Identifying the Financial, Economic and Structural Components Effective in Preventing Financial Frauds

Abdul Amir Majdam
Department of Accounting, Kish International Branch, Islamic Azad University, Kish Island, Iran
Abdolamiir.mojadam@42gmail.com

Shahram Chaharmahali
Assistant Professor, Accounting, Faculty of Economics and Accounting, Islamic Azad University, South Tehran Branch, Tehran, Iran (Responsible Author)
sh.chaharmahali@gmail.com

Mohammad Kohandel
Assistant Professor, Accounting, Faculty of Economics and Accounting, Islamic Azad University, South Tehran Branch, Tehran, Iran
kohanmo@yahoo.com

Nowruz Nouraleh Zadeh:
Assistant Professor, Accounting, Faculty of Economics and Accounting, Islamic Azad University, South Tehran Branch, Tehran, Iran
nour547@yahoo.com

Submit: 14/05/2023 **Accept:** 20/08/2023

ABSTRACT

The aim of the present study is to identify the financial, economic and structural components effective in preventing financial frauds. Heeding the issue of fraudulent financial reporting in the country, considering the increase in the number of companies admitted to the Tehran Stock Exchange, the membership of the Stock Exchange and Securities Organization among the members of the International Organization of Securities Commissions, the requirement to improve the quality of financial information, special attention to attracting foreign investors are deemed necessary in the post-sanction conditions and the continuation of the privatization process in the country. The statistical population of this study is all the companies accepted in the Tehran Stock Exchange from 2010 to the end of 2011, and their number is 570 companies. The hypothesis of the research is to evaluate the effectiveness of discovering the possibility of fraud in financial statements using the Morchegan algorithm in comparison with the regression method, neural network, K average and genetic algorithm. The results indicate that the performance of the Morchegan algorithm is based on entropy in the correct classification. Companies are similar to the performance of the genetic algorithm. The error in the classification of companies in the entropy-based Morchegan algorithm is significantly more than the error in the classification of companies using logistic regression methods and the neural network algorithm. The entropy-based algorithm in identifying companies suspected of fraud. It works significantly less successfully than the K-means method. The distance-based Morchegan algorithm is significantly more unsuccessful than the logistic regression method in correctly identifying companies as suspected fraud companies.

Keywords: Financial, Economic Components, Financial Frauds.

1. Introduction

Heeding the issue of fraudulent financial reporting in the country, considering the increase in the number of companies admitted to the Tehran Stock Exchange, the membership of the Stock Exchange and Securities Organization among the members of the International Organization of Securities Commissions, the requirement to improve the quality of financial information, special attention to Attracting foreign investors are deemed necessary in the post-sanction conditions and the continuation of the privatization process in the country. Despite the aforementioned necessity, the lack of a legal institution to discover and publicly announce cases of fraud in financial reporting and the lack of consensus among experts and elites regarding the definition of examples of the risk of fraud in financial statements, in addition to creating opportunities for fraudulent companies to abuse more, is a fundamental limitation in research in the field. has created fraud. This limitation can be seen in the previous research conducted in the country, Therefore, the previous researches have used examples of identifying fraudulent financial statements in the form of tastes and based on the judgment of the researcher. Rezaei and Riley (2019) believe that the extent of fraudulent financial statements has created concerns at the public level and has undermined public trust in the financial reporting process and audit performance. This phenomenon threatens the quality, integrity and reliability of the financial reporting process, and brings economic losses to investors and creditors. Joseph Wells (2015), the founder and president of the Association of Official Fraud Examiners, also states in the introduction to the second edition of the book "Fraudulent Financial Statements" that fraudulent financial statements have become an increasingly serious problem for governments, investors and businesses in the last century, which It gradually undermines public confidence in the capital market, corporate leaders and the auditing profession. Financial statements are a basic document of a company that shows its financial status (Ravisonkaro et al., 2018). Financial statements are the main basis for decision-making in a large part of investors, creditors and other persons who need accounting information, as well as expressing It is a special aspect of business performance, financial status and social responsibility of our company. Fraudulent financial statements are intentional and illegal acts that lead to misleading financial statements or misleading financial disclosures. Shareholders are unknowingly affected by misleading reports (Elliott and Wellingham, 2017). However, in recent years, cases of fraudulent financial statements have become increasingly serious (Yeh et al., 2015). Considering these incidents, it is important to be able to detect fraudulent behavior before it occurs. Cavic is the key tool for dealing with complex data analysis and classification. This tool summarizes the valuable information hidden in large amounts of data for analysis into a structured model to provide a reference for decision making. Data mining has many different functions, such as classification, association, clustering and prediction (Ceifert, 2019). The classification function is often used. Classification results can be used as a basis for decision making and for forecasting purposes. Fuzzy financial statements can be viewed as a typical classification problem (Kirkes et al., 2018). The classification problem involves performing calculations using variable features from some known classification data in order to derive rules. Classification is relevant. After that, the unknown classification data is fed into the rules to achieve the final classification results. Data mining is an iterative process where progress is defined by discovery either through automated or manual methods. It is most useful in an exploratory analysis scenario where there is no preconceived notion of what is not a "preferred" result. Many law enforcement and special investigative units whose mission is to identify fraudulent activities have successfully used data mining. However, as opposed to other researched fields such as predicting bankruptcy or financial dissatisfaction, research on the use of DM methods has been used for the purpose of detecting management fraud. The operational definition of fraud risk in financial reporting is a research gap. The current research seeks to present a comprehensive model of fraudulent financial reporting in the country and develop its detection methods. According to the mentioned materials, the main problem of the research can be stated as follows. Zhao and Kapur (2023) investigated the effectiveness of data mining methods such as logistic regression, decision tree, neural network and Bayesian networks. In addition, they adapted data mining methods to new fraud schemes. The results of their research confirm the successful performance of data mining methods in discovering fraudulent financial reporting. Advanced detection methods with new fraud designs were more efficient and effective than other models.K Roux and colleagues (2023) investigated the performance and efficiency of decision tree methods, artificial neural network and Bayesian speculative network as data mining techniques in detecting fraud in financial statements. The variables of this research include ratios obtained from financial statements. The results of this research have confirmed the efficiency of data mining methods in discovering the fraud of financial statements, so that the Bayesian speculative network method has the most efficiency in this field.

In view of the significant losses and severe damages that have been and will be inflicted on the economy of the countries of the world due to "fraud" and also considering the increasing impact of the audit reports of companies on the stability of financial markets, in recent years, extensive efforts have been made to improve audit operations. Tools and techniques have been designed so that auditors can perform their audit operations with the help of computers. Salem, (2019) that one of these techniques and tools is "expert systems". Therefore, since the last decade, the desire to use the systems that they call "expert systems" has increased widely. The main difference between these software-data systems that process information and other software is that expert systems process knowledge, while other software may only process data and information. Mahdavi and Gharmani, (2022).One of the most important tools that can be used as a support for auditors to assess fraud is artificial neural networks (ANNs). These networks are one of the broad and widely used branches of artificial intelligence that regulate and explore methods and algorithms based on which computers and systems gain the ability to learn from past experiences. Machine learning methods are divided into three categories: "supervised learning", "unsupervised learning" and "reinforcement learning", each of which is subject to its own model and method. Ismailpour et al., (2021).

"What are the examples of the risk of fraud in financial statements in the country and which method is suitable to discover it?"

Foreign Research Background

Alia and et al (2023) 1 The effects of variables such as financial goals, financial stability, pressure externalities, inefficient supervision, dual duties of the CEO and political connection in financial reporting investigated fraudulently. The results of their research using multiple regression analysis

was analyzed and showed that the financial stability and the dual duties of the CEO can be used to detect fraud in financial reporting. However, financial objectives, external pressure, supervision Inefficient, political connection and continuity of the company's activity cannot be used to detect fraud in reporting Financially used.

Rasouli Pareshkoh et al. (2023) investigated the relationship between audit quality, stability and financial goal with the possibility of fraud in financial statements. The results of the research showed that there is a negative and significant relationship between audit quality and the possibility of fraud. In other words, with the increase in audit quality in the company, the possibility of fraud in financial statements is reduced. Also, the research findings showed that stability and financial goal encourage fraud in the company. Based on the results of the research, the hypothesis of personal interests is accepted in the Tehran Stock Exchange.

Yusef and Lai (2023), in a study titled "Application of Fraud Models in Malaysian Stock Exchange Companies", investigated the possibility of financial statement fraud in Malaysian stock companies using the Fraud Triangle Model, the Fraud Diamond Model, and the Fraud Pentagon Model. gave The results of this research introduce measurable criteria and new fraud risk factors such as greed and ignorance.

In a research, Tajiknia et al. (2023) investigated the cost of not choosing profit management of follower companies in response to the fraudulent financial reporting of the leading company. The results of the test show the existence of a relationship between not performing profit management and a reduction in the benefits of senior management. In the periods when companies are faced with the motivation of this action and refuse to perform profit management, they experience a decrease in the benefits of their senior managers, and also the cost of capital in companies without profit management is facing a decrease, but based on the results, this decrease cannot be attributed to not choosing the option. Earnings management generalized and no significant relationship was observed between reducing the credit rating and not choosing the earnings management option.

Khoshbakht et al. (2023) investigated religiosity, accountants' professional ethics and financial reporting frauds in a research. The findings showed that accountants' religiosity has a negative and significant effect on financial reporting frauds and has a positive and significant effect on accountants' professional ethics. Also, religiosity causes accountants' to have a higher moral level, and this higher moral level is an effective factor in reducing fraud in financial reporting by accountants' , which shows that the professional ethics of accountants' explains the relationship between their religiosity and reporting frauds.

Melki Kakler et al. (2022) in a research investigated the presentation of the developed pentagonal model of fraud in financial reporting with an emphasis on the structure of internal controls. The findings of the research show that the variables of pressure, opportunity, justification, capability and structure of internal controls have no significant effect on fraudulent financial reporting.

Shah Moradi and Tabatabaei Nesab (2022) investigated the impact of audit quality on the relationship between economic uncertainty and profit management caused by accruals in companies listed on the Tehran Stock Exchange. The analysis of the results indicates that in the economic uncertainty, the use of large auditing companies and the increase in tenure lead to a decrease in profit management. In addition, the provision of non-standard audit reports has strengthened the direct impact of economic uncertainty on corporate profit management.

Muhammad and Schachler (2022) in a case study in Malaysia, through interviews with managers, auditors and legislators, have examined the potential methods of combating, detecting and interacting with fraud in financial statements. The results of their research emphasize the importance of developing internal (control) systems to prevent and detect fraudulent financial reporting.

Lin and colleagues (2022) investigated the efficiency of logistic regression, decision tree, and artificial neural network models. The results of their research confirm the efficiency of neural network and decision tree models in comparison with logistic regression. A significant point in their research is the correct classification rate of more than ninety percent of fraudulent companies in the use of neural networks.

Kanapina and Grondin (2021) proposed a model to detect fraud in financial statements by means of financial ratios. They studied 18 fraudulent financial statements and 480 non-fraudulent financial statements (whose audit report was acceptable) using the logistic regression model. The results of their research have confirmed the effectiveness of the presented model in predicting fraudulent financial reporting.

Orgin and Odom (2021) analyzed the cost benefit of the punishments considered in the Sarbanes-Oxley law, and investigated its deterrent power in preventing fraudulent financial reporting. The results of their research showed that the increase in prison time and its economic consequences in the life of criminals (such as the loss of job opportunities) have increased the deterrent power of the law.

In a research, Rezaei (2020) has emphasized the role of accounting in this field and introduced strategies to prevent fraud by examining cases of fraud in financial statements. In addition, he has examined the role of Sarbanes-Oxley law in improving corporate governance and improving the quality of financial reporting. He identifies five effective factors in the occurrence, detection and avoidance of fraud in financial reporting. These factors are respectively: the fraudulent person, fraud plan, fraud motivation, monitoring process and the final results of fraud in financial statements.

Santoso (2019) predicted the fraud of financial statements using the diamond fraud model in Indonesian manufacturing companies. The samples used in this study included 86 manufacturing companies that were registered in Bahadur Indonesia Stock Exchange in 2014-2012. In this study, secondary data in the form of financial reports of companies were used. The results showed that variables such as pressure, opportunity, rationality and ability simultaneously determine financial statement fraud. Also, based on the partial test, external variable pressures, financial goal and financial capability can also be used to predict financial statement fraud.

Ahmadi et al. (2019) in a research investigated the presentation of a model for predicting fraudulent financial statements and comparing financial statements and ratios with Benford's law. The results show that according to the accuracy rate of 64.6%, this model has an effective role in discovering the fraud of financial statements. Also, the results of T-test and Lone's test in 35 independent variables examined showed that there is a significant difference in 20 variables in the cheating and non-cheating groups. In addition, compatibility and deviation from Benford's law were investigated in four different situations and the following results were obtained: Paying attention to the figures of the profit and loss statement and the balance sheet in non-fraudulent companies showed that Benford's distribution correctly identified non-fraudulent companies, but incorrectly identified fraudulent companies as non-fraudulent. Paying attention to the financial ratio of total assets to sales and the payment period of accounts payable in fraudulent companies showed that the Benford distribution has assessed these companies as fraudulent and classified them correctly, but has incorrectly evaluated non-fraudulent companies as fraudulent.

Osenberg et al. (2018) proposed a computational fraud detection model to detect fraud in financial statements. The mentioned model uses the approach of quantifying the information of textual data regarding the detection of fraud. They checked the performance of their proposed model with the "Management Discussion and Analysis" test. Their proposed model had a successful performance in detecting fraudulent cases.

Rahimian and Haji Heydari (2018) have investigated fraud using the modified Benish model and identified financial ratios sensitive to fraud. In terms of the method of data collection, the descriptive research is of the regression analysis type. First, out of 150 companies using Benish's modified model, fraudulent companies were identified and separated from non-fraudulent companies, and 25 financial ratios were selected from among the ratios introduced by previous researches, and after the implementation of the Kolmogorov-Smirnov test, 10 The financial ratio of fraudulent and non-fraudulent companies has been determined. After running the regression model in three stages, the results show that the ratio of sales to total assets and the ratio of equity to total assets are two financial ratios sensitive to fraud.

Research Method

Statistical Community of Research

The purpose of conducting research is to recognize and predict a phenomenon in a statistical community. In order to gain knowledge about that phenomenon, samples from that population are selected and analyzes are performed on the selected sample and then the results obtained are generalized to the entire statistical population. The statistical population of this research is all the companies accepted in the Tehran Stock Exchange from 2010 to the end of 2019, and their number is 570 companies until the end of 2019.

Research Variables

In order to define the variable of financial fraud, first, the design and examples of fraud were extracted from the qualitative part of the research. After designing a comprehensive model of fraud in financial reporting in Iran, by referring to the report of the independent auditor and the legal auditor of sample companies, the content of the clauses of the auditor's report has been adapted to the examples of fraud in financial reporting. If the auditor refers to any of the fraud schemes in financial reporting, the said financial statement is suspected of fraud. The independent variable in this research is defined as 0 and 1. The number 1 is defined to indicate a suspected fraudulent financial statement and the number 0 is defined to indicate a non-fraudulent financial statement. The variables of the current research are defined as follows.

Current Ratio

This ratio is calculated by dividing current assets by current liabilities. The larger the company's current ratio, the less problems it faces in paying current debts (Narson, 2015):

$$\text{Current Ratio } \frac{\text{Current Assets}}{\text{Current Liabilities}} \qquad (1)$$

Periodicals Collection

The receivables collection period shows the company's management of accounts receivable (Higgins, 2014). If fraud in financial reporting is of the type of fake income recognition, it leads to claims that will never be collected (Cruncher, Riley and Wells, 2011). This ratio is calculated as follows:

$$\text{Period Receivables} = \frac{365 \times \text{Average Account Receivables}}{\text{Sales Credit}} \qquad (2)$$

Operating Cash to Sales Ratio

The inability of a business entity to generate cash flow from operations despite reporting profit and profit growth is one of the risk factors of fraud, which is mentioned by the auditing standard 240. In order to operationalize the definition of this factor, the ratio of operating cash to sales can be used. This ratio is calculated as follows:

$$\text{Sales to Cash Operations Ratio } = \frac{\text{Cash Operations}}{\text{Sales}} \quad (3)$$

Debt Payment Period

Cranchler et al. (2019) believe that the debt payment period is one of the signs of fraud detection risk. This ratio is equal to the average accounts payable on the company's credit purchase in 365 days, the calculation method of which is presented below.

$$\text{Period Payment Debt} = \frac{365\times\text{Average Account Payables}}{\text{Purchase Credit}} \quad (4)$$

Inventory Turnover Period

This ratio shows the time it takes for raw materials to be converted into goods and sold. An increase in the period of inventory turnover causes an increase in storage costs, the risk of product scarcity and a decrease in the market price (Cruncher, Riley and Wells, 2011). The way to calculate this ratio is presented below:

$$\text{Circulation Inventory } = \frac{365\times\text{Average Inventory}}{\text{Sold Goods at Full Price}} \quad (5)$$

Asset Turnover Ratio

The asset turnover ratio calculates the realized sales per one riyal of assets. This ratio shows the dependence of sales on assets. Low turnover of assets means a business dependent on assets, and high turnover of assets has the opposite meaning (Higgins, 4939). The ratio of turnover of assets is calculated as follows.

Average Total Assets / Net Sales = Asset Turnover Ratio (6)

Operating Profit Margin

The profit margin is a part of each riyal of sales that is added to the profit through the profit and loss statement. This ratio is especially important for operational managers. The net profit margin is calculated as follows:

Sales / Net Profit = Net Profit Margin (7)

When analyzing profitability, it is necessary to distinguish between variable costs and fixed costs. Variable costs change with changes in sales, while fixed costs remain unchanged. Companies that have a lot of fixed costs are more vulnerable than other companies to a decrease in sales because they are not able to reduce fixed costs with a decrease in sales. Gross profit margin provides the possibility to differentiate between fixed and variable costs as much as possible (Higgins, 2013).

Higgins (2013) believes that the profit margin shows the organization's pricing strategy and the organization's ability to control operating costs. Therefore, he defined the operating profit margin as follows:

Sales/Operating Profit = Operating Profit Margin (8)

Efficiency on Equity

The most common measure of financial performance among investors and senior managers is efficiency on equity, which is defined as follows (Higgins, 2013).

Net Profit/Equity = Efficiency on Equity (9)

Efficiency on Assets

Efficiency on assets is a basic indicator regarding the effectiveness of allocation and management of company resources (Higgins, 2014). This ratio is calculated as follows:

Net Profit/Total Assets = Efficiency on Assets (10)

The ratio of facilities to capital

This ratio represents the percentage of the company's capital that is provided through facilities. A high ratio of facilities to capital means high financial risk and low flexibility. This ratio is calculated according to the following relationship (Henri et al., 2013).

Equity + Total Facilities/Total Facilities=Ratio of Facilities to Capital (11)

Debt Ratio

With this ratio, the total funds obtained from debts are calculated (Peri Nervo, 2013). This ratio shows the percentage of funds secured by debt. This ratio can be calculated as described below.

Total Assets/Total Debt=Debt Ratio (12)

Financial Leverage

This ratio shows the amount of assets supporting each riyal of shareholders' equity. The amount of financial leverage indicates the amount of debt used in financing assets (Henri et al., 2018).

Average Assets/Average Equity=Leverage Ratio (13)

Interest Coverage Ratio

Using this ratio, the ability of the company to pay the loan interest is calculated. Sometimes reducing this ratio means increasing financial risk. If the amount of this ratio exceeds a certain level, the lending organizations demand more interest from the loan applicant and set tougher conditions (Henri et al., 2018). This ratio is calculated as described in the following relationship.

Earnings before Interest and Taxes/Finance Charge=Interest Coverage Ratio (14)

Article 141 of the Trade Law

Article 141 of the law amending a part of the commercial law states that "if at least half of the company's capital is lost due to losses, the board of directors is obliged to immediately call an extraordinary general meeting of shareholders. Therefore, the issue of liquidation or survival of the company can be voted on." If the aforementioned meeting does not vote to liquidate the company, it must reduce the company's capital to the amount of the existing capital in the same meeting and in accordance with the provisions of Article 6 of this law. In case the board of directors does not invite an extraordinary general meeting contrary to this article, or if the meeting that is invited cannot be held in accordance with the legal provisions, any beneficiary can request the dissolution of the company from the competent court." This article states the conditions of dissolution and bankruptcy of the business unit. To operationalize the conditions stated in the law, the ratio of accumulated profit and loss to the capital of the company can be used. The more this ratio tends to -0.5 and smaller, the more likely the company will be liquidated.

Accumulated Profit and Loss/Ordinary Share Capital=Article 141 of the Commercial Law

Inferential Analysis of Research

In this section, according to the confirmation of the theoretical adequacy of the research questions of each of the desired variables, in this section, the suitability of the model is applied to test the research hypothesis. Based on the analysis done in the previous section and also, the study of the paragraphs of the auditor's report can provide useful information about the possibility of fraud in the financial statements. In this way, according to the examples of fraud in financial reporting, by reading the paragraphs of the auditor and legal inspector's reports and if the auditor refers to any of the fraud plans in financial reporting, the said financial statement is suspected of fraud. The independent variable in this research is defined as 0 and 1. The number 1 is defined to show the financial statement suspected of fraud and the number 0 is defined to show the clean financial statement. Also, based on the analysis done in the previous section, the analysis of financial ratios is useful in detecting fraud. Therefore, the financial ratios of current ratio, receivables collection period ratio, operating cash to sales, debt payment period ratio, inventory turnover, asset turnover ratio, operating profit margin, return on equity, return on assets, ratio of facilities to capital, Debt ratio, leverage ratio, interest coverage ratio were extracted from financial statements. In addition, the interviewees believed that examining the limitations of Article 141 of the Trade Law can be useful in detecting fraud. Therefore, the limitation variable of Article 141 was defined as the division of accumulated profit and loss into capital.

Logistic Regression

The Importance of Independent Variables

The importance of independent variables helps to identify the extent to which the values predicted by the network are changed by changing the values of the independent variable. Table 1 shows the importance of each of the independent variables. Based on this table, return on equity, operating cash for sale, return on assets and debt collection period ratio are the most important, respectively. The importance of other variables is less than 40%.

Confusion Table

The results of the Confusion table (based on the number) obtained from 30 executions of the artificial neural network algorithm in the evaluation phase are presented in table 2. Based on the information in table 3, the performance criteria obtained from each execution of the artificial neural network in the evaluation stage were calculated. The performance of the artificial neural network algorithm can be checked based on the criteria of accuracy, error rate, sensitivity, specificity and accuracy. Table 3 presents the results of calculating the performance criteria of the neural network algorithm. The summary of the results related to the performance criteria of the artificial neural network algorithm in the evaluation phase is presented in table 2.

Table 1 the Importance of Independent Variables Resulting from Artificial Neural Network

Research variables	Importance	Normalized Significance
Current Ratio	٠.٠٤٧	٢٦.٧%
Ratio of Receivables Collection Period	٠.٠٧٣	٤١.٥%
Operating Cash to Sell	٠.١٦٧	٩٤.٤%
Debt Payment Period Ratio	٠.٠٢٨	١٥.6%
Inventory Turnover	٠.٠٦٢	٣٥%
Circulation of Assets	٠.٠٣٥	٢٠.١%
Operating Profit Margin	٠.٠٤٩	٢٧.٩%
Return on Equity	٠.١٧٧	١٠٠%
Return on Assets	٠.١٥١	٨٥.٥%
The Ratio of Facilities to Capital	٠.٠٤٤	٢٥.٢%
Debt Ratio	٠.٠٤٨	٢٧.١%
Financial Leverage	٠.٠٣٤	١٩%
Interest Coverage Ratio	٠.٠٤٦	٢٦%
Article 141	٠.٠٣٩	٢٢.٣%

Table 2 Results of Artificial Neural Network Algorithm Implementation

Performance	True Negative	False Positive	False Negative	True Positive	Performance	True Negative	False Positive	False Negative	True Positive
1	١٢٧	١٨	٩٧	٢٣	١٦	١٣٨	١٩	٩٣	٢٩
2	١٥٣□	٢٧	٦٨	٤٦	١٧	١٤٣	١٠	٩٣	١٢
3	١٤٧	١٤	٩٥	٢٠	١٨	١٤٩	١١	١٠٠	٢٠
4	١٥٦	١٠	١٠٦	٧	١٩	١٤٦	١	١٠١	٣
5	١٣٠	٢٠	٩١	٢٣	٢٠	١٤٠	٣٠	٨٨	٣٠
6	١٥٣	١٤	٩٣	٢٠	٢١	١٣٦	٢١	٩٣	٢٧
7	١١٣	٣٤	٦٩	٥٢	٢٢	١٢٥	٢٧	٧٥	٤٠
8	١٣٩	٢٣	٧٩	٢٠	٢١	١٣٦	٢١	٩٣	٢٧
9	١١٩	٣٧	٦٨	٤٣	٢٤	١٧٢	٥	١٠٨	٩
10	١٢٦	٢٠	٩١	٣٤	٢٥	١٣١	٢٠	٩٦	٣٠
11	١٢٣	٢٨	٧١	٣٢	٢٦	١٥٦	٦	٩٥	٦
12	١٤١	٨	١٠٣	١٣	٢٧	١٤١	٢٥	٧٦	٣١
13	١٢٤	٢٣	١٠٤	٢٦	٢٨	١٣٧	٧	١٠٠	١١
14	١٣٧	٣	١٠٠	٣	٢٩	١٢٨	٣٢	٧٠	٤٤
15	١٣٠	٢٥	٦٥	٣٣	٣٠	١٢٠	١٨	٧٥	□٣٢

Table 3 Performance Criteria Obtained from 30 Executions of Artificial Neural Network

Performance	Precision	Error Rate	Sensitivity	Feature	Correctness	Performance	Precision	Error Rate	Sensitivity	Feature	Correctness
1	٠,٥٨٢	٠,٤١٨	٠,٢٥٤	٠,٨٧٦	٠,٦٤٧	١٦	٠,٥٩٩	٠,٤٠١	٠,٢٣٨	٠,٨٧٩	٠,٦٠٤
2	٠,٦٧٧	٠,٣٢٣	٠,٤٠٤	٠,٨٥٠	٠,٦٣٠	١٧	٠,٦٠١	٠,٣٩٩	٠,١١٤	٠,٩٣٥	٠,٥٤٥
3	٠,٦٠٥	٠,٣٩٥	٠,١٧٤	٠,٩١٣	٠,٥٨٨	١٨	٠,٦٠٤	٠,٣٩٦	٠,١٦٧	٠,٩٣١	٠,٦٤٥
4	٠,٥٨٤	٠,٤١٦	٠,٠٦٢	٠,٩٤٠	٠,٤١٢	١٩	٠,٥٩٤	٠,٤٠٦	٠,٠٢٩	٠,٩٩٣	٠,٧٥٠
5	٠,٥٨٠	٠,٤٢٠	٠,٢٠٢	٠,٨٦٧	٠,٥٣٥	٢٠	٠,٥٩٠	٠,٤١٠	٠,٢٥٤	٠,٨٢٤	٠,٥٠٠
6	٠,٦١٨	٠,٣٨٢	٠,١٧٧	٠,٩١٦	٠,٥٨٨	٢١	٠,٥٨٨	٠,٤١٢	٠,٢٢٥	٠,٨٦٦	٠,٥٦٣
7	٠,٦١٦	٠,٣٨٤	٠,٤٣٠	٠,٧٦٩	٠,٦٠٥	٢٢	٠,٦١٨	٠,٣٨٢	٠,٣٤٨	٠,٨٢٢	٠,٥٩٧

Performance	Precision	Error Rate	Sensitivity	Feature	Correctness	Performance	Precision	Error Rate	Sensitivity	Feature	Correctness
8	٠,٦٠٩	٠,٣٩١	٠,٢٠٢	٠,٨٥٨	٠,٤٦٩	٢٣	٠,٦٠٤	٠,٣٩٦	٠,٢٢٧	٠,٩١٦	٠,٦٩٠
9	٠,٦٠٧	٠,٣٩٣	٠,٣٨٧	٠,٧٦٣	٠,٥٣٨	٢٤	٠,٦١٦	٠,٣٨٤	٠,٠٧٧	٠,٩٧٢	٠,٦٤٣
10	٠,٥٩٠	٠,٤١٠	٠,٢٧٢	٠,٨٦٣	٠,٦٣٠	٢٥	٠,٥٨١	٠,٤١٩	٠,٢٣٨	٠,٨٦٨	٠,٦٠٠
11	٠,٦١٠	٠,٣٩٠	٠,٣١١	٠,٨١٥	٠,٥٣٣	٢٧	٠,٦١٦	٠,٣٨٤	٠,٠٥٩	٠,٩٦٣	٠,٥٠٠
12	٠,٥٨١	٠,٤١٩	٠,١١٢	٠,٩٤٦	٠,٦١٩	٢٧	٠,٦٣٠	٠,٣٧٠	٠,٢٩٠	٠,٨٤٩	٠,٥٥٤
13	٠,٥٤٢	٠,٤٥٨	٠,٢٠٠	٠,٨٤٤	٠,٥٣١	٢٨	٠,٥٨٠	٠,٤٢٠	٠,٠٩٩	٠,٩٥١	٠,٦١١
14	٠,٥٧٦	٠,٤٢٤	٠,٠٢٩	٠,٩٧٩	٠,٥٠٠	٢٩	٠,٦٢٨	٠,٣٧٢	٠,٣٨٦	٠,٨٠٠	٠,٥٧٩
15	٠,٦٤٤	٠,٣٥٦	٠,٣٣٧	٠,٨٣٩	٠,٥٦٩	٣٠	٠,٦٢٠	٠,٣٨٠	٠,٢٩٩	٠,٨٧٠	٠,٦٤٠

Table 4 Summary of Artificial Neural Network Performance

	True Negative	False Positive	False Negative	True Positive	Accuracy	Error Rate	Sensitivity	Specificity	Accuracy
Mean	١٣٧	١٨,٣	٨٨,٧	٢٥,٣	٠,٦٠٣	٠,٣٩	٠,٢٢٠	٠,٨٨	٠,٥٨
Max.	١٧٢	٣٧	١٠٨	٥٢	٠,٦٧٧	٠,٤٥	٠,٤٣٠	٠,٩٩	٠,٧٥
Min.	١١٣	١	٦٥	٣	٠,٥٤٢	٠,٣٢	٠,٠٢٩	٠,٧٦	٠,٤١
SD	١٣,٢	٩,٦	١٣,٣	١٣,٢	٠,٢٥	٠,٠٢	٠,١١٤	٠,٠٦	٠,٠٦

K Mean Method

MATLAB software has been used to implement the K-means clustering method. The results of the Confusion table (based on the number) obtained from 30 times of K-means clustering method are presented in table 5.

Table 5 the Results of the Implementation of the K Average Method

Performance	Negative	Positive	Negative	Positive	Performance	Negative	Positive	Negative	Positive
1	263	264	175	200	16	263	264	175	200
2	264	263	200	175	17	263	264	175	200
3	264	263	200	175	18	264	263	200	175
4	264	263	200	175	19	263	264	175	200
5	263	264	175	200	20	263	264	175	200
6	264	263	200	175	21	264	263	200	175
7	264	263	200	175	22	263	264	175	200
8	263	264	175	200	23	263	264	175	200
9	264	263	200	175	24	263	264	175	200
10	264	263	200	175	25	264	263	200	175
11	264	263	200	175	26	٥١٥	12	358	7
12	263	264	175	200	27	263	264	175	200
13	264	263	200	175	28	263	264	175	200
14	264	263	200	175	29	263	264	175	200
15	263	264	175	200	30	263	264	175	200

The performance criteria obtained from 30 executions of the K-means clustering method including accuracy, error rate, sensitivity, specificity and accuracy are presented in table 6.

The summary of the results related to the performance criteria obtained from 30 executions of the K-means clustering method is presented in table 7.

Table 6 Performance Criteria Obtained from 30 Executions of K Average Method

Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness	Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness
1	0.513	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١	١٦	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
2	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	١٧	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
3	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	١٨	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠
4	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	١٩	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١

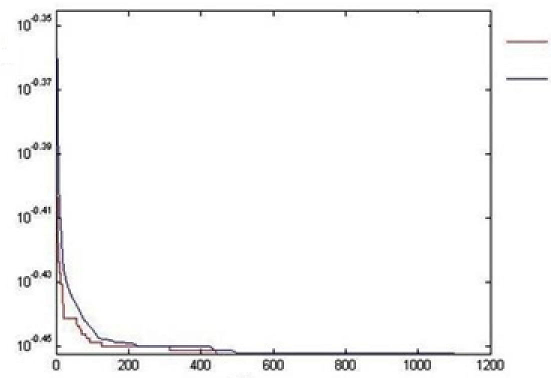
Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness	Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness
5	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١	٢٠	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
6	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢١	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠
7	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٢	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
8	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١	٢٣	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
9	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٤	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
10	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٥	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠
11	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٦	٠,٥٧٩	٠,٥٢١	٠,٠١٩	٠,٩٧٧	٠,٣٦٨
12	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١	٢٧	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
13	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٨	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
14	٠,٤٨٧	٠,٥١٣	٠,٤٦٧	٠,٥٠١	٠,٤٠٠	٢٩	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١
15	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١	٣٠	٠,٥١٣	٠,٤٨٧	٠,٥٣٣	٠,٤٩٩	٠,٤٣١

Table 7 Summarizes the Performance of the K Average Method

	True Negative	False Positive	False Negative	True Positive	Accuracy	Error Rate	Sensitivity	Specificity	Accuracy
Mean	٢٧١	٢٥٥	١٩٢	١٨٢,٧	٠,٥٠	٠,٤٩	٠,٤٨٧	٠,٥١	٠,٤١
Max.	٥١٥	٢٦٤	٣٦٨	٢٠٠	٠,٥٧	٠,٥١	٠,٥٣٣	٠,٩٧	٠,٤٣
Min.	٢٦٣	١٢	١٧٥	٧	٠,٤٨	٠,٤٢	٠,٠١٩	٠,٥	٠,٣٦
SD	٤٥,٩	٤٥,٩	٣٥,٤	٣٥,٤	٠,١	٠,٢	٠,٠٩٥	٠,٠٩	٠,٠٢

Genetic Algorithm

For the implementation of the genetic algorithm, the merit function is defined as the sum of the distances of the data to the center of each cluster. The number of chromosomes is equal to 200%, the intersection is equal to 90%, and the percentage of mutation is equal to 10%. 4- is provided.



Suitability Repetition

Diagram 1 sensitivity analysis diagram of the maximum number of repetitions and the convergence process of the genetic algorithm.

Based on the sensitivity analysis of the parameter of the maximum number of repetitions, by increasing the number of repetitions of the algorithm, the results obtained from the implementation of the genetic algorithm improve. The process of improving the results continues until the 600th repetition and after that no improvement in the results is observed. Therefore, the parameter of the maximum number of repetitions equal to 600 repetitions has been selected. The results of the anxiety table obtained from 30 executions of the genetic algorithm are presented in Table 4-31.

Based on the information in table 8, the performance criteria obtained from each execution of the genetic algorithm was calculated. The performance of the genetic algorithm can be checked based on the criteria of accuracy, error rate, sensitivity, specificity and accuracy. Table 9 presents the results of the calculation of the performance criteria of the genetic algorithm. The summary of the results related to the genetic algorithm performance criterion is presented in table 10.

Table 8 Results from the Implementation of the Genetic Algorithm

Performance	True Negative	False Positive	False Negative	True Positive	Performance	True Negative	False Positive	False Negative	True Positive
1	٥١٢	١٥	٣٦٠	١٥	١٦	٥٢٥	٢	٣٧٣	٢
2	٥٢٦	١□	٣٧٤	١□	١٧	٥١٩	٨	٣٧٤	١
3	٥١٤	١٣	٣٦٢	١٣	١٨	٥٢٢	٥	٣٧٤	١
4	٥٢٦	١	٣٧٤	١	١٩	٥٢٢	٥	٣٧٤	١
5	٥٢٦	١	٣٧٣	٢	٢٠	٥٢٠	٧	٣٧١	٤
6	٤٨٧	٤٠	٣٥١	٢٤	٢١	٥٢٧	٠	٣٧٣	٢
7	٥٢٥	٢	٣٧٣	٢	٢٢	٥٢١	٦	٣٧٥	٠
8	٣٨٥	١٤٢	٢٧٤	١٠١	٢٣	٥٢٦	١	٣٧٤	١
9	٥٠٠	٢٧	٣٦٥	١٠	٢٤	٥١٢	١٥	٣٦٧	٨
10	٥٢٥	٢	٣٧٣	٢	٢٥	٥٢٧	٠	٣٧٤	١
11	٤٩٥	٣٢	٣٥١	٢٤	٢٦	٥٢٦	١	٣٧٣	٢
12	٥٢٧	٠	٣٧٤	١	٢٧	٤٦٨	٥٩	٣٣٣	٤٢
13	٣٦٨	١٥٩	٢٧٩	٩٦	٢٨	٥٢٦	١	٣٧٣	٢
14	٤٤٢	٨٥	٣٤٤	٣١	٢٩	٥٢٤	٣	٣٧٥	٠
15	٥٢٥	٢	٣٧٣	٢	٣٠	٣٦٧	١٦٠	٢٨٠	٩٥

Table 9 Performance Criteria Obtained from 30 Executions of the Genetic Algorithm

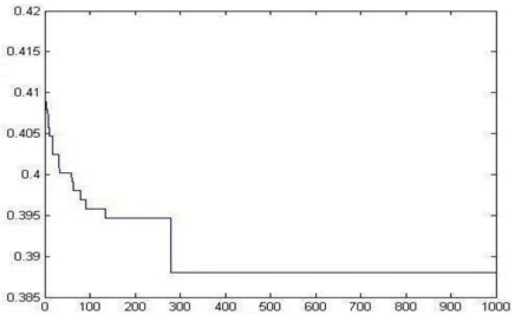
Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness	Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correctness
١	٠,٥٨٤	٠,٤١٦	٠,٠٤٠	٠,٩٧٢	٠,٥٠٠	١٦	٠,٥٨٤	٠,٤١٦	٠,٠٠٥	٠,٩٩٦	٠,٥٠٠
٢	٠,٥٨٤	٠,٤١٦	٠,٠٠٣	٠,٩٩٨	٠,٥٠٠	١٧	٠,٥٧٦	٠,٤٢٤	٠,٠٠٣	٠,٩٨٥	0.168
٣	٠,٥٨٤	٠,٤١٦	٠,٠٣٥	٠,٩٧٥	٠,٥٠٠	١٨	٠,٥٨٠	٠,٤٢٠	٠,٠٠٣	٠,٩٩١	٠,١٦٧
٤	٠,٥٨٤	٠,٤١٦	٠,٠٠٣	٠,٩٩٨	٠,٥٠٠	١٩	٠,٥٨٠	٠,٤٢٠	٠,٠٠٣	٠,٩٩١	٠,١٦٧
٥	٠,٥٨٥	٠,٤١٥	٠,٠٠٥	٠,٩٩٨	٠,٦٦٧	٢٠	٠,٥٨١	٠,٤١٩	٠,٠١١	٠,٩٨٧	٠,٣٦٤
□□	٠,٥٦٧	٠,٤٣٣	٠,٠٦٤	٠,٩٢٤	٠,٣٧٥	٢١	٠,٥٨٦	٠,٤١٤	٠,٠٠٥	١,٠٠٠	١,٠٠٠
٧	٠,٥٨٤	٠,٤١٦	٠,٠٠٥	٠,٩٩٦	٠,٥٠٠	٢٢	٠,٥٧٨	٠,٤٢٢	٠,٠٠٠	٠,٩٨٩	٠,٠٠٠
٨	٠,٥٣٩	٠,٤٦١	٠,٢٦٩	٠,٧٣١	٠,٤١٦	٢٣	٠,٥٨٤	٠,٤١٦	٠,٠٠٣	٠,٩٩٨	٠,٥٠٠
٩	٠,٥٦٥	٠,٤٣٥	٠,٠٢٧	٠,٩٤٩	٠,٢٧٠	٢٤	٠,٥٧٦	٠,٤٢٩	٠,٠٢١	٠,٩٧٢	٠,٣٤٨
١٠	٠,٥٨٤	٠,٤١٦	٠,٠٠٥	٠,٩٩٦	٠,٥٠٠	٢٥	٠,٥٨٥	٠,٤١٥	٠,٠٠٣	١,٠٠٠	١,٠٠٠
١١	٠,٥٧٥	٠,٤٢٥	٠,٠٦٤	٠,٩٣٩	٠,٤٢٥	٢٦	٠,٥٨٥	٠,٤١٥	٠,٠٠٥	٠,٩٩٨	٠,٦٦٧
١٢	٠,٥٨٥	٠,٤١٥	٠,٠٠٣	١,٠٠٠	١,٠٠٠	٢٧	٠,٥٦٥	٠,٤٣٥	٠,١١٢	٠,٨٨٧	٠,٤١٦
١٣	٠,٥١٤	٠,٤٨٦	٠,٢٥٦	٠,٦٩٨	٠,٣٧٦	٢٨	٠,٥٨٥	٠,٤١٥	٠,٠٠٥	٠,٩٩٨	٠,٦٦٧
١٤	٠,٥٢٤	٠,٤٧٦	٠,٠٨٣	٠,٨٣٩	٠,٢٦٧	٢٩	٠,٥٨١	٠,٤١٩	٠,٠٠٠٠	٠,٤٩٤	٠,٠٠٠
١٥	٠,٥٨٤	٠,٤١٦	٠,٠٠٥	٠,٩٩٦	٠,٥٠٠	٣٠	٠,٥١٢	٠,٤٨٨	٠,٢٥٣	٠,٦٩٦	٠,٣٧٣

Table 10 Summary of Genetic Algorithm Performance

	True Negative	False Positive	False Negative	True Positive	Accuracy	Error Rate	Sensitivity	Specificity	Correctness
Mean	٥٠٠,٥	٢٦,٥	٣٥٨,٨	١٦,٢	٠,٥٧٣	٠,٤٢٧	٠,٠٤٣	٠,٩٥٠	٠,٤٥٣
Max.	٥٢٧	١٦٠	٣٧٥	١٠١	٠,٥٨٦	٠,٤٨٨	٠,٢٦٩	١,٠٠٠	١,٠٠٠
Min.	٣٦٧	٠	٢٧٤	٠	٠,٥١٢	٠,٤١٤	٠,٠٠٠	٠,٦٩٦	59555
SD	٤٧,٣	٤٧,٣	٢٩,٤	٢٩,٤	٠,٠٢١	٠,٠٢١	٠,٠٧٨	٠,٠٩٠	٠,٢٥٣

Distance Based Morchegan Algorithm

To implement the algorithm of Morchegan based on distance by trial and error method, the initial value of the number of Morchegan was defined as 200 Morchegan. One of the performance evaluation criteria of innovative algorithms is its convergence process. The process of convergence and improvement of the results in the algorithm of Morchegan based on the distance is presented in the graph 2.



Suitability Repetition

Diagram 2 is the sensitivity analysis diagram of the maximum number of repetitions and convergence of the distance-based Morchegan algorithm.

The results of the disturbance table (based on the number) obtained from 30 times of running the distance-based Morchegan algorithm are presented in table (11). Based on the information in table (11), the performance criteria obtained from each execution of the distance-based Morchegan algorithm was calculated.

The performance of the distance-based Morchegan algorithm can be checked based on the criteria of accuracy, error rate, sensitivity, specificity and accuracy. Table 12 presents the results of calculating the performance criteria of the distance-based Morchegan algorithm. The summary of the results related to the distance based Morchegan algorithm performance criteria is presented in table 13.

Entropy-based Morchegan Algorithm

To implement the entropy-based Morchegan algorithm, the initial values of the number of repetitions and the number of Morchegan have been determined as in the distance-based Morchegan algorithm. The results of the disturbance table (based on the number) obtained from 30 times of running the entropy-based Morchegan algorithm are presented in table 14.

Table 15 presents the performance rate of each entropy-based Morchegan algorithm implementation, including accuracy, error rate, sensitivity, specificity and accuracy.

The summary of the results related to the performance criteria obtained from 30 executions of the entropy-based Morchegan algorithm is presented in table 16.

Table 11 The Results of the Implementation of the Distance-based Morchegan Algorithm

Performance	True Negative	False Positive	False Negative	True Positive	Performance	True Negative	False Positive	False Negative	True Positive
١	٤٥٩	٦٨	٢٧٠	١٠٥	١٦	٤٩٧	٣٠	٣٢٢	٥٣
٢	٤٦٣	٦٤	٢٨٦	٨٩	١٧	٤٦٤	٦٣	٢٨٨	٨٧
٣	٤٦٩	٥٨	٢٨٩	٨٦	١٨	٤٩٠	٣٧	٣١٥	٦٠
٤	٤٧٤	٥٣	٢٩٥	٨٠	١٩	٥١٦	١١	٣٤٥	٣٠
٥	٤٣٢	٩٥	٢٥٥	١٢٠	٢٠	٥١٨	٩	٣٤٦	٢٩
٦	٥٠١	٢٦	٣٢٨	٤٧	٢١	٤٢١	١٠٦	٢٣٧	١٣٨
٧	٥٠٦	٢١	٣٣٧	٣٨	٢٢	٤٧٥	٥٢	٣٠١	٧٤
٨	٥١٣	١٤	٣٤٤	٣١	٢٣	٤٦٣	٦٤	٢٨٠	٩٥
٩	٥٠١	٢٦	٣٢٣	٥٢	٢٤	٥١٨	٩	٣٥١	٢٤
١٠	٤٦٣	٦٤	٢٨٦	٨٩	٢٥	٥١٩	٨	٣٤٧	٢٨
١١	٤٩١	٣٦	٣١٨	٥٧	٢٦	٤١٨	١٠٩	٢٤٠	١٣٥
١٢	٥١٥	١٢	٣٤٣	٣٢	٢٧	٤٨٨	٣٩	٣١٠	٦٥
١٣	٥٠٦	٢١	٣٣١	٤٤	٢٨	٤٦٢	٦٥	٢٨٥	٩٠
١٤	٥١٤	١٣	٣٤١	٣٤	٢٩	٥٠٣	٢٤	٣٣٥	٤٨
١٥	٤٧٣	٥٤	٢٩٤	٨١	٣٠	٥٠٥	٢٢	٣٢٨	٤٧

Table 12 Performance Criteria Obtained from 30 Executions of Distance-Based Morchegan Algorithm

Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correction	Performance	Accuracy	Error Rate	Sensitivity	Specificity	Correction
1	0.625	0.375	0.280	0.871	0.607	16	0.610	0.390	0.141	0.943	0.639
2	0.612	0.388	0.237	0.879	0.582	17	0.611	0.389	0.232	0.880	0.580
3	0.615	0.385	0.229	0.890	0.597	18	0.610	0.390	0.160	0.930	0.619
4	0.614	0.386	0.213	0.899	0.602	19	0.605	0.395	0.080	0.979	0.732
5	0.612	0.388	0.320	0.820	0.558	20	0.606	0.394	0.077	0.983	0.763
6	0.608	0.392	0.125	0.951	0.644	21	0.620	0.380	0.368	0.799	0.566
7	0.603	0.397	0.101	0.960	0.644	22	0.609	0.391	0.197	0.901	0.587
8	0.603	0.397	0.083	0.973	0.689	23	0.619	0.381	0.253	0.879	0.597
9	0.613	0.387	0.139	0.951	0.667	24	0.601	0.399	0.064	0.983	0.727
10	0.612	0.388	0.237	0.879	0.582	25	0.606	0.394	0.075	0.985	0.779
11	0.608	0.392	0.152	0.932	0.613	26	0.613	0.387	0.360	0.793	0.553
12	0.606	0.394	0.085	0.977	0.727	27	0.613	0.387	0.173	0.926	0.625
13	0.610	0.390	0.117	0.960	0.677	28	0.612	0.388	0.240	0.877	0.581
14	0.608	0.392	0.091	0.975	0.723	29	0.605	0.395	0.125	0.954	0.667
15	0.614	0.386	0.216	0.898	0.600	30	0.612	0.388	0.125	0.958	0.681

Table 13 Summary of the Performance of the Distance-based Morchegan Algorithm

	True Negative	False Positive	False Negative	True Positive	Accuracy	Error rate	Sensitivity	Specificity	Correctness
Mean	٤٨٤,٦	٤٢,٤	٣٠٩,٠	٦٦,٣	٠,٦١١	٠,٣٨٩	٠,١٧٧	٠,٩١٩	٠,٦٤٠
Max.	٥١٩	١٠٩	٣٥١	١٣٨	٠,٦٢٥	٠,٣٩٩	٠,٣٦٨	٠,٩٨٥	٠,٧٧٨
Min.	٤١٨	٨	٢٣٧	٢٤	٠,٦٠١	٠,٣٧٥	٠,٠٦٤	٠,٧٩٣	٠,٥٥٣
SD	٢٨,٩	٢٨,٩	٣٢,٥	٣٢,٣	٠,٠٠٥	٠,٠٠٥	٠,٠٨٦	٠,٠٥٥	٠,٠٦٤

Table 14 Results of the Implementation of the Entropy-based Morchegan Algorithm

Performance	True Negative	False Positive	False Negative	True Positive	Performance	True Negative	False Positive	False Negative	True Positive
١	٥٠٥	٢٢	٣٤٨	٢٧	١٦	٥٠٥	٢٢	٣٥٧	١٨
٢	٢٤٤	٢٨٣□	١٦٠	٢١٥	١٧	٤٦٢	٦٥	٣٢٩	٤٦
٣	٥٢٢	٥	٣٧١	٤	١٨	٥٢١	٦	٣٧٢	٣
٤	٤٦٠	٦٧	٣٢٥	٥٠	١٩	٢٩١	٢٣٦	٢١٠	١٦٥
٥	٥٢١	٦	٣٧٣	٢	٢٠	٥١١	١٦	٣٦٨	٧
□	٥٢٢	٥	٣٦٨	٧	٢١	٥٢٣	٤	٣٧٣	٢□
٧	٤٨١	٤٦	٣٣٢	٤٣	٢٢	٥٢٥	٢□	٣٧٤	١
٨	٥١٨	٩	٣٦٩	٦	٢٣	٥٢٦	١	٣٧٤	١□
٩	٥٢٠	٧□	٣٧٤	١	٢٤	٢٨٨	٢٣٩□	١٩٠	١٨٥
١٠	٥١٨	٩	٣٧٤	١	٢٥	٥٢٣	٤	٣٧٤	١
١١	٥٢٣	٤□	٣٧٤	١□	٢٦	٣٣١	١٩٦	٢٢٢	١٥٣
١٢	٥٢٠	٧	٣٧٣	٢	٢٧	٤٩٠	٣٧	٣٦٢	١٣□
١٣	٢٨٦	٢٤١□	١٧٧	١٩٨	٢٨	٥٠٤	٢٣□	٣٤٩	٢٦
١٤	٤٥٨	٦٩	٣٢٧	٤٨	٢٩	٤٥٣	٧٤	٣٢٣	٥٢
١٥	٥٢٣	٤	٣٧٤	١	٣٠	٤١٩	١٠٨	٣٠٥	٧٠

Table 15 Results of the Implementation of the Entropy-Based Morchegan Algorithm

Performance	Accuracy	Error Rate	Sesesivity	Specificity	Correction	Performance	Accuracy	Error Rate	Sesesivity	Specificity	Correction
1	0.590	0.410	0.072	0.958	0.551	16	0.580	0.420	0.048	0.958	0.450
2	0.509	0.491	0.573	0.463	0.432	17	0.563	0.437	0.123	0.877	0.414
3	0.583	0.417	0.011	0.991	0.444	18	0.581	0.419	0.008	0.989	0.333
4	0.565	0.435	0.133	0.873	0.427	19	0.506	0.494	0.440	0.552	0.411
5	0.580	0.420	0.005	0.989	0.250	20	0.574	0.426	0.019	0.970	0.304
6	0.586	0.414	0.019	0.991	0.583	21	0.582	0.418	0.005	0.992	0.333
7	0.581	0.419	0.115	0.913	0.483	22	0.583	0.417	0.003	0.996	0.333
8	0.581	0.419	0.016	0.983	0.400	23	0.584	0.416	0.003	0.998	0.500
9	0.578	0.422	0.003	0.987	0.125	24	0.524	0.476	0.493	0.546	0.436
10	0.579	0.425	0.003	0.983	100	25	0.581	0.419	0.003	0.992	0.200
11	0.581	0.419	0.003	0.992	0.200	26	0.537	0.463	0.408	0.628	0.438
12	0.579	0.421	0.005	0.987	0.222	27	0.558	0.442	0.035	0.930	0.260
13	0.537	0.463	0.528	0.543	0.451	28	0.588	0.412	0.069	0.956	0.531
14	0.561	0.439	0.128	0.869	0.410	29	0.560	0.440	0.139	0.860	0.413
15	0.581	0.419	0.003	0.992	0.200	30	0.542	0.458	0.187	0.795	0.393

Table 16 Summary of the Performance of the Entropy-based Morchegan Algorithm

	True Negative	False Positive	False Negative	True Positive	Accuracy	Error Rate	Sensitivity	Specificity	Correctness
Mean	٤٦٦,٤	٦٠٦	٣٣٠	٤٥,٠	٠,٥٦٧	٠,٤٣٣	٠,١٢٠	٠,٨٨٥	٠,٣٦٨
Max.	٥٢٦	٢٨٣	٣٧٤	٢١٥	٠,٥٩٠	٠,٤٩٤	٠,٥٧٣	٠,٩٩٨	٠,٥٨٣
Min.	٢٤٤	١	١٦٠	١	٠,٥٠٦	٠,٤١٠	٠,٠٠٣	٠,٤٦٣	٠,١٠٠
SD	٨٦,٢	٨٦,٢	٦٦,٥	٦٦,٥	٠,٢٣	٠,٢٣	٠,١٧٧	٠,١٦٤	٠,١٢٥

Hypothesis Test

In order to test research hypotheses and compare the performance of fraud detection methods, the results of 30 executions of each method including accuracy, error rate, sensitivity, specificity and accuracy have been compared. To check the validity of the first hypothesis, the F test or one-way analysis of variance was used for K independent groups. F test or one-way analysis of variance is used to test the difference in the mean of a variable in more than two groups.

Homogeneity of Variances Test

To choose the right test, the homogeneity of variances must be checked first. Therefore, Lone variance homogeneity test was used. Table 17 shows the results of the Lone variance homogeneity test. According to the significance level smaller than 0.05, the results obtained from Lon's test show that the variance of the groups for the performance criteria of accuracy, error rate, sensitivity, specificity and accuracy are not equal. Therefore, one-way analysis of variance should be performed with the assumption of heterogeneity of variances.

Table 17 Homogeneity Of Variances Test

	Loon Statistics	Degree of Freedom 1	Degree of Freedom 2	Significance Level
Accuracy	١٢,٩٢٢	٥	١٧٤	.٠٠٠
Sensitivity Error	١٢,٩٢٢	٥	١٧٤	.٠٠٠
	١١,٦٣٧	٥	١٧٤	.٠٠٠
Specificity	١٢,٦٦٤	٥	١٧٤	.٠٠٠
Correctness	٢٠,٦٧٣	5	174	.٠٠٠

Research Hypothesis Test

The hypothesis of the research is presented as follows: To discover the possibility of fraud in financial statements, Morchegan algorithm is more effective in comparison with regression method, neural network, K average and genetic algorithm.

