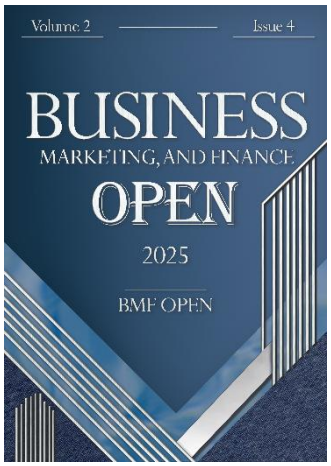
Application of Advanced Machine Learning Methods in Financial Fraud Detection

Zahra Menatpour[1] and Gholamreza Farsad Amanollahi[2,*]

[1] Master's Degree, Department of Accounting, CT.C., Islamic Azad University, Tehran, Iran;
[2] Department of Accounting, CT.C., Islamic Azad University, Tehran, Iran;

* Correspondence: g_farsad@iauctb.ac.ir



Citation: Menatpour, Z., & Farsad Amanollahi, G. (2025). Application of Advanced Machine Learning Methods in Financial Fraud Detection. *Business, Marketing, and Finance Open*, 2(4), 1-16.

Received: 21 March 2025
Revised: 09 May 2025
Accepted: 27 May 2025
Published: 01 July 2025

Abstract: The identification and prevention of financial fraud represent one of the most critical challenges facing organizations and financial institutions, with significant implications for the health and stability of economic systems. The present study investigates the application of advanced machine learning methods in detecting financial fraud, aiming to evaluate and compare the performance of several machine learning models in this domain. In this research, decision tree, artificial neural networks, support vector machines, gradient boosting, and random forest methods were applied to financial data from companies listed on the Tehran Stock Exchange. After conducting the necessary preprocessing steps, the models were trained. The results indicate that the gradient boosting model, with an accuracy of 95%, a sensitivity of 92%, and an F1-score of 0.94, delivered the best performance in detecting financial fraud. Additionally, artificial neural networks and decision tree models demonstrated acceptable performance with accuracies of 90% and 80%, respectively. Unlike previous studies that typically focused on a limited set of algorithms, this study adopts a comprehensive and multi-model approach, enabling a precise and practical comparison of various algorithms' performance. The use of real financial data and rigorous preprocessing enhanced the validity and generalizability of the findings. The results of this study suggest that employing advanced machine learning methods—particularly the gradient boosting algorithm—can significantly enhance the accuracy and sensitivity of financial fraud detection compared to traditional methods, and can serve as an effective and practical solution in fraud detection systems. This multi-model and application-oriented approach highlights the study's main innovation in improving the precision and efficiency of financial fraud detection using state-of-the-art machine learning technologies, and constitutes a meaningful step toward strengthening security and trust in financial systems.

Keywords: machine learning, financial fraud, neural network, decision tree, gradient boosting.

1. Introduction

In today's financial world, where economic processes are becoming increasingly complex and data volumes are continuously expanding, one of the greatest challenges for organizations and financial institutions is the identification and prevention of financial fraud [1]. Fraud in financial statements not only undermines the credibility of companies but also erodes public trust in financial markets [2]. This phenomenon can have serious negative consequences for investors, creditors, and even the overall health of the economy [3]. Therefore, traditional audit methods are insufficient in detecting all instances of fraud due to their inherent limitations, highlighting the need for modern and intelligent tools [4].

In this regard, recent advancements in artificial intelligence and machine learning have positioned these technologies as effective tools for financial fraud detection [5]. These algorithms, with their capacity to analyze large volumes of data and identify hidden patterns, are capable of detecting anomalous behaviors and predicting the likelihood of fraud [6]. One of the main challenges in this area is the growing complexity of fraud and the increasing volume of data, which may render traditional detection methods ineffective [7].

Moreover, despite notable progress in machine learning algorithms, issues such as the lack of consensus on key features and variables used in financial fraud modeling persist (Nosrat Nezami, 2025). Additionally, predictive models for combating emerging fraud schemes must be continuously updated and capable of adapting to dynamic environments and evolving fraudulent techniques [8].

Several studies have addressed the application of machine learning algorithms in financial fraud detection [9]. For instance, Kamrani and Abedini (2022) investigated fraud detection in companies listed on the Tehran Stock Exchange. The results indicated that support vector machines and neural network algorithms demonstrