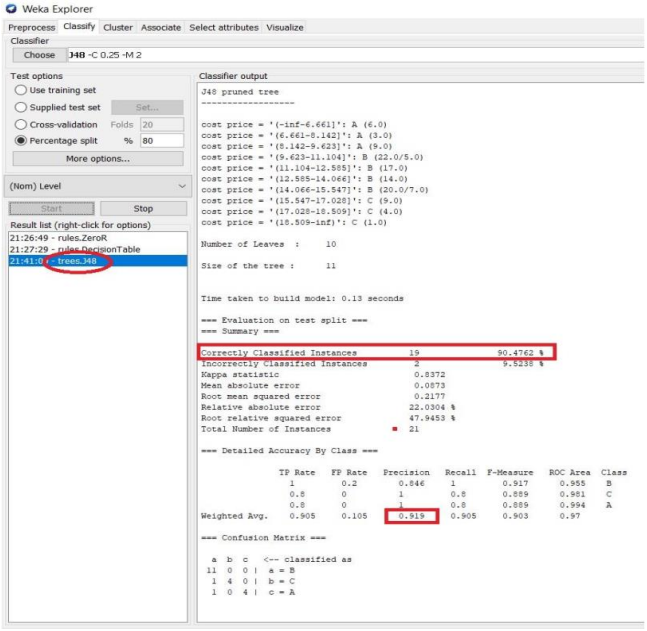

Fig. 2. Classification output in Weka decision tree J48

In total, 19 samples were categorized correctly, and two samples were categorized incorrectly.

After the preprocessed data was entered into Weka software, the following rules were generated using the decision tree algorithm:

Rule 1: If the cost of the branch is lower than 6.661, the branch is in the first level in terms of cost division.

Rule 2: If the branch cost is between 6,661 and 8,142, the branch is in the second level in cost segmentation.

Rule 3: If the branch cost is between 8.142 and 9.623, the branch is in the third level in the cost division.

C. Naive Bayes

The Bayesian is a supervised learning method. It is characterized by elegance, simplicity, and robustness. For this reason, it has become widely used for classification purposes (Hamid, Ahmed, 2016).

The results of implementing the Naive Bayes algorithm are shown in the following fig:

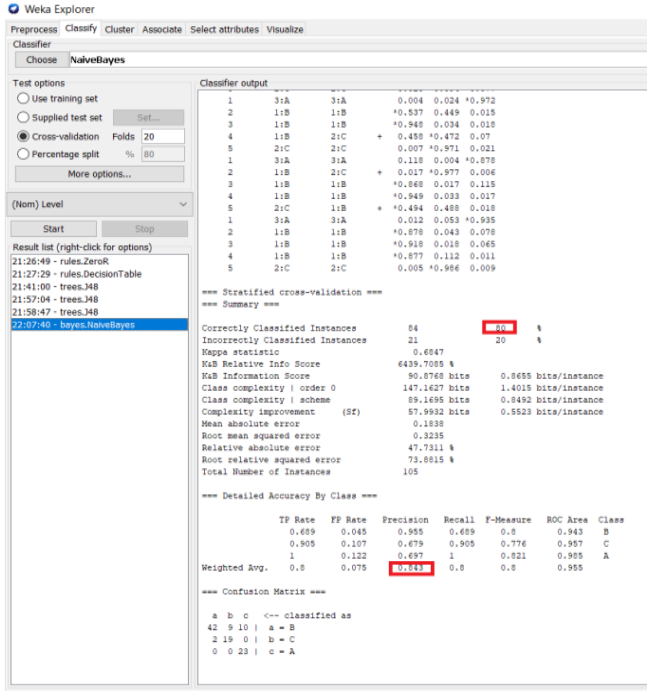

Fig. 3. Results from the implementation of the Naive Bayes algorithm

As it is clear from the classification results, the accuracy of the algorithm is 0.843. A total of 80 samples are classified correctly, and 20 samples are classified incorrectly, which has a higher error rate than the J48 algorithm.

D. Neural networks

The above results were obtained after implementing neural networks and a multi-layer perceptron algorithm in Weka software:

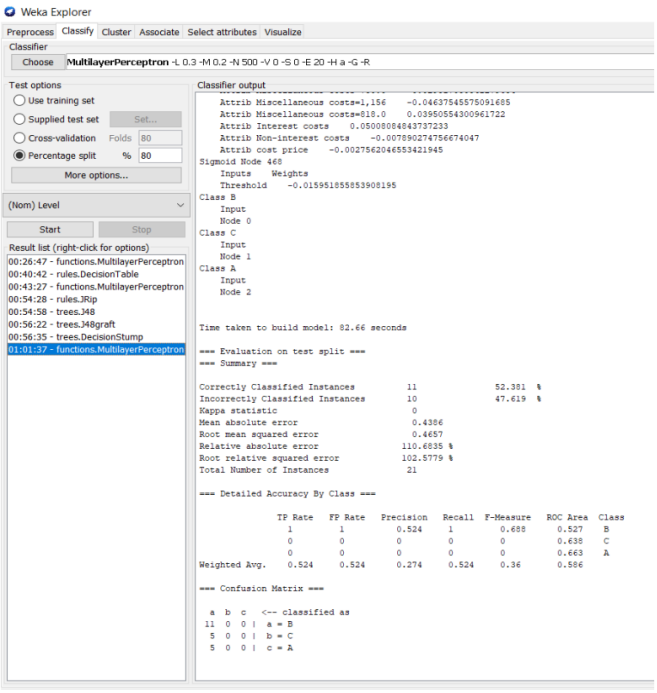

Figure 4. Implementation of neural networks and Multi-layer perceptron algorithm

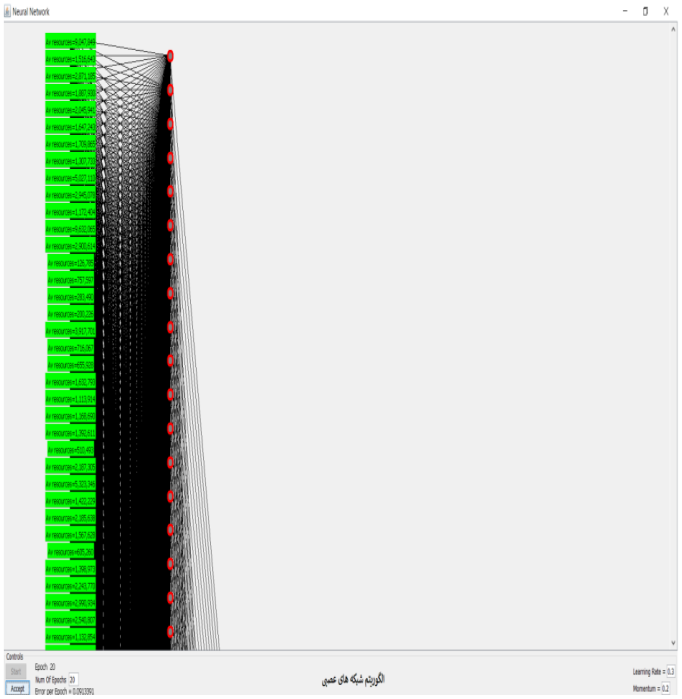

Fig. 5. Neural networks algorithm

The accuracy of this algorithm was as much as 0.274 after executing the neural network model.

IV. ANALYSIS OF FINDINGS (DISCUSSION)

The selected model is implemented in Weka software with 13 attributes and 501 samples. Then, the data are classified into three categories according to level A, level B, and level C, the result of which is specified in the following diagram:

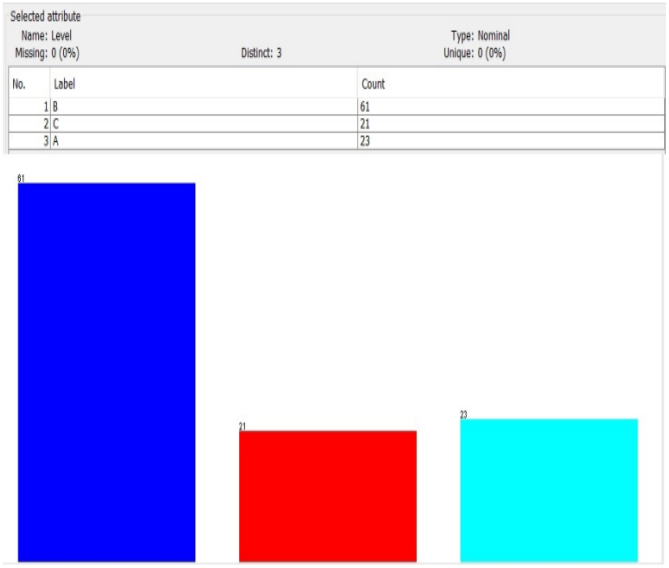

Fig. 6. Number of level A, B, C data

In addition, data collection and understanding of data obtained from the database of National Bank branches and other branches of state-owned banks in the western region of Tehran province have been used as raw data for this research.

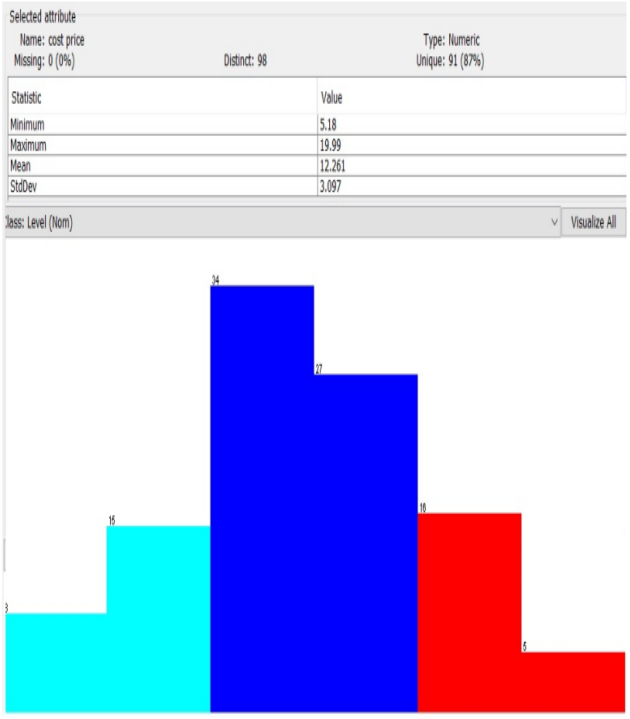

Fig. 7. Costs

These data are applied as input features for processing and executing the J48 decision tree algorithm in model prediction (Rajesh, Reena, & Maiti, 2018).

In this study, raw data is saved as an Excel file. The "Cost of Money" database file is incorporated into a Microsoft Excel spreadsheet and converted to CSV format so that it may be analyzed by Weka and Rosetta data mining tools. Following the completion of data integration, data cleaning was performed. Disproportionate data, data with no value (no file), and excessive data (data comprising several files with the same values) have been removed since their existence might degrade the quality and accuracy of data mining findings. After combining and cleaning the data, a 12-column table comprising the branches' financial data is created. The specifics of these traits are provided in Table I.

If the bank seeks to reduce the cost of money, it must combine the bank's resources to minimize operating costs and invest in the most profitable deposits. Therefore, bank managers are advised to focus on attracting long-term deposits and using these deposits to pay for more profitable facilities, as well as reduce non-operating costs (Rashidian, Mohammadipour, 2017).

The cost of money indicates the proper management of the bank's resources and expenses, and the arrangement and combination of bank resources can minimize the cost of money. All hypotheses were confirmed based on the data analysis and their reliability, and it was concluded that there is a significant relationship between the percentage of long-term deposits, short-term deposits, and Qardh al-Hasan bank deposits with the cost price of money. Therefore, if the bank seeks to reduce the cost of money, it must combine its resources to minimize future costs and invest in deposits where it is most profitable. Bank managers are advised to acquire low-cost deposits and profitable facilities and always consider optimal consumption (Sadeghi et al., 2016).

Iranian state-owned banks can reduce the cost of money by aiming to reform the composition of deposits. One of the essential factors in increasing the price of banking services and payment facilities by banks is the cost of money in the banking system, although the increase and decrease of this price depend on several factors.

According to banking experts and activists, one reason for the high cost of banking facilities is the rising cost of money in banks, and customers, and economic activists constantly criticize the banking system.

According to the author, reducing the cost of money in the banking system is a solution that can be considered to make it possible to acquire deposits in banks and to reduce the interest on facilities.

The cost price of money indicates the proper management of resources and costs and expenses of a bank, and the study of ways to reduce the cost of money in banks can be an indicator of the efficiency of managers and the bank's performance.

The largest share of the cost of money in the banking system is the interest rate on deposits and issues in the Iranian banking system, which is determined by the central bank based on the inflation rate. Considering the fierce competition between banks to acquire financial resources, the acquisition of low-cost resources significantly reduces the cost of money and subsequently provides low-cost facilities to customers.

V. Discussion and conclusion

The results showed that the J48 decision tree model has higher accuracy (0.919) compared to the three models based on the Naive Bayes rule (0.843), the neural network model (0.274), and the RoughSet model (0.095). However, two models of the J48 decision tree and the Naive Bayes model offer better predictions on the factors that affect the cost of money. Therefore, the appropriateness of the data, appropriate preprocessing, and the appropriate data mining strategy lead to better results concerning banking data.

TABLE I
The Result from the Algorithm

Technique	J48	Naive Bayes	neural network	RoughSet
Correctly classified instance percent	0.919	0.843	0.274	0.095

The J48 algorithm is best because it has high accuracy and low mean absolute error, as shown in the result. (Hamid, Ahmed, 2016). Since the J48 decision tree has higher optimization and accuracy than all data mining classification algorithms used in its research, the knowledge extracted from this tree is the most reliable knowledge obtained from the study data that can be the basis for extracting data mining rules.

The most important indicator in this algorithm is the cost; indicators of average resources, average legal deposit, contingency reserve, average free resources, net interest paid, administrative staff costs, depreciation costs of movable and immovable property, and other expenses are less critical, respectively.

In addition, the J48 algorithm successfully creates a classification model in the form of an illustration (decision tree) that can be easily studied and interpreted to predict the cost of money in banks.

As a result, the J48 decision tree is predicted to be a useful and informative method for conducting data mining.

Illustrations through the decision tree, which make reading

and interpreting easy, also help staff in the foundation to draw information (Patil, Toshniwal, & Joshi 2009). It is essential to make sure that the calculated sample ratio is balanced before performing the data mining. This ratio does not need to be the same, but it is suitable for the computing system to learn the classification for each created category.

This research was carried out on the branches of Bank Melli in 2019-2024. As a result, this bank became one of the most successful banks in reducing the cost of money in recent years with proper management of the effective targeting path to improve the composition of deposits by modifying deposit acquisition schemes, issuing general deposit certificates by Central Bank directives, reviewing the location of ATMs and POSs, and implementing a marketing plan. In other words, this bank has taken an effective step to reduce the cost of withdrawing money in terms of the growth rate of Rial resources among state-owned banks so that the cost of money in this bank has decreased compared to previous years.

One of the keys to success in reducing the cost of money is to design a variety of low-cost services and products. Providing new services and products based on the customers' needs causes the durability and survival of banks in a competitive environment and increases their credibility. In the meantime, the design of low-cost products has a special place in reducing the cost of money.

Bank Melli Iran also started designing new projects to meet the diverse needs of customers and target markets and design low-cost products in the current years, which was a practical step in providing accessible and cheap banking facilities to customers and reducing the cost of money.

The development of e-banking is also one of the indirect and crucial factors that affect the reduction of the cost of money considered in the operational management of Bank Melli. This topic can be placed in another subset, such as increasing productivity or reforming the organizational structure. In this type of banking, the costs of branches, employees, branch equipment, updating this equipment, and all costs related to branch banking are reduced, which leads to a reduction in interest rates on deposits and interest on facilities by reducing the cost price.