To test the hypothesis of the research, the average results of 30 times of each of the methods, including accuracy, error rate, sensitivity, specificity and accuracy, using one-way variance test and t method. Two samples have been compared. The two-sample t-test is a conservative test that compares the pairs of groups based on the value of the t-statistic. In the following, the results of the said test are presented in table (18) for each of the performance criteria of regression method, neural network, K average, genetic algorithm and Morchegan algorithm. According to the significance level smaller than 0.05 for each of the criteria of accuracy, error rate, sensitivity, specificity and accuracy, it can be concluded that there is a difference between the performance of different regression methods, neural network, K average, genetic algorithm and Morchegan algorithm in detecting the possibility of fraud. There is a significant Morchegan difference.

In order to evaluate the performance of pairs of fraud detection methods, the results of multiple comparisons are examined, which are presented in the following to separate the performance criteria including accuracy, error rate, sensitivity, specificity and accuracy.

Table 18 One-Way Variance Analysis

		Sum of Squares	Degree of Freedom	Average of Squares	F Statistics	Significance Level
Accuracy	Between Group	.231	5	0.046	١٣٧,٦٧٦	.٠٠٠
	Within Group	.٠٥٨	١٧٤	.٠٠٠		
	Total	.٢٩٠	١٧٩			
Error	Between Group	.٢٣١	٥	.٢٣٦	١٣٧,٦٧٦	.٠٠٠
	Within Group	.٠٥٨	١٧٤	.٠٠٠		
	Total	.٢٩٠	١٧٩			
Sensitivity	Between Group	٣,٩٨٩	٥	.٧٩٨	٧١,٥١٠	.٠٠٠
	Within Group	١,٩٤١	١٧٤	.٠١١		
	Total	٥,٩٣٠	١٧٩			
Specificity	Between Group	٤,٣٨٦	٥	.٨٧٧	١٠٦,٩٢٩	.٠٠٠
	Within Group	١,٤٢٧	١٧٤	.٠٠٨		
	Total	٥,٨١٣	١٧٩			
Correctness	Between Group	٢,٧٧٢	٥	.٥٥٤	٣٧,٥٠١	.٠٠٠
	Within Group	٢,٥٧٣	١٧٤	.٠١٥		
	Total	٥,٣٤٥	١٧٩			

Accuracy Criterion

Table 19 presents the results of the multiple comparison of the accuracy criterion resulting from the one-way analysis of variance test for K independent groups for the results of 30 times of the regression, neural network, K average and genetic algorithm and the Morchegan algorithm in detecting the possibility of fraud.

Regarding the distance-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average accuracy criteria of the distance-based Morchegan algorithm and the neural network method. In addition, the significance level smaller than 0.05 also shows that there is a significant difference between the average accuracy criterion of the distance-based Morchegan algorithm and logistic regression methods, K-means and genetic algorithm. Therefore, the accuracy of the Morchegan algorithm based on distance is higher than logistic regression, K mean and genetic algorithm. Regarding the entropy-based Morchegan algorithm, a significance level greater than 0.05 shows that there is no significant difference between the average accuracy criteria of the entropy-based Morchegan algorithm and the genetic algorithm. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the average accuracy criterion of the entropy-based Morchegan algorithm and logistic regression methods, neural network and K average. Thus, the accuracy of the logistic regression method and neural network is higher than the accuracy of the entropy-based Morchegan algorithm and the accuracy of the entropy-based Morchegan algorithm is higher than the K-means method.

Table 19 Results of Multiple Comparisons of Accuracy Criteria

Group 4	Group 8	Group Mean Difference ٤ and 8	Error Standard	Significance Level	٪٣٠ Confidence Interval	
					Upper Level	Lower Level
Morchegan Algorithm based on the Distance	Logistic Regression of K-Means Neural Network Genetic Algorithm	٠١١□	□٠٠٠□	٠,٧٧٧	□٠٠٨□	٠,٠١٤
		□٠٠٧□	□٠٠٤□	٠,٨٣٧	□□٠٠٧□	□٠٢٢□
		□١٠٦□	□٠٠٣□	□٠٠٠□	□٠٩٥□	٠,١١٨
		□٠٣٧□	□٠٠٤	□٠٠٠□	□٠٢٤	□٠٥٠□
Morchegan Algorithm based on Entropy	Logistic Regression of Neural Network Average K Genetic Algorithm	□□٠٣١□	□٠٠٤□	□٠٠٠□	□□٠٤٥□	-٠,٠١٨□
		□□٠□٦□	□٠٠٦□	□٠٠٠	□□٠٥٥□	□□٠١٧□
		□٠٦٣	□٠٠٥□	□٠٠٠	□٠٤٦□	□٠٧٩□
		□□٠٠٥	□٠٠٥	□٩٩٦	□□٠٢٣	□٠١١

Error Criterion

Table 20 shows the results of the multiple comparison of the error criterion resulting from the one-way analysis of variance test for K independent groups for the results of 30 times of the regression, neural network, K average, genetic algorithm and Morchegan algorithm in discovering the possibility of fraud. Regarding the distance-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average error criterion of the distance-based Morchegan algorithm and the neural network method. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the average error criterion of the distance-based Morchegan algorithm and logistic regression methods, K average and genetic algorithm. Thus, the error of the Morchegan algorithm based on the distance is less than the logistic regression methods, K-means and genetic algorithm. Regarding the entropy-based Morchegan algorithm, a significance level greater than 0.05 shows that there is no significant difference between the average error criterion of the entropy-based Morchegan algorithm and the genetic algorithm. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the average error criterion of the entropy-based Morchegan algorithm and logistic regression, neural network and K-means methods. Thus, the error of the logistic regression and neural network method is less than the error of the entropy-based Morchegan algorithm and the error of the entropy-based Morchegan algorithm is less than the K-means method.

Table 20 the Results of Multiple Comparison of the Error Criterion

Group 1 Group 2	Average Difference of Group 1 and 2	Error Standard	Significance Level	95% Confidence Interval	
				Upper Level	Lower Level
logistic regression	-.٠١١	.٠٠٠	.٠٠٠	-.٠١٤	-.٠٠٨
Morchegan Algorithm based on	-.٠٠٧	.٠٠٤	.٨٣٧	-.٠٢٢	.٠٠٧

Neural Network			.٠٠٠		
K average on the distance	-.١٠٦	.٠٠٣	.٠٠٠	-.١١٨	-.٠٩٥
Genetic algorithm	-.٠٣٧	.٠٠٤		-.٠٥٠	-.٠٢٤
Logistic Regression of Neural Network based on Morchegan Algorithm	.٠٣١ .٠٣٦	.٠٠٤ .٠٠٦	.٠٠٠ .٠٠٠	.٠١٨ .٠١٧	.٠٤٥ .٠٥٥
	-.٠٦٣	.٠٠٥	.٠٠٠	-.٠٧٩	-.٠٤٦
	.٠٠٥	.٠٠٥	.٩٩٦	-.٠١١	.٠٢٣

Sensitivity Criterion

Table 21 presents the results of the multiple comparison of the sensitivity criterion resulting from the one-way analysis of variance test for K independent groups for the results of 30 times of the regression, neural network, K average and genetic algorithm and the Morchegan algorithm in discovering the possibility of fraud.

Regarding the distance-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average sensitivity criterion of the distance-based Morchegan algorithm and the neural network. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the average sensitivity criterion of the distance-based Morchegan algorithm and logistic regression methods, K average and genetic algorithm. Thus, the sensitivity of logistic regression methods and genetic algorithm is lower than the sensitivity of distance-based Morchegan algorithm. The sensitivity of the K-means method is higher than the sensitivity of the distance-based Morchegan algorithm. Regarding the entropy-based Morchegan algorithm, a significance level greater than 0.05 shows that there is no significant difference between the average sensitivity of the entropy-based Morchegan algorithm, logistic regression, neural network, and genetic algorithm. In addition, a significance level smaller than 0.05 indicates that there is a significant difference between the average sensitivity criterion of the entropy-based Morchegan algorithm and the K-means method. Thus, the sensitivity of the K average method is higher than the sensitivity of the entropy-based Morchegan algorithm.

Table 21 the Results of Multiple Comparison of the Sensitivity Criterion

Group 1	Group 2	Group Mean Difference 1 and 2	Error Standard	Significance Level	95% Confidence Interval	
					Upper Level	Lower Level
Morchegan Algorithm based on Distance	Logistic Regression	.١١٧	.٠١٥	.٠٠٠	.٠٦٧	.١٦٨
	Neural Network	-.٠٤٣	.٠٢٦	.٨٠١	-.١٢٣	.٠٣٦
	K Average	-.٣١٠	.٠٢٣	.٠٠٠	-.٣٨١	-.٢٣٩
	Genetic Algorithm	.١٣٣	.٠٢١	.٠٠٠	.٠٦٨	.١٩٨
Morchegan Algorithm based on Entropy	Logistic Regression	.٠٦١	.٠٣٢	.٦٥٤	-.٠٤١	.١٦٤
	Neural Network	-.١٠٠	.٠٣٨	.١٦٨	-.٢١٨	.٠١٨
	K Average	-.٣٦٧	.٠٣٦	.٠٠٠	-.٤٨	-.٢٥٣
	Genetic Algorithm	.٠٧٦	.٠٣٥	.٤٢٦	-.٠٣٣	.١٨٦

Feature Criteria

Table 22 presents the result of the multiple comparison of the feature criterion resulting from the one-way analysis of variance test for K independent groups for the results of 30 times of the regression, neural network, K average and genetic algorithm and the Morchegan algorithm in discovering the possibility of fraud. Regarding the distance-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average feature criteria of distance-based Morchegan algorithm, neural network and genetic algorithm. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the mean feature criteria of distance-based Morchegan algorithm and logistic regression methods and K mean. Thus, the characteristic of the logistic regression method is more and the characteristic of the K-mean method is less than the characteristic of the distance-based Morchegan algorithm. Regarding the entropy-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average feature criteria of the entropy-based Morchegan algorithm, the neural network, and the genetic algorithm. In addition, a significance level smaller than 0.05 shows that there is a significant difference between the mean feature criterion of the entropy-based Morchegan algorithm and logistic regression methods, K mean. Therefore, the characteristic of the logistic regression method is more and the characteristic of the K-mean method is less than the characteristic of the entropy-based Morchegan algorithm.

Table 22 the Results of the Implementation of the Multiple Comparison of the Feature Criteria

Group 1	Group 2	Average difference of group 1 and 2	Error Standard	Significance Level	95% Confidence Interval	
					Upper Level	Lower Level
Morchegan Algorithm based on Distance	Logistic Regression	-.٠٦٣	.٠١٠	.٠٠٠	-.٠٩٥	-.٠٣١
	Neural Network	.٠٣٦	.٠١٥	.٢٢٧	-.٠٠٨	.٠٨٢
	K Average	.٤٠٣	.٠١٨	.٠٠٠	.٣٤٥	.٤٦١
	Genetic Algorithm	-.٠٣٠	.٠١٩	.٨٥٨	-.٠٨٩	.٠٢٨
Morchegan Algorithm based on Entropy	Neural Network	-.٠٩٧	.٠٢٩	.٠٤٠١,٠	-.١٩٣	-.٠٠٢
	Logistic Regression	.٠٠٢	.594		-.٠٩٧	.١٠٢
	K Average	.٣٦٩	.٠٣٣	.٠٠٠	.٢٦٤	.٤٧٤
	Genetic Algorithm	-.٠٦٤	.٠٣٤	.٦٣١	-.١٦٩	.٠٤٠

Correctness Criteria

Table 23 presents the results of the multiple comparison of the accuracy criterion resulting from the one-way analysis of variance test for K independent groups for the results of 30 times of the regression, neural network, K average and genetic algorithm and the Morchegan algorithm in discovering the possibility of fraud.

