Data mining approach in detecting inaccurate financial statements in government-owned enterprises

Amra Gadžo[1,*], Mirza Suljić[2], Adisa Jusufović[1], Slađana Filipović[1] and Erna Suljić[3]

[1] *Faculty of Economics, University of Tuzla, Urfeta Vejzagića 8, 75000 Tuzla, BiH*
E-mail: ⟨{amra.gadzo, adisa.jusufovic, sladjana.filipovic}@untz.ba⟩

[2] *Center for Quality Assurance and Internal Evaluation, University of Tuzla, Armije RBiH bb, 75000 Tuzla, BiH*
E-mail: ⟨mirza.suljic@untz.ba⟩

[3] *Public Elementary School "Simin Han", Sarajac 4, 75207 Tuzla, BiH*
E-mail: ⟨erna.cerimagic@hotmail.com⟩

Abstract. The study aims to assess the capability of various data mining techniques in detecting inaccurate financial statements of government-owned enterprises operating in the Federation of Bosnia and Herzegovina (FBiH). Inaccurate financial statements indicate potential financial fraud. Prediction models of four classification algorithms (J48, KNN, MLP, and BayesNet) were examined using a dataset comprising 200 audited financial statements from government-owned enterprises under the supervision of the Audit Office of the Institutions in the Federation of Bosnia and Herzegovina. The results obtained through data mining analysis reveal that a dataset encompassing seven balance sheet items provides the most comprehensive depiction of financial statement quality. These seven attributes are: opening entry of accounts receivable, profit (loss) at the end of the period, operating assets at the end of the period, accounts receivable at the end of the period, opening entry of operating assets, short term financial investments at the end of the period, and opening entry of short-term financial investments. By employing these seven attributes, the MLP algorithm was implemented to construct the most precise predictive model, achieving a 76% accurate classification rate for financial statements. Leveraging the identified attributes, a mathematical model could potentially be formulated to effectively predict financial statements of government-owned enterprises in FBiH. This, in turn, could considerably facilitate the process of selecting GOEs for inclusion in the annual work plan of state auditors. Presently, due to resource constraints, government-owned enterprises in FBiH do not undergo regular annual scrutiny by state auditors, with only 10 to 15 such enterprises being subject to audits each year. The results of this research can also be beneficial to both the public and the Financial Intelligence Agency in the FBiH. The paper contributes to filling the gap in the literature regarding the applied methodology, particularly in the part concerning the attributes used in the research.

Keywords: data mining, financial statement frauds, government-owned enterprises, prediction of financial statements accuracy

Received: April 20, 2024; accepted: October 4, 2024; available online: February 4, 2025

DOI: 10.17535/crorr.2025.0001

Original scientific paper.

---

*Corresponding author.

1. Introduction

Transitional countries, including Bosnia and Herzegovina, often face problems such as an inefficient public sector, a high corruption index, and the employment of politically affiliated personnel in the management structure and supervisory boards of government-owned enterprises. The resources of these enterprises are often used to support political campaigns and favor certain suppliers in the procurement of goods and services, under the guise of the Public Procurement Law. The challenges faced by transitional countries like Bosnia and Herzegovina, including an inefficient public sector, high corruption index, and politically affiliated personnel in government-owned enterprises, have been widely documented [3]. These issues have eroded public trust in state institutions and led to financial embezzlement in government-owned enterprises. Anti-corruption policies and good governance practices have been implemented to address these challenges, but their effectiveness in rebuilding trust in public institutions remains a subject of debate [3]. The restructuring of state-owned enterprises, with a focus on corporate governance and the role of the state, has been proposed as a potential solution. However, the complex relationship between tax evasion, state capacity, and trust in transitional countries, as well as the prevalence of public procurement corruption in Bosnia and Herzegovina, further complicate the situation. Government-owned enterprises that are owned by the state or lower levels of government are frequently accused of financial embezzlement and do not enjoy public trust in the accuracy of their financial statements. Financial statement fraud involves the intentional concealment or omission of critical information resulting from a deliberate failure to report financial data in line with generally accepted accounting principles. Financial fraud is a serious problem worldwide and is particularly pronounced in companies that are state-owned in countries such as Bosnia and Herzegovina. Lalić, Jovičić & Bošnjaković [20] highlights the inflow of money from abroad and the presentation of operating losses as common examples of financial fraud in the region. Buljubašić Musanović & Halilbegović [2] further underscores the manipulation of financial statements in failing small and medium-sized enterprises, with significant differences in accruals, asset quality, leverage, profitability, and liquidity between failing and non-failing SMEs. Isaković-Kaplan et al. [13] explores the application of Benford's Law in detecting potential earnings manipulation in income statements of economic entities, emphasizing the need for additional forensic investigations. Yadiati, Rezwiandhari & Ramdany [35] highlight a broader perspective by identifying factors such as financial stability, external pressure, industry nature, director changes, and collaboration with government projects as possible indicators of fraudulent financial reporting in state-owned enterprises. Another issue also arises from the fact that financial statements of government-owned enterprises are not subject to annual audits by state auditors. Due to the extensive number of institutions under state audit supervision (more than 2,000 institutions), only a few of government-owned enterprises are incorporated into the annual audit plan. The Audit Office of the Institutions in the Federation of Bosnia and Herzegovina does not utilize modern data mining tools to aid in the detection of inaccurate financial statements. In the process of planning which government-owned enterprises will be included in the audit plan for the year, state auditors are guided by media reports, anonymous complaints, and internal information from previous financial audits. The fundamental research question in this study is which balance sheet items and data mining technique provide the best probability of predicting inaccurate financial statements in government-owned enterprises? Therefore, the objectives of this paper are to identify the balance sheet items that best indicate inaccurate financial statements and to determine which data mining technique achieves the best results in predicting inaccuracies in financial statements for government-owned enterprises. Data mining techniques will be employed to detect balance sheet items from financial position reports of enterprises that offer the most accurate predictions of the state auditors' assessments of financial statement quality. Furthermore, various data mining techniques will be utilized to evaluate their predictive efficacy through comparative analysis of forecasting results. This paper aims to contribute to addressing the literature gap concerning the attributes used to identify balance sheet items that best predict inaccuracies in financial statements of government-owned enterprises in FBiH. These objectives will be achieved through a systematic literature review, identifying gaps in the literature. Following this, a detailed explanation of the research methodology and the application of data mining techniques will follow. Finally, the research findings will be presented, accompanied by their explanations.

2. Literature review

A range of studies have explored the use of various models and techniques to predict inaccurate financial statements based on audit opinion. Gadžo et al. [5] found that the Beneish M-score model, particularly its partial indicators, can accurately predict the quality of financial statements in public enterprises. Similarly, Sánchez-Serrano et al. [27] developed a model for predicting audit opinion in consolidated financial statements, achieving high accuracy. Wu & Li [33] further improved on this by using a BP neural network with Adam optimizer to predict audit opinions in listed companies, with a high accuracy rate. Yue, Shen & Chu [37] also identified specific financial ratios, such as net assets per share and earnings per share, as strong indicators of false financial affairs, on the basis of the model of logistic regression analysis. Research consistently shows that financial ratios are a key factor in predicting an auditor's qualified opinion on financial statements [6]. These ratios, such as retained earnings to total assets, equity to total liabilities, and net income to total assets, are used in various models to accurately classify qualified and unqualified opinions. Rudkhani & Jabbari [6] found that only two financial ratios, "earnings per share" and "fixed asset turnover," were needed for an accuracy rate of 64.1% . Similarly, Gadžo [5] achieved a high accuracy rate of 98-100% using eight partial indicators from the Beneish M-score model. Sormunen [28] further noted that the classification ability of certain financial ratios may diminish over time, suggesting the need for ongoing evaluation. In our literature review, we did not identify studies that predict inaccurate financial statements based on actual values of balance sheet items at the beginning and end of the period (without using ratio indicators). That is the research gap that this scientific paper aims to fill. The basis for determining inaccurate financial statements in government-owned enterprises is the analysis of auditors' findings and criticisms, as well as the written grounds for issuing opinions on the financial statements. According to our own research [5], the most common causes of irregularities in financial reporting of government-owned enterprises in FBiH include inadequate accounting estimates of accounts receivable from customers, inadequate valuation of fixed assets, inventory, and provisions. A range of methodologies have been proposed for predicting the auditor's opinion on financial statements. Sánchez-Serrano et al. [27] and Stanišić, Radojević & Stanić [30] both highlight the use of artificial neural networks and machine learning algorithms. Data mining can be an extremely useful tool for detecting irregularities within a large volume of data in financial statements. Large amounts of data often contain hidden patterns and trends that may indicate potential fraudulent activities. Therefore, data mining is used to extract knowledge from vast data sets in order to identify behavioral patterns that may indicate fraud. The study conducted by evaluated the capability of various Data Mining classification methods to identify companies that released fraudulent financial statements (FFS), with a particular emphasis on recognizing the key factors linked to such fraud. The study explored the application of Decision Trees, Artificial Neural Networks (ANN), and Bayesian Belief Networks as tools for detecting fraudulent financial reporting. The results demonstrated that the Bayesian Belief Network model achieved the best performance, successfully classifying 90.3% of the validation sample in a 10-fold cross-validation procedure. According to the research conducted by [26], utilized data mining techniques, including Multilayer Feed Forward Neural Network (MLFF), Support Vector Machines (SVM), Genetic Programming (GP), Group Method of Data Handling (GMDH), Logistic Regression (LR), and Probabilistic Neural Network (PNN), to identify companies involved in financial statement fraud. These techniques were tested on a dataset of 202 Chinese companies and compared with and without feature selection. Without feature selection, PNN outperformed all other techniques, while with feature selection, both GP and PNN achieved nearly equal accuracy, outperforming the others. According to the research conducted by [21] on financial statement fraud, utilizing data mining methods including logistic regression, decision trees (CART- C4.5. algoritam), and artificial neural networks (ANN), the results indicate that artificial neural networks and decision trees resulted in much more accurate classification compared to logistic regression. According to the research conducted by [29], the methods used in the fraud detection process were Linear Regression, Artificial Neural Networks (ANN), k-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Stump, M5P Tree, Random Forest, and J48. The results from the experiments indicated that data mining methods were able to detect the fraud factors between the financial statements and the e-ledger. In this study, the Decision Stump Algorithm exhibited the best performance. In addition to the mentioned studies, numerous other research studies have focused on the effectiveness of different data mining techniques in detecting fraud in financial statements [17]. Tatusch et al. [31] introduced a modified version of DBSCAN, a density-based clustering algorithm, which outperformed prior methods in detecting restated financial statements. Ravisankar et al. [26] and Gill & Gupta [7] both found that probabilistic neural network (PNN) and neural network techniques were effective in identifying companies resorting to financial statement fraud. These studies collectively suggest that a combination of clustering and classification techniques, particularly those that incorporate temporal variation and financial ratios, can be effective in predicting inaccurate financial statements. All of these authors have investigated fraud detection in the profit sector. However, We were unable to identify significant scientific studies on detecting inaccurate financial statements in government-owned enterprises. Numerous authors have demonstrated the effectiveness of utilizing the Beneish M-Score model in practice, although they did not employ data mining techniques but rather relied solely on the mathematical formula of the model [10].

3. Research elaboration

3.1. Methodology

This paper examines the data extracted from the financial statements of Government-Owned Enterprises (GOEs) in the FBiH and the corresponding audit reports prepared by the Audit Office of the Institutions in FBiH. The study encompasses a time span of 16 years, from 2004 to 2019. It comprises a total of 200 financial statements and their associated audit reports, all of which were available at the time of conducting the study. Financial statements serve as the fundamental basis and starting point for analyzing business operations and assessing the condition of a company. They can also be regarded as confidential records generated by organizations that contain their financial transactions, including expenses, realized profits, income from loans, etc. [8]. Financial statements provide an indication of the organization's financial reality and also include management notes on business performance and projected future trends. Furthermore, inaccurate financial statements deceive users of financial reports by creating the impression that organizations are performing favorably. The task of auditing is to protect the interests of capital owners and provide a reliable information foundation for rational decision-making and management of state-owned companies. In accordance with the International Standard on Auditing (ISA) 240 (sections 2 and 3), misstatements in financial statements can occur due to either fraud or error (the International Standards of Supreme Audit Institutions-ISSAI does not specifically define inaccurate financial statements like ISA 240 does). The key differentiating factor between fraud and error lies in whether the underlying action that leads to the misstatement in financial statements is deliberate or unintentional.

While fraud is a broad legal concept, auditors, under the ISAs, focus on fraud that causes significant misstatements in financial statements. There are two types of intentional misstatements that are relevant to auditors: misstatements arising from fraudulent financial reporting and misstatements resulting from misappropriation of assets. Although auditors may suspect or, in rare cases, identify instances of fraud, they do not make legal determinations regarding the occurrence of fraud. Consequently, the term "misstatement" will be used going forward without explicitly specifying whether an error is intentional or not . Pursuant to International Standard on Auditing (ISA) 240 (sections 2 and 3), misstatements in the financial statements can arise from either fraud or error (International Standards of Supreme Audit Institutions-ISSAI does not specifically define inaccurate financial statements like ISA 240 does). The key difference between fraud and error lies in the intent behind the action that leads to inaccuracies in the financial statements—whether it is intentional or unintentional. Although the concept of fraud is broad in legal terms, under the International Standards on Auditing, the auditor focuses on frauds that cause material misstatements in the financial statements. There are two types of intentional misstatements relevant to the auditor: those arising from fraudulent financial reporting and those resulting from the misappropriation of assets [11]. Although the auditor may suspect or, in rare cases, identify the occurrence of fraud, the auditor does not make legal determinations of whether fraud has actually occurred. Therefore, the term misstatement shall be used onwards without specifying if error is intentional or not.

International Standard on Auditing (ISA) 705 (paragraph 2) defines three types of modified opinions: a qualified opinion, an adverse opinion, and a disclaimer of opinion. The decision regarding which type of modified opinion is appropriate depends on the nature of the matter causing the modification, i.e., whether the financial statements are materially misstated or, in cases where sufficient and appropriate audit evidence cannot be obtained, whether they could be materially misstated; and the auditor's judgment regarding the pervasiveness of the effects or possible effects of the issue on the financial statements [12]. Our model uses balance sheet items from financial statements as predictive attributes and the type of opinion on the financial statements as the target variable. As the input set of data for the model, we used 24 balance sheet positions: Opening and closing balances of accounts receivable (attribute code: A3, A1), Sales income for the current and previous accounting period (A4, A2), Operating expenses of the current and previous accounting periods (A6, A5), Opening and closing balances of operating assets (A11, A7), Opening and closing balances of Property, Plant, and Equipment (A12, A8), Opening and closing balances of short-term financial investments (A13, A9), Opening and closing balances of business assets (A14, A10), Opening and closing balances of Depreciation (A15, A16), Administrative expenses of the current and previous accounting period (A17, A18),

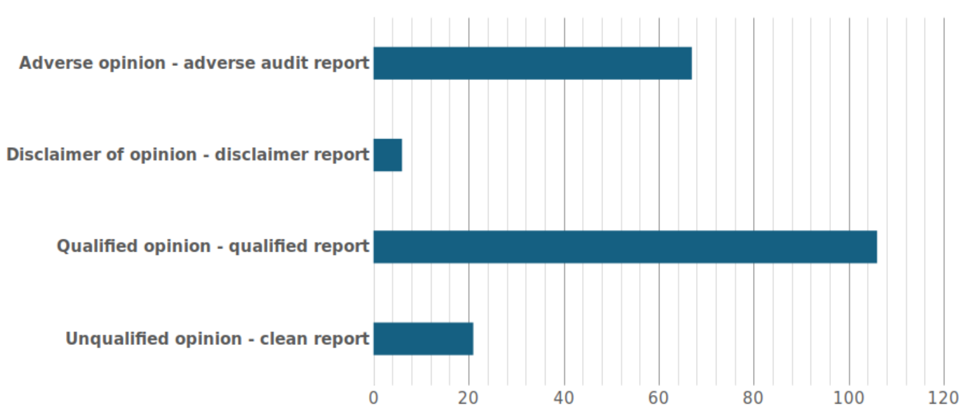

Figure 1: *Audit reports distribution for GOEs in FBiH.*

Opening and closing balances of short-term liabilities (A21, A19), Opening and closing balances of long-term liabilities (A22, A20), Business profit/loss for the current accounting period (A23) and Net cash flow from operating activities for the current accounting period (A24). These balance sheet items were selected based on the fact that auditors have predominantly identified irregularities in the valuation of these balance sheet positions as grounds for issuing qualified opinions. Auditors currently highlight non-compliance with the provisions of IFRS 9, IFRS 15, IAS 2, IAS 16, IAS 36, IAS 37, IAS 38, as well as cash flows from operating activities and business results in the financial statements of government-owned enterprises in FBiH. The output variable was the quality of financial statements measured relative to the audit report of state auditors. The results of the final audit reports for GOEs subject to the analysis are given in Figure 1. Number of enterprises is presented on the x-axis.

The output variable – audit report for GOEs in FBiH can be grouped as follows:

- As four categories or classes, so that audit report is a class, as presented in Table 1,
- As two classes coded as YES category – unqualified and qualified opinion, and NO category – adverse opinion and disclaimer of opinion, as presented in Table 2.

Class	Audit report	Sample number	Percentage
1	Unqualified opinion	21	10.50%
2	Qualified opinion	106	53.00%
3	Disclaimer of opinion	6	3.00%
4	Adverse opinion	67	33.50%
SUM		**200**	**100.00%**

Table 1: *Four classes, according to the final audit report.*

Class	Assessment	Sample number	Percentage
1	No	73	36.50%
2	Yes	127	63.50%
SUM		**200**	**100.00%**

Table 2: *Two classes, according to the final audit report.*

It is evident that prediction error in the first case would be much higher due to different distribution of the final audit report by classes. Hence, this research gave advantage to the second case. The enterprises are divided into two basic groups:

- The first group includes the enterprises whose financial statements were given unqualified or qualified opinion (127 enterprises).
- The second group includes the enterprises whose financial statements were given adverse of disclaimer of opinion (73 enterprises)

Such formulation of the output variable categorizes the problem as classification problem, where the aim of the model is to learn how to recognize the proper classification of the final audit report. The primary goal of prediction is to develop a model that derives insights about a specific characteristic of the dependent variable by utilizing a combination of independent variables. Predictive modeling involves determining the output variable for a constrained dataset, where the symbols represent the values of the output variable in particular instances. The choice of variables from the available dataset significantly influences the precision and accuracy of the resulting predictive models.

3.2. Data mining

Data mining, a field of knowledge discovery in databases [1], can be utilized for discovering financial frauds. Different types of data mining techniques can be employed for this purpose. Classification is a method used to evaluate and find a function that assigns items from a data set to predetermined classes of the output variable, based on the input variables' values [18]. Through classification, models can be created to classify unknown datasets into specific categories or classes [18]. The classification process typically involves the following steps:

- Selecting classifiers for implementing the classification algorithm.
- Choosing the class attribute (output variable).
- Dividing data sets into two: training data and test data.
- Training the classifiers on the training data set with known values of the class attribute.
- Testing the classifiers on the test data set with hidden values of the class attribute.

In the classification process, existing techniques are applied to evaluate the proposed prediction model using collected instances. In the case of selecting a classifier for datasets obtained from the financial statements of government-owned enterprises, the situation is quite clear. This dataset is very small; the procedure of conducting the experimentation phase is more difficult due to the fact that the data is dynamically changeable. Financial statements data are mostly of a numerical and categorical type, and since they are manually extracted from databases of financial statements, they require less cleaning in the preprocessing phase. Some applications of different classifiers on such data are described in the papers mentioned in the listed references. In the paper [15], the authors employed DT model for financial fraud detection in companies. Jan [14] used decision trees in combination with other data mining techniques to achieve a high accuracy rate of 90.83% in detecting financial statements fraud. Kirkos, Spathis & Manolopoulos [15] also explored the effectiveness of decision trees in detecting fraudulent financial statements, comparing their performance with other data mining techniques. These studies collectively highlight the potential of decision trees in detecting fraud in financial statements of government-owned enterprises. However, some recent approaches have been introduced for financial fraud detection [22]. Kırda & Özçelik [16] found KNN to be a highly effective classifier, achieving an accuracy rate of 91.73% in detecting financial statement fraud. Yao et. al. [36] also highlighted the effectiveness of KNN, particularly when combined with support vector machine (SVM) and stepwise regression, in detecting fraudulent financial statements. These studies collectively underscore the potential of KNN in detecting fraud in government-owned enterprises . A range of studies have demonstrated the effectiveness of Multilayer Perceptron (MLP) in detecting fraud in financial statements. Trigueiros & Sam [32] and Mubarek & Adali [24] both found that MLP, when used in conjunction with other machine learning techniques, outperformed traditional methods in fraud detection. Ravisankar et al. [26] and Kwon & Feroz [19] further support these findings, with Ravisankar [26] noting the superior performance of MLP in identifying companies engaged in financial statement fraud, and Kwon [19] reporting an 88% accuracy rate in predicting SEC investigation targets using MLP. These studies collectively highlight the potential of MLP in detecting fraud in financial statements, particularly in government-owned enterprises. For example, some studies have employed Bayesian networks to detect inaccurate financial statements. Deng [4] found the algorithm to be effective in this context, noting its potential for proactive fraud detection. Handoko, Wiyardi & Handoko [9] also found success in using the Beneish M-Score method, which includes variables related to financial statement manipulation, in detecting fraud in Indonesian government-owned enterprises. In this work, and in accordance with the above mention, the following algorithms were chosen: Decision Trees (J48) [14, 15], K-nearest neighbors [16, 36], Neural Network (MLP)

[8, 19, 24, 26, 32], and Bayesian network (Naive Bayes) [4, 9] for application on the training dataset. The following subsection offers an detailed overview of the machine learning techniques applied in this research to detect financial fraudulent activities. Decision Trees (DT) are one of the most well-known classification techniques often used to model data in the form of a tree structure. These algorithms establish relationships between input features and outputs using a tree-like structure, named for its resemblance to an inverted tree. The name "decision tree" derives from the fact that it resembles an inverted tree. The most commonly used and widely recognized decision tree algorithm is C4.5, and its implementation in the Weka software tool is known as the J48 algorithm. The J48 (C4.5) algorithm [25] presents an extension of Professor Ross Quinlan's earlier ID3 algorithm, and is known for its exceptionally high accuracy. The advantage of the J48 algorithm is the ability to work with numerical and categorized data [34], in addition, it is easy to implement and effectively deals with noise and missing values [23], and it has the ability to display results graphically. Basic construction of J48 algorithm uses a method known as divide and conquer to display output [34, 23]:

- Select the dataset as input for the rule-making process.
- Calculate the normalized information gain for each attribute.
- Choose the attribute with the maximum information gain as the best attribute. This attribute becomes the root node and corresponding to the best predictor.
- Repeat the above step until a stopping criterion is met, calculating the information gain for each attribute and adding that attribute as a child node.