VI. REFERENCES

[1] Cleary, R. J. (2006). Applied data mining. Statistical methods for business and industry.

[2] David Hand, Heikki Mannila, & Smyth, P. (2001). Principles of Data Mining. The MIT Press.

[3] Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. Artificial intelligence in medicine, 34(2), -113 .127

[4] Eisazadeh, S., & Shaeri, Z. (2010). The effect of institutional factors on efficiency and cost in the Iranian banking system. The20 is the conference on banking and money research.

[5] G, R., & E, J. (2011). Data Mining Your Data for Healthcare Quality Improvement [Online]. Retrieved from http://www2.sas.com/proceedings/sugi22/EMERGING/PAPER139.PDf/

[6] Giudici, P. (2005). Applied data mining: statistical methods for business and industry. John Wiley & Sons.

[7] Koh, H. C., &Tan, G. (2011). Data mining applications in healthcare. Journal of Healthcare Information Management, 19(2), .65

[8] Marbán, O., Segovia, J., Menasalvas, E., & Fernández-Baizán, C. (2009). Toward data mining engineering: a software engineering approach. Information Systems, 34(1), 107-87. Retrieved from https://www.sciencedirect.com/science/article/pii/S0306437908000355

[9] Patil, B. M., Toshniwal, D., & Joshi, R. C. (2009). Predicting burn patient survivability using decision tree in Weka environment. In 2009IEEE International Advance Computing Conference (pp. 1356-1353). IEEE.

[10] Qudsi, D. H. (2016). Predictive Analytics Data Mining in Imbalanced Medical Dataset. Journal Computer Terapan .204-195 2 ,2,

[11] Rajesh, R., Reena, M., & Maiti, J. (2018). Decision Tree for Manual Material Handling Tasks

[12] Using WEKA. Ergonomic Design of Products and Work Systems - 21st Century Perspectives of Asia, 24-13

[13] Trybula, W. J. (1997). Data mining and knowledge discovery. Annual Review of Information Science and Technology (ARIST), 32, .229-197

[14] Yousofi, M. H., Esmaeili, M., & Sharifian, M. S. (2016). A study on image mining; its importance and challenges. American Journal of Software Engineering and Applications, 5(1-3), .9-5

[15] Abrishami, H. (2006). Fundamentals of econometrics. (University of Tehran Press)

[16] Hassanvand, R. (2016). The cost of money. Paper presented at the 6th International Conference on Accounting and Management and the 3rd International Conference on Entrepreneurship and Open Innovation. https://civilica.com/doc/577568.

[17] Hayati, Z., Sadeghi Mojarad, M., & Jafari, N. (2010). Discovering the movement of electronic information users using the dependency law algorithm in data mining: A case study of the UTS University Library website, Australia. Library and Information Science, (No. 1 (consecutive 49)), 1-4. Retrieved from http://lis.aqr-ed.pdf0818e2e0da09504bbd44af934b5397_43576libjournal.ir/article

[18] Rahmani, M., & Haji Zin Al-Abedini, M. (2015). Applications of data mining in information science and epistemology.

[19] Rashidian, S., & Mohammadi-Pour, and. V. R. A. (2017). Investigating the optimal combination of resources in Ilam Post Bank and determining its effects on the cost of money. The first conference on accounting, management, and economics with a dynamic approach to the national economy, Malayer.

[20] Sadeghi, K., & Azam Sharifi, V. (2016). Investigating the Impact of Money Cost on Market Inflation Based on Banking Resources. Paper presented at the 5th International Conference on Accounting and Management and the 2nd International Conference on Entrepreneurship and Open Innovation. https://civilica.com/doc/501479.