Regarding the distance-based Morchegan algorithm, the significance level greater than 0.05 shows that there is no significant difference between the average feature criteria of the distance-based Morchegan algorithm and the neural network. In addition, a significance level smaller than 05.05 shows that there is a significant difference between the average accuracy criteria of the distance-based Morchegan algorithm and logistic regression methods, K average and genetic algorithm. Therefore, the accuracy of the distance-based Morchegan algorithm is less than the logistic regression method and more than the K-means and genetic algorithm methods. Regarding the entropy-based Morchegan algorithm, a significance level greater than 0.05 shows that there is no significant difference between the average accuracy criteria of the entropy-based Morchegan algorithm, the K-means method, and the genetic algorithm. In addition, the significance level smaller than 0.05 also shows that there is a significant difference between the average accuracy criterion of the entropy-based Morchegan algorithm and the methods of logistic regression and artificial neural network. Therefore, the accuracy of the entropy-based Morchegan algorithm is lower than the logistic regression method and the artificial neural network. Based on the findings of this section regarding the performance of the distance-based Morchegan algorithm and the entropy-based Morchegan algorithm, the first hypothesis of the research cannot be confirmed.

Table 23 Results of the Multiple Comparison of Accuracy Criteria

Group 1	Group 2	Average difference of group 1 and 2	Error Standard	Significance Level	95% Confidence Interval	
					Upper Level	Lower Level
Morchegan Algorithm based on Distance	Logistic Regression	□□٠٦٩□	□٠١١	□٠٠٠	□□١٠٦□	□□٠٣٢□
	Neural Network	□٠٥٩	□٠١٧	□٠١٤□	□٠٠٧□	٠,١١٢
	K Average	٠,٢٢٤	□٠١٢	□٠٠٠	□١٨٦□	□٢٦٢
	Genetic Algorithm	٠,١٨٧	□٠٤٧	□٠٠٦	□٠٣٧□	٠,٣٣٧
Morchegan Algorithm based on Entropy	Neural Network	-٠,٣٤١□	□٠٢٢□	□٢٢٢□	-٠,٤١٤	-٠,٢٦٩
	Logistic Regression	-٠,٢١٢□	□٠٢٦	□٠٠٠□	-٠,٢٩٣□	□□١٣٢□
	K Average	□□٠٤٧□	□٠٢٣□	□٥١٩□	-٠,١٢□	□٠٢٥
	Genetic Algorithm	□□٠٨٤□	□٠٥١	□٨١٦	-٠,٢٤٤□	□٠٧٤

Examining the Results of the Research Hypothesis Test

Research Hypothesis

The hypothesis of the research is to evaluate the effectiveness of discovering the possibility of fraud in financial statements using the Morchegan algorithm in comparison with the regression method, neural network, K-means and genetic algorithm. The results of the test of this hypothesis regarding the performance of the distance-based Morchegan algorithm are presented as follows:

1) The accuracy of distance-based Morchegan algorithm (average = 61.1%) is not significantly different from the accuracy of neural network algorithm (average = 60.3%). In other words, the performance of distance-based Morchegan algorithm in the correct classification of companies is similar to the performance of neural network algorithm.

2) The accuracy of distance-based Morchegan algorithm (average = 61.1 percent) is significantly better than the accuracy of logistic regression methods (average = 59.9 percent), genetic algorithm (average = 57.3) and k average (average = 4). /50 percent). In other words, the performance of distance-based Morchegan algorithm in the correct classification of companies is significantly better than the performance of logistic regression methods, genetic algorithm and k-means.

3) The error rate of the distance-based Morchegan algorithm (average = 38.9 percent) is not significantly different from the error rate of the neural network algorithm (average = 39.7 percent). In other words, the performance of distance-based Morchegan algorithm in misclassifying companies is similar to the performance of neural network algorithm.

4) The error rate of distance-based Morchegan algorithm (average = 38.9 percent) is significantly lower than the error rate of logistic regression methods (average = 40.1 percent), genetic algorithm (average = 42.7 percent) and k average (average = 49.6 percent). In other words, the error in the classification of companies in the distance-based Morchegan algorithm is significantly less than the error in the classification of companies using logistic regression, genetic algorithm and k-means methods.

5) The sensitivity of the distance-based Morchegan algorithm (average = 17.7%) is not significantly different from the sensitivity of the neural network algorithm (average = 22%). In other words, the performance of the distance-based Morchegan algorithm in identifying suspected fraud companies is similar to the performance of the neural network algorithm.

6) The sensitivity criterion of distance-based Morchegan algorithm (mean = 17.7 percent) is significantly higher than the sensitivity criterion of logistic regression method (mean = 5.9 percent) and genetic algorithm (mean = 4.3 percent). In other words, the performance of distance-based Morchegan algorithm in identifying companies suspected of fraud is significantly better than the performance of logistic regression methods and genetic algorithm. In addition, the sensitivity of the distance-based Morchegan algorithm is significantly lower than the sensitivity of the K-means method (mean = 48.7%). In other words, the distance-based Morchegan algorithm is significantly more unsuccessful than the K-means method in identifying companies suspected of fraud.

7) The feature criterion of distance-based Morchegan algorithm (mean = 91.9 percent) is not

significantly different from the feature of neural network algorithm (mean = 88.2 percent) and genetic algorithm (mean = 95 percent). In other words, the performance of distance-based Morchegan algorithm in identifying healthy companies is similar to the performance of neural network algorithm and genetic algorithm.

8) The feature criterion of distance-based Morchegan algorithm (mean = 91.9 percent) is significantly higher than the feature of the K-means method (mean = 51.6 percent). In other words, the performance of the distance-based Morchegan algorithm in identifying healthy companies is significantly better than the performance of the K-means method. In addition, the feature of the distance-based Morchegan algorithm (mean = 38.9%) is significantly lower than the feature of the logistic regression method (mean = 98.3%). In other words, the distance-based Morchegan algorithm is significantly more unsuccessful than the logistic regression method in identifying healthy companies.

9) The accuracy of the distance-based Morchegan algorithm (average = 64%) is not significantly different from the accuracy of the neural network algorithm (average = 58%). In other words, the performance of the distance-based Morchegan algorithm regarding the correctness of identifying companies as suspected of fraud is similar to the performance of the neural network algorithm.

10) The accuracy of distance-based Morchegan algorithm (average = 64%) is significantly higher than the accuracy of K-means methods (average = 41.5%) and genetic algorithm (average = 45.3%). In other words, the performance of distance-based Morchegan algorithm is significantly better than the performance of K-means and genetic algorithm in terms of the accuracy of identifying companies as suspected fraud companies. In addition, the accuracy of distance-based Morchegan algorithm (average = 64%) is significantly lower than the accuracy criterion of the logistic regression method (average = 71%). In other words, the distance-based Morchegan algorithm is significantly more unsuccessful than the logistic regression method in correctly identifying companies as suspected of fraud. Table (24) summarizes the results of the research hypothesis test regarding the performance of the distance-based Morchegan algorithm.

Based on the results obtained from the comparison of the performance of the distance-based Morchegan algorithm with other methods, the research hypothesis regarding the performance of the distance-based Morchegan algorithm cannot be confirmed.

In general, the results of the first hypothesis test regarding the superior performance of the artificial neural network algorithm compared to the logistic regression method are consistent with the findings of the researchers (2022). The results of Santoso et al.'s (2019) research are consistent.

Table 24 summarizes the results of the research hypothesis test regarding the performance of the distance-based Morchegan algorithm.

Similar Methods	Comparison of Methods	Performance Criteria
Neural Network	Logistic Regression, K Average, Genetic Algorithm < Morchegan Algorithm	Accuracy
Neural Network	Logistic Regression, K Average, Genetic Algorithm > Morchegan Algorithm	Error rate
Neural Network	Logistic Regression, Genetic Algorithm < Morchegan Algorithm < K Average	Sensitivity
Neural Network Genetic Algorithm	K Mean < Morchegan Algorithm < Logistic Regression	Specificity
Neural Network	K Mean, Genetic Algorithm < Morchegan Algorithm < Logistic Regression	Correctness

In the following, the results of the test of Pezoh's hypothesis regarding the performance of the Morchegan algorithm based on entropy are presented:

1) The accuracy of the Morchegan algorithm based on entropy (average = 56.7%) is not significantly different from the accuracy of the genetic algorithm (average = 57.3%). In other words, the performance of the entropy-based Morchegan algorithm in the correct classification of companies is similar to the performance of the genetic algorithm.

2) The accuracy of the entropy-based Morchegan algorithm (average = 56.7%) is significantly better

than the accuracy of the k-means method (average = 50.4%). In other words, the performance of the entropy-based Morchegan algorithm in the correct classification of companies is significantly better than the performance of the k-means method.

In addition, the accuracy criterion of the entropy-based Morchegan algorithm (average = 56.7%) is significantly lower than the accuracy of the logistic regression method (average = 59.9%) and the neural network algorithm (average = 60.3%). In other words, entropy-based Morchegan algorithm is significantly unsuccessful in the correct classification of companies compared to logistic regression methods and neural network algorithm.

3) The error rate of the entropy-based Morchegan algorithm (average = 43.3%) is not significantly different from the error rate of the genetic algorithm (average = 42.7%). In other words, the performance of the entropy-based Morchegan algorithm in misclassifying companies is similar to the performance of the genetic algorithm.

4) The error rate of the entropy-based Morchegan algorithm (mean = 43.3%) is significantly lower than the error rate of the k-means method (mean = 49.6%). In other words, the error in classifying companies in the entropy-based Morchegan algorithm is significantly less than the error in classifying companies using the k-means method. In addition, the error rate of the entropy-based Morchegan algorithm (average = 43.3%) is significantly higher than the error rate of logistic regression methods (average = 40.1%) and the artificial neural network algorithm (average = 39.7%). In other words, the error in classifying companies in the entropy-based Morchegan algorithm is significantly more than the error in classifying companies using logistic regression and neural network algorithms.

5) The sensitivity of the entropy-based Morchegan algorithm (average = 12 percent) with the sensitivity of logistic regression methods (average = 5.9 percent), neural network algorithm (average = 22 percent) and genetic algorithm (average = 4.3 percent), the difference It has no meaning. In other words, the performance of entropy-based Morchegan algorithm in identifying suspected fraud companies is similar to the performance of logistic regression methods, neural network algorithm, and genetic algorithm.

6) The sensitivity of the entropy-based Morchegan algorithm (average = 12%) is significantly lower than the sensitivity of the K-average method (average = 48.7%). In other words, the entropy-based Morchegan algorithm is significantly more unsuccessful than the K-means method in identifying companies suspected of fraud.