Just as a tree starts from the root, branches into individual branches, and ends in leaves, decision trees use branches to represent decision paths, with the final outcome represented by the leaves. The final result is a tree with decision nodes and leaves. Each decision node has two or more divisions, while each leaf represents an outcome or decision.

K-nearest neighbors (KNN) algorithm represents a straightforward and easily understandable supervised learning method commonly applied in both classification and regression problems . It belongs to a group of algorithms known as instance-based learning, sometimes referred to as 'lazy learning methods' because the processing of the training data is delayed until a test instance needs to be classified. KNN operates on the principle that objects with similar characteristics tend to belong to the same class or have similar output values. The main goal of KNN is to group n objects into k groups (classes) based on their attributes or features. When a testing example is considered, it is placed in an n-dimensional metric space of attribute values. The algorithm then determines the distance between the testing example and all training examples in this space. For classification, the most popular classification among the k nearest training examples is the target estimation for the classifier. To define 'nearest,' various metrics can be used, including the standard Euclidean distance and Hamming distance, with Euclidean distance being commonly used. If k is greater than 1, those closer to the testing example will have a greater weight in the classification. The procedure starts with an initial division of the set items into a selected number of groups. The distance between every object and every group is determined [22], and objects are located into groups closest to them based on given characteristics. After joining an object to a group, the centroid of the group is recalculated. The distance of every object from the group centroid is recalculated, and the distribution of objects among groups continues until the selected function of the criterion suggests otherwise. Because of all the above, the KNN algorithm is easy to implement and understand, and is more suitable for small data sets. However, for large datasets, it can be computationally demanding because it requires calculating the distances to all instances in the training set for each new instance.

Multilayer Perceptron (MLP) algorithm is one of the most frequently used and known neural network. The network consists of a set of perceptrons that make an input layer, one or more hidden layers of process elements and an output layer [34]. MLP is particularly suitable for the approximation of classification functions (when we have little knowledge on the ratio between input and output attributes) that map the example determined by the attribute value vector into one or more classes.

Multilayer perceptrons (MLPs) utilizes a feed-forward structure where signals flow in one direction, from the input layer to the output layer without any feedback loops, providing a nonlinear input-output mapping. The basic nonlinear mappings and summation layers are determined by the structure of the MLP. Layers that are not directly connected to the environment are called hidden layers. In this research, we chose an MLP with one hidden layer, using a hyperbolic tangent activation function $\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}$ for the hidden neurons, and a logistic sigmoid function as the output activation function. However, any other activation function, such as a threshold function, could also be used. The MLP model can be described as follows:

$$f(x,w) = g\left(\sum_{j=0}^{M} w_j^{(2)} \tanh\left(w_{ji}^{(1)} x_i\right)\right), \tag{1}$$

where w contains the weights and the bias parameter and i and j stand for input and hidden units, respectively. Since the logistic function's output value can be interpreted as a probability and ranges from 0 to 1, it was selected as the output activation function in this case.

Indeed, the Multilayer Perceptron (MLP) algorithm has shown significant potential for use in the financial sector, particularly in detecting Financial Statement Fraud [8]. MLP's ability to handle complex patterns and nonlinear relationships in data makes it well-suited for identifying irregularities and anomalies in financial statements. By analyzing historical financial data and patterns, MLP can learn to detect suspicious activities and behaviors that may indicate fraudulent practices.

Naive Bayesian (NB) algorithm is widely used in machine learning technique for classification and has shown effectiveness in various domains. Naive Bayes is a simple classification method that predicts membership probabilities for each class based on Bayes' theorem [4, 34]. It operates under the assumption that the attributes are statistically independent of each other. Naive Bayes classifier uses equation to classify the value of the target variable y based on the assumption that the attribute variables, $x_1, \ldots, x_p$, are independent of each other:

$$y_{MAP} \approx \mathrm{argmax}_{y \in Y} p(y) \prod_{i=1}^{p} P(x_i|y). \tag{2}$$

It calculates the probability of a data instance belonging to a particular class by combining prior probabilities and conditional probabilities. By using these probabilities, NB can make predictions and classify new data instances into the most probable class [34]. The term "naïve" refers to its simplification of the problem by relying on two key assumptions: it assumes that prognostic attributes are conditionally independent given the known classification, and it assumes there are no hidden attributes that could influence the prediction process. This approach is a promising method for probabilistic knowledge detection and enables an efficient algorithm for financial fraud detection.

To assess the robustness of classifiers [1], we used triple cross validation, that is we randomly divided a set of data into three subsets of equal size. Two subsets were used for training and one was used for cross checking the validity. This procedure was executed three times so that each subset was tested once. The test results stand for the given average value of each of the cross checking. For the implementation of the selected classifiers, we used the Weka 3.8.5 data mining tool.

4. Results and discussions

In the previous stages of developing the mathematical model for detecting fraudulent financial statements in the government-owned enterprises of FBiH, it was essential to identify key variables and indicators that would enhance predictive capability. After collecting and extracting the data, we gained insights into their structure and informative value to prepare for the implementation of data mining algorithms. To achieve this, we utilized simple statistical techniques and visualization tools. Histograms were created to visualize the distribution of nominal attribute values, numerical attribute distributions were examined, and graphical presentations were generated to analyze attribute values relative to the class or other attributes. These techniques helped in identifying unexpected data patterns, indefinite attribute values, duplicate data, and other anomalies. In order to gain more detailed insight into the importance of input variables, the values of input attributes relative to the output attribute were assessed, along with analyzing the influence of individual input variables on the model relative to the output variable. The aim of this assessment and attribute selection was to extract irrelevant and redundant attributes from the training data set. Data filtering included techniques to assess attribute values based on heuristics derived from general data features. Filtering methods proved to be a more practical solution for intelligent data analysis, as they significantly shortened attribute selection and assessment processes. Their independence from machine learning algorithms allowed for their implementation alongside any data modeling technique. The filtering methods used were InfoGain and GainRatio with the Ranker search method. InfoGain assesses the value of an attribute by measuring its information gain relative to class. GainRatio assesses the value of an attribute by measuring relative information gain to the class. The attributes assessed less than 0.01 need to be excluded from the analyzed data set. The results of the assessment and ranking the attributes based on their individual values are given in Table 3.