[21] Attaran, J., Diwandari, A., & Adinov, H. (2012). Identifying the effective factors on market consolidation (achieving sustainable competitive advantage) Banking services in Bank Mellat based on a resource-based perspective. Retrieved from Business Management (Management Knowledge), 4 (2 (12)). https://www.sid.ir/fa/journal/ViewPaper.aspx?id=205955

[22] Issazadeh, S., & Shaeri, Z. (2012). Investigating the effect of bank profit margins on the efficiency of the banking industry (Case study: Middle East and North Africa). Twenty-second Annual Conference on Monetary and Foreign Exchange Policies, Tehran.

[23] Fallah Shams, M., & Mahdavi Rad, H. (2012). Designing a credit model and forecasting the credit risk of leasing facility customers: Case Study: Iran Khodro Leasing Company). Retrieved from .234-213, (44) 12 Econo.

[24] Rokach, L., & Maimon, O. (2005). Decision trees. Data mining and knowledge discovery handbook, 165–192. Retrieved from https://doi.org/10.1007/0-387- 25465-X_9

[25] Quinlan, J.R., (1987). Simplifying decision trees, International Journal of Man-Machine Studies, 27, 221-234.

[26] Shahvaroughi Farahani, Milad, (2021). Prediction of Interest Rate Using Artificial Neural Network and Novel Meta-Heuristic Algorithms, Vol. 5 No. 1, https://civilica.com/doc/1328211

[27] Pei, D., Yang, T., & Zhang, C. (2020). Estimation of diabetes in a high-risk adult Chinese population using J48 decision tree model. Diabetes, Metabolic Syndrome and Obesity, 4621-4630.

[28] Kapoor, P., Rani, R., & JMIT, R. (2015). Efficient decision tree algorithm using J48 and reduced error pruning. Int. J. Eng. Res. Gen. Sci, 3(3), 1613-1621.

[29] Aljawarneh, S., Yassein, M. B., & Aljundi, M. (2019). An enhanced J48 classification algorithm for the anomaly intrusion detection systems. Cluster Computing, 22, 10549-10565.

[30] Cruz, A. P. D., & Tumibay, G. M. (2019). Predicting tuberculosis treatment relapse: a decision tree analysis of J48 for data mining. Journal of Computer and Communications, 7(7), 243-251.

[31] Ahishakiye, E., Taremwa, D., Omulo, E. O., & Niyonzima, I. (2017). Crime prediction using decision tree (J48) classification algorithm. International Journal of Computer and Information Technology, 6(3), 188-195.

[32] Rahi, M., Ebrahimnejad, A., & Motameni, H. (2024). Evaluation of predicted fault tolerance based on C5. 0 decision tree algorithm in irrigation system of paddy fields. International Journal of Intelligent Computing and Cybernetics, 17(2), 253-305.

[33] Ayinla, I. B., & Akinola, S. O. (2020). An improved collaborative pruning using ant colony optimization and pessimistic technique of C5. 0 decision tree algorithm. International Journal of Computer Science and Information Security (IJCSIS), 18(12), 111-123.
[34] Sum, R. M., Ismail, W., Abdullah, Z. H., FATHIHIN, N., SHAH, M. N., & HENDRADI, R. (2022). A New Efficient Credit Scoring Model for Personal Loan Using Data Mining Technique Toward for Sustainability Management. Journal of Sustainability Science and Management, 17(5), 60-76.
[35] Hamid, A. J., & Ahmed, T. M. (2016). Developing prediction model of loan risk in banks using data mining. Machine Learning and Applications: An International Journal, 3(1), 1-9.
[36] Zhu, F., Tang, M., Xie, L., & Zhu, H. (2018). A Classification Algorithm of CART Decision Tree based on MapReduce Attribute Weights. International Journal of Performability Engineering, 14(1), 17.
[37] Moslehi, F., Haeri, A., & Gholamian, M. R. (2019). Investigation of effective factors in expanding electronic payment in Iran using data mining techniques. Journal of Industrial and Systems Engineering, 12(2), 61-94.
[38] Dhakar, A., Singh, B., & Gupta, P. (2023). Fault diagnosis of air compressor set-up using decision tree-based J48 classification algorithm. Journal of Engineering Research.