7) The feature criterion of the entropy-based Morchegan algorithm (average = 88.5%) is not significantly different from the feature of the neural network algorithm (average = 88.2%) and the genetic algorithm (average = 95%). In other words, the performance of the entropy-based Morchegan algorithm in identifying healthy companies is similar to the performance of the neural network algorithm and the genetic algorithm.

8) The characteristic criterion of the entropy-based Morchegan algorithm (mean = 88.5%) is significantly higher than the characteristic of the K-means method (mean = 51.6%). In other words, the performance of the entropy-based Morchegan algorithm in identifying healthy companies is significantly better than the performance of the K-means method. In addition, the feature criterion of the entropy-based Morchegan algorithm (mean = 88.5%) is significantly lower than the characteristic criterion of the logistic regression method (mean = 98.3%). In other words, the distance-based Morchegan algorithm is significantly more unsuccessful than the logistic regression method in identifying healthy companies.

9) The accuracy criterion of the entropy-based Morchegan algorithm (average = 36.8 percent) is not significantly different from the accuracy criterion of K-methods (average = 41.5 percent) and genetic algorithm (average = 45.3 percent). In other words, the performance of the entropy-based Morchegan algorithm regarding the accuracy of identifying companies as suspected of fraud is similar to the performance of the K-means and genetic algorithm methods.

10) The accuracy of the entropy-based Morchegan algorithm (average = 36.8%) is significantly lower than the accuracy criteria of logistic regression methods (average = 71%) and neural network algorithm (average = 58%). In other words, the entropy-based Morchegan algorithm performs significantly more unsuccessfully than logistic regression and neural network algorithms in correctly identifying companies as suspected of fraud. Table

(25) summarizes the results of the first hypothesis test regarding the performance of the entropy-based Morchegan algorithm.

Based on the results obtained from comparing the performance of the entropy-based Morchegan algorithm with other methods, the research hypothesis regarding the performance of the entropy-based Morchegan algorithm cannot be confirmed.

The outputs of this project include scientific documents, data-driven methods for identifying financial fraud, providing a model for implementing the anti-fraud system in the bank, and finally, intercepting software technologies and implementing the entire research process.

Reviewing data and commenting on them follow two general methods, which are also true in most financial anti-fraud systems. First, the use of indicators using linear and statistical methods, and second, the use of data mining algorithms, which are divided into supervised and unsupervised, and both methods have been used in this project.

Another point is that the use of different methods can be done based on different tools, in the end, each tool implements the desired method using a specific algorithm, and this variety has a different effect on the results and implementation method. The most important difference between these algorithms is their speed and efficiency. In addition, according to the amount and type of input data in terms of numbers or text or other features, the choice of algorithm can change. Accordingly, in this project, for example, to implement the clustering method with the aim of identifying non-conventional data, the four methods of EM, Corresponding Group, Cummins and Cohen were used. Except for the EM method that was implemented with the Vaka software, the rest of the methods were implemented by the well-known data mining software, that is, Clementine. The purpose of the results of the implementation of this model is to provide useful reports for the bank inspection unit in the form of numerical and graphical reports. According to the nature of the data mining process, all the stages of the implementation of the model were implemented in accordance with the Crisp model, and its implementation documents have been prepared.

The results of this research are consistent with the research of Alia and et al (2023) ,Muhammad and Schachler (2022) , Lin and colleagues (2022) and Osenberg et al. (2018).

Table 25 Summary of the results of the research hypothesis test regarding the performance of the Morchegan algorithm based on entropy.

Similar Methods	Comparison of Methods	Performance Criteria
Genetic Algorithm	K Mean < Morchegan Algorithm < Logistic Regression, Neural Network	Accuracy
Genetic Algorithm	K Mean > Morchegan Algorithm > Logistic Regression, Neural Network	Error rate
Neural Network Regression Genetic Algorithm	Morchegan Algorithm < K Average	Sensitivity
Neural Network Genetic Algorithm	K Mean < Morchegan Algorithm < Logistic Regression	Specificity
K Average Genetic Algorithm	Morchegan Algorithm <Logistic Regression, Neural Network	Correctness

Practical Suggestions for Research

1- By amending the commercial law, embedding control and legally binding tools, considering appropriate punitive measures and raising the cost of fraud, provide the basis for reducing fraudulent financial reporting.

2- The duties of each of the regulatory bodies in the process of combating fraud should be precisely defined and parallel work should be avoided by making the necessary arrangements.

3- The judicial authority and supervisory and professional bodies such as the Securities and Exchange Organization should pay more attention to the category of fraudulent financial reporting and by forming a specialized board to deal with crimes related to financial reporting, the field of prevention, detection, follow-up and dealing with this type Provide fraud.

4-Economic enterprises should prevent any kind of fraud, including fraud in financial reporting, by establishing a corporate governance system and strengthening internal controls.

5-Defining new goals and duties with the aim of preventing and fighting against fraudulent financial reporting in the field of auditing and financial reporting of the Securities and Exchange Organization can be useful. In addition to this, the announcement of the list of fraudulent companies and the introduction of used fraud schemes will make auditors, inspectors and researchers more familiar with the fraud process in financial reporting.

6- The use of data mining methods and the provision of fraud detection software, due to the large volume of data, facilitate and accelerate the process of handling and discovering fraudulent financial reporting cases by auditors, inspectors, the stock exchange organization and regulatory bodies. Faster detection of fraud cases prevents further damage to the victims.

7- Auditors, inspectors and investors should pay more attention to return on equity, operating cash to sales, return on assets, and debt collection period ratio in analyzing financial statements.

8- The judicial authority and regulatory and professional bodies such as the Securities and Exchange Organization should pay more attention to the category of fraudulent financial reporting and by forming a specialized board to deal with crimes related to financial reporting, the field of prevention, detection, follow-up and dealing with this type of fraud provide

9-Economic enterprises should prevent any kind of fraud, including fraud in financial reporting, by establishing a corporate governance system and strengthening internal controls.

10-Definition of new goals and duties with the aim of preventing and fighting against fraudulent financial reporting in the area of auditing and financial reporting of the Securities and Exchange Organization can be useful. In addition, the announcement of the list of fraudulent companies and the introduction of used fraud schemes will make auditors, inspectors and researchers more familiar with the fraud process in financial reporting.

11- The use of data mining methods and the provision of fraud detection software, due to the large volume of data, facilitates and accelerates the process of handling and discovering fraudulent financial reporting cases by auditors, inspectors, stock exchange organization and regulatory bodies. Faster detection of fraud cases prevents further damage to the victims.

Research limitations

Due to the fact that in qualitative researches, the desired phenomenon is studied in the context where it occurs, therefore, the possibility of generalizing the results and findings of the research to other conditions and situations is limited. Also, in qualitative research, it is possible for the researcher's presuppositions and prejudices to appear and interfere, which may damage the findings and results of the research. Another limitation of qualitative research is that there is a possibility of different interpretations of the investigated phenomenon (Mashaikhi et al., 2012).

In addition, the list of fraudulent companies in the country is not announced, therefore, in the current research, based on the findings of the foundational theorizing section, the risk of fraud in financial reporting is defined indirectly and by referring to the report of the auditor and legal inspector.Due to the fact that there is no standard tool for the implementation of genetic algorithm, distance-based ant algorithm and entropy-based ant algorithm, the required tools were designed and programmed in this research. The performance of the prepared tool depends on the researcher's creativity in designing the relevant algorithm and skill in programming.

References

1) Ahmadi, Seyyed Jalal, Faghani Makrani, Khosro, & Fadeli, Naghi. (2019). Presenting a model for predicting fraudulent financial statements and comparing financial statements and ratios with Benford's law. *Accounting Knowledge and Management Audit*, Volume 9, No. 35, pp. 221-237.
2) Aliott, W. S., C. C. Albrecht and C. O. Albrecht. (2023). Fraud and Corporate Executives: Agency, Stewardship and Broken Trust. *The Journal of Forensic Accounting*, Vol. 5, pp. 109-.031.
3) Canapinta, S., Gerundin, C., and Manolopoulos, Y. (2021). Data mining techniques for the detection of

fraudulent financial statements. *Expert Systems with Applications*, Vol.32, pp. 995-1003.

4) Ciefert C. S. (2019). Qualified audit opinions and share prices: an investigation, Auditing. *Journal of Practice and Theory*, Vol. 1, pp. 35-.35.
5) Clirces. M. Boumediene, Salem Lotfi (2018). Detection and prediction of managerial fraud in the financial statements of Tunisian banks. *Accounting & Taxation*, Vol. 6, No. 2, 2014, pp. 1-10
6) Jozef, N., & Welz, S. M. (2015). Financial statement fraud risk mechanisms and strategies: the case studies of Malaysian commercial companies, Procedia - *Social and Behavioral Sciences*, NO. 145, 321 – 329.
7) Khoshbakht, Ismail, Taibi Naqandari, Amir Hossein, & Zinli, Hadith. (2022). Religiosity, accountants' professional ethics and financial reporting frauds. *Financial Accounting and Auditing Researches*, *14* (53), pp. 223-254.
8) Malki Kakler, Hassan, Bahri Tahali, Jamal, Jabarzadeh Kangarloui, Saeed, Ashtab, Ali, Kazemi, Tawheed, (2023). Presenting the developed pentagonal model of fraud in financial reporting with an emphasis on the structure of internal controls. *Audit Knowledge Quarterly*, No. 84, pp. 85-102.
9) Mohammad. M. Gupta, R., & Singh Gill, N. (2023). Prevention and detection of financial statement fraud: An implementation of data mining framework. *International Journal of Advanced Computer Science and Applications*, Vol. 3, No.8.
10) Osunberg. C., Helmut Krcmar (2018). Reducing false positives in fraud detection: Combining the red flag approach with process mining. *International Journal of Accounting Information Systems*, Volume 31, 1-16.
11) Rahimian, Nizamuddin, & Haji Heydari, Razieh. (2018). Fraud detection using modified Benish model and financial ratios. *Accountant's Empirical Researche*s, volume 9, number 1, pp. 47-70.
12) Rasuli Parshkoh S., Shahri M., & Haji Hashemi Verno Safadarani M. (2022). The relationship between audit quality, stability and financial objective with the possibility of fraud in financial statements. *Specialized Scientific Quarterly of New Research Approaches in Management and Accounting*, Volume 6, Number 20, pp. 151-175.
13) Ravisankar, P., Ravi, V., Raghava Rao, G., Bose, I., (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support System*s. No. 50, pp. 491–500.
14) Rezaee, Z., & Raieei. N, (2019). Financial Statement Fraud, Prevention and Detection, John Wiley and Sons Inc, Chapter 3, pp.44-68.
15) Sanousoo. N., Josh (2019). Using Machine Learning to Detect Financial Fraud. Elizabethtown College JayScholar. 1-48.
16) Shah Moradi, Nasim, & Tabatabaei Nesab, Zahra. (2022). Investigating the impact of audit quality on the relationship between economic uncertainty and profit management caused by accruals in companies listed on the Tehran Stock Exchange. *Financial Accounting Research*, Volume 13, Number 1, pp. 67-86.
17) Tajik Nia, Elham, Yaqub Nejad, Ahmad, Jahanshad, Azita, & Kiqbadi, Amir Reza. (2022). The cost of not choosing profit management of the follower companies in response to the fraudulent financial reporting of the leading company. *Management Accounting*, Volume 15, Number 52. pp. 93-110.