Table 3 shows that out of 24 balance sheet positions, seven balance sheet positions best describe misstatement of financial statements of GOEs in FBiH. They are: opening entry of accounts receivable (A3), profit (loss) at the end of the period (A23), operating assets at the end of the period (A7), accounts receivable at the end of the period (A1), opening entry of operating assets (A11), short term financial investments at the end of the period (A9), and opening entry of short-term financial investments (A13).

If we observe the audit bases used for giving opinions on financial statements, it can be concluded that the results largely match. Namely, as the auditors found, the most frequent error in the financial statements is that public procurement procedures are not in line with the Law on Public Procurements. This directly affects financial statements through booking procurement of fixed assets, services or consumable goods. This objection given by the auditors is evident in 42 out of 200 enterprises. The error registered in 35 enterprises refers to accounts receivable not being value adjusted. What follows is the error identified in 34 enterprises that do not measure reimbursable value of fixed assets. Stocks not being properly showed is an error placed fourth, registered in 26 companies, same as the error of reserves not being properly shown.

Table 3 indicates that the attributes A03, A23, A07, A01 and A09 have the highest ranking while A11 and A13 have the lowest ranking. Since the remaining attributes have the values below 0.01, we will have the fourth set of input data to test their predictive capability. This fourth set of input data includes only the attributes (seven balance sheet positions with informative value).

The algorithms Decision Trees (J48), KNN, Neural network (MLP) and Naive Bayes for creating prediction model were used in this research (Table 4). The model estimate was made by a triple cross validation method.

Table 4, in section B, shows that for the set of data with seven attributes and informative value above 0.01, using the MLP algorithm, the most accurate prediction model was created.

All atributes			**Balance sheet positions**
ATTRIBUTES	**InfoGain**	**GainRatio**	**Description**
A03	0.1665	0.1305	Accounts receivable t-1
A23	0.1050	0.1758	Business profit/loss t
A07	0.0892	0.1271	Operating assets t
A01	0.0649	0.1165	Accounts receivable t
A11	0.0639	0.0885	Operating assets t-1
A09	0.0589	0.1404	Short-term financial investments t
A13	0.0516	0.0678	Short-term financial investments t-1
A02	0	0	Sales income t
A04	0	0	Sales income t-1
A05	0	0	Operating expenses t-1
A06	0	0	Operating expenses t
A08	0	0	Property, plant and equipment t
A10	0	0	Business assets t
A12	0	0	Property, plant and equipment t-1
A14	0	0	Business assets t-1
A15	0	0	Depreciation t-1
A16	0	0	Depreciation t
A17	0	0	Administrative expenses t
A18	0	0	Administrative expenses t-1
A19	0	0	Short-term liabilities t
A20	0	0	Long-term liabilities t
A21	0	0	Short-term liabilities t-1
A22	0	0	Long-term liabilities t-1
A24	0	0	Net cash flow from operating activities t

Table 3: *Results of evaluating and ranking attributes.*

The MLP algorithm generated the model with 76% correctly classified items (CCI), precision of 75.6% (0.756), and classification above the segment of the ROC curve ($0.754 > 0.5$). Predictive capability for the first set of data A, made of 24 balance sheet positions is between 57.5% and 71%, depending on the implementation of different classification algorithms.

Ultimately, the results indicate that the MLP algorithm provides the most accurate model for the set of 7 attributes (A3, A23, A7, A1, A11, A9 and A13) with high precision and classification above the ROC curve segment. This suggests that MLP was an efficient choice for data analysis, considering the selected attributes and learning methodology.

It is important to emphasize that the implementation of the MLP neural network and the validation by means of dividing the data set in such a way that two thirds are used for the creation of the model and one third for its validation, result in 4% on average better prediction indicators for all sets of input data. This leads to the conclusion that a larger set of input data might prove better results of predictive capability, using the same methodology. A small data set is the main limitation of this research.

Limitations of this study relate to a relatively small research sample. A sample of 200 financial statements is not large enough for the applied data mining methodology. However, collecting a larger volume of historical data would not be relevant due to changes and adoption of new international accounting and financial reporting standards.

DATA SET	EVALUATION CRITERIA	CLASSIFIERS			
		J48	KNN	MLP	BayesNet
A	Timing to build model (in Sec)	0	0	0.41	0
	Correctly classified instances	142	141	139	115
	Incorrectly classified instances	58	59	61	85
	Prediction accuracy	71	70.5	69.5	57.5
	Kappa statistic	0.3316	0.397	0.3552	0.1264
	Mean absolute error (MAE)	0.3163	0.2916	0.3403	0.4175
	Root mean squared error (RMSE)	0.5021	0.4626	0.4445	0.4887
	Relative absolute error (RAE)	68.16%	62.82%	73.33%	89.96%
	Root relative squared error (RRSE)	104.29%	96.07%	92.32%	101.51%
B	Timing to build model (in Sec)	0	0	0.08	0
	Correctly classified instances	129	124	152	115
	Incorrectly classified instances	71	76	48	85
	Prediction accuracy	64.5	62	76	57.5
	Kappa statistic	0.1635	0.2501	0.4536	0.1264
	Mean absolute error (MAE)	0.4051	0.3685	0.3584	0.4175
	Root mean squared error (RMSE)	0.4997	0.5313	0.4348	0.4887
	Relative absolute error (RAE)	87.29%	79.39%	77.23%	89.96%
	Root relative squared error (RRSE)	103.77%	110.34%	90.30%	101.51%

Table 4: *The performances obtained by the previous classification algorithms using the 24 selected attributes (A) and the seven selected attributes (B).*

5. Conclusions

Aiming to provide a specific solution to the problem of detecting misstatements in financial statements of GOEs in FBiH whose majority capital is mainly owned by the federal, cantonal and local levels of authority, we implemented four different algorithms of data mining (J48, KNN, MLP and BayesNet) on four sets od input data. Using filtering methods InfoGain and GainRatio with the Ranker search method, we established that seven balance sheet positions have the largest significance for future development of a mathematical model of prediction of misstatements in GOEs' financial statements. They are: opening entry of accounts receivable, profit (loss) at the end of the period, operating assets at the end of the period, accounts receivable at the end of the period, opening entry of operating assets, short term financial investments at the end of the period, and opening entry of short-term financial investments. These seven balance sheet positions have as many as 76% of predictive capability, i.e., correct classification of financial statements, with the estimation of the model by a triple cross validation method. If the validation is set in such a way that two thirds of data are used for the creation of the model and one third for its validation, we get even better indicators of predictive value. This points to the fact that a larger data set might prove higher predictive capability. The relatively small sample of 200 financial statements from government-owned enterprises represents a limitation of the conducted research.

This research also contributes to filling the literature gap regarding the attributes used to detect inaccurate financial statements in government-owned enterprises. Unlike other similar studies, this work uses balance sheet positions from financial position reports of enterprises instead of financial ratios as attributes.

Future development of mathematical model for predicting misstatements of financial statements made by GOEs in FBiH should focus on seven identified balance sheet positions. This can serve as the recommendation for future research in this field. The benefits of development and application of this model by the public, Financial-intelligence Agency of FBiH (FIA) and the Audit Office of the Institutions in FBiH can be identified as follows:

- at the point of financial statement submission, the state agency would be able to detect GOEs with high probability of misstatements in their financial statements;
- the Audit Office of the Institutions in FBiH would spend less resources as they would have a focused action strategy and audit those GOEs with high probability of misstatements in their financial statements;
- the community would be more efficient in detecting misstatements in fiancial statements and consequently spend less resources and have a better control of the process.

References

[1] Banjanović-Mehmedović, L. (2011). *Inteligentni sistemi*. Tuzla: Harfograf Tuzla. url: https://www.knjiga.ba/inteligentni-sistemi-k6991.html [Accessed 29/6/2024]

[2] Buljubasic Musanovic, E. and Halilbegovic, S. (2021). Financial statement manipulation in failing small and medium-sized enterprises in Bosnia and Herzegovina. Journal of Eastern European and Central Asian Research (JEECAR), 8(4), 556-569. doi: 10.15549/jeecar.v8i4.692

[3] Chandler, D. (2006). Building trust in public institutions? Good governance and anti-corruption in Bosnia-Herzegovina. Ethnopolitics, 5(1), 85-99. doi: 10.1080/17449050600558272

[4] Deng, Q. (2010). Detection of fraudulent financial statements based on Naive Bayes classifier. In: *2010 5th International Conference on Computer Science & Education*. doi: 10.1109/ICCSE.2010.5593407

[5] Gadžo, A., Halilbegović, S., Ðaković, A. O and Hodžić, A. (2022). Predicting Auditor's Opinion on Financial Statements of Public Enterprises Based on Indicators of the Beneish M-score Model. Journal of Forensic Accounting Profession, 2(2), 1-13. doi: 10.2478/jfap-2022-0006

[6] Ghale Rudkhani, T. M. and Jabbari, H. (2014). The effect of financial ratios on auditor opinion in the companies listed on TSE. European Online Journal of Natural and Social Sciences: Proceedings, 2(3(s)), 1363-1373. url: https://european-science.com/eojnss_proc/article/view/3826 [Accessed 29/6/2024]

[7] Gill, N. S. and Gupta, R. (2012). Analysis of Data Mining Techniques for Detection of Financial Statement Fraud. The IUP Journal of Systems Management, X(1), 7-15. url: https://ssrn.com/abstract=2152771 [Accessed 29/6/2024]

[8] Hajek, P. and Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud–A comparative study of machine learning methods. Knowledge-Based Systems, 128, 139-152. doi: 10.1016/j.knosys.2017.05.001

[9] Handoko, B. L., Wiyardi, E. S. and Handoko, J. S. (2022). The Application of Beneish M-Score Method in Detecting Fraudulent Manipulation on Financial Statements (Case Study of Indonesian Government-State Owned Enterprise (SOE) Registered on Indonesian Stock Exchange 2016-2020. In: *Proceedings of the 2022 13th International Conference on E-business, Management and Economics*. doi: 10.1145/3556089.3556106

[10] Herawati, N. (2015). Application of Beneish M-Score models and data mining to detect financial fraud. Procedia-Social and Behavioral Sciences, 211, 924-930. doi: 10.1016/j.sbspro.2015.11.122

[11] The International Auditing and Assurance Standards Board (2004). The Auditor's Responsibility to Consider Fraud and Error in an Audit of Financial Statements. url: https://www.iaasb.org/consultations-projects/auditors-responsibility-consider-fraud-audit-financial-statements [Accessed 29/6/2024]

[12] The International Auditing and Assurance Standards Board (2015). Forming an Opinion and Reporting on Financial Statements. url: https://www.iaasb.org/publications/international-standard-auditing-isa-705-revised-modifications-opinion-independent-auditor-s-report-3 [Accessed 29/6/2024]

[13] Isaković-Kaplan, Š., Demirović, L. and Proho, M. (2021). Benford's law in forensic analysis of income statements of economic entities in Bosnia and Herzegovina. Croatian Economic Survey, 23(1), 31-61. doi: 10.15179/ces.23.1.2

[14] Jan, C. (2018). An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan. Sustainability, 10(2), 513. doi: 10.3390/su10020513

[15] Kirkos, E., Spathis, C. and Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. Expert systems with applications, 32(4), 995-1003. doi: 10.1016/j.eswa.2006.02.016

[16] Kırda, K. and Özçelik, M. K. (2021). Finansal Tablo Hilesi Riski Taşiyan Şirketlerin Veri Madenciliği Ile Belirlenmesi. Journal of Accounting and Taxation Studies, 14(2), 609-639. doi: 10.29067/muvu.802703

[17] Kopun, D. (2018). A Review of the Research on Data Mining Techniques in the Detection of Fraud in Financial Statements. Journal of Accounting and Management, VIII(1), 1-18. url: https://hrcak.srce.hr/207781 [Accessed 29/6/2024]

[18] Kumar, R. and Verma, R. (2012). Classification algorithms for data mining: A survey. International Journal of Innovations in Engineering and Technology (IJIET), 1(2), 7-14. url: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=cf8621c96c73220ebf0c02dc53795b38da3f8376 [Accessed 29/6/2024]

[19] Kwon, T. M. and Feroz, E. H. (1996). A multilayered perceptron approach to prediction of the SEC's investigation targets. IEEE transactions on neural networks, 7(5), 1286-1290. doi: 10.1109/72.536321

[20] Lalić, S., Jovičić, Ž. and Bošnjaković, T. (2021). The most common examples of financial fraud in Bosnia and Herzegovina: A practical insight. Journal of Forensic Accounting Profession, 1(2), 80-88. doi: 10.2478/jfap-2021-0010

[21] Lin, C.-C., Chiu, A.-A., Huang, S. Y. and Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. Knowledge-Based Systems, 89, 459-470. doi: 10.1016/j.knosys.2015.08.011

[22] Malini, N. and Pushpa, M. (2017). Analysis on credit card fraud identification techniques based on KNN and outlier detection. In: *2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB)*. doi: 10.1109/AEEICB.2017.7972424

[23] Mesarić, J. and Šebalj, D. (2016). Decision trees for predicting the academic success of students. Croatian Operational Research Review, 7(2), 367-388. doi: 10.17535/crorr.2016.0025

[24] Mubarek, A. M. and Adalı, E. (2017). Multilayer perceptron neural network technique for fraud detection. In: *2017 International Conference on Computer Science and Engineering (UBMK)*. doi: 10.1109/UBMK.2017.8093417

[25] Quinlan, J. R. (2014). *C4. 5: Programs for Machine Learning*. San Mateo: Elsevier. url: https://archive.org/details/c45programsforma0000quin [Accessed 29/6/2024]

[26] Ravisankar, P., Ravi, V., Rao, G. R. and Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. Decision support systems, 50(2), 491-500. doi: 10.1016/j.dss.2010.11.006

[27] Sánchez-Serrano, J. R., Alaminos, D., García-Lagos, F. and Callejón-Gil, A. M. (2020). Predicting audit opinion in consolidated financial statements with artificial neural networks. Mathematics, 8(8), 1288. doi: 10.3390/math8081288

[28] Sormunen, N. (2009). The Auditorś Going Concern Decision and Alternative Financial Ratios. url: https://osuva.uwasa.fi/handle/10024/5486

[29] Sorkun, M. C. and Toraman, T. (2017). Fraud detection on financial statements using data mining techniques. Intelligent Systems and Applications in Engineering, 5(3), 132-134. doi: 10.18201/ijisae.2017531428

[30] Stanisic, N., Radojevic, T. and Stanic, N. (2019). Predicting the type of auditor opinion: Statistics, machine learning, or a combination of the two? The European Journal of Applied Economics, 16(2), 1-58. doi: 10.5937/ejae16-21832

[31] Tatusch, M., Klassen, G., Bravidor, M. and Conrad, S. (2020). Predicting erroneous financial statements using a density-based clustering approach. In: *2020 The 4th International Conference on Business and Information Management*. doi: 10.1145/3418653.3418673

[32] Trigueiros, D. and Sam, C. (2016). Software to Streamline the Detection of Fraud in Published Financial Reports of Companies. SSRN Electronic Journal, 95-106. doi: 10.2139/ssrn.4501531

[33] Wu, H.-P. and Li, L. (2021). The BP neural network with adam optimizer for predicting audit opinions of listed companies. IAENG International Journal of Computer Science, 48(2). url: https://www.iaeng.org/IJCS/issues_v48/issue_2/IJCS_48_2_16.pdf [Accessed 29/6/2024]

[34] Witten, I. H. and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edition.* San Francisco: Morgan Kaufmann Publishers. url: http://academia.dk/BiologiskAntropologi/Epidemiologi/DataMining/Witten_and_Frank_DataMining_Weka_2nd_Ed_2005.pdf [Accessed 29/6/2024]

[35] Yadiati, W. and Rezwiandhari, A. (2023). Detecting Fraudulent Financial Reporting In State-Owned Company: Hexagon Theory Approach. JAK (Jurnal Akuntansi) Kajian Ilmiah Akuntansi, 10(1), 128-147. doi: 10.30656/jak.v10i1.5676

[36] Yao, J., Pan, Y., Yang, S., Chen, Y. and Li, Y. (2019). Detecting fraudulent financial statements for the sustainable development of the socio-economy in China: a multi-analytic approach. Sustainability, 11(6), 1579. doi: 10.3390/su11061579

[37] Yue, D., Wu, X., Shen, N. and Chu, C.-H. (2009). Logistic regression for detecting fraudulent financial statement of listed companies in China. In: *2009 International Conference on Artificial Intelligence and Computational Intelligence.* doi: 10.1109/AICI.2009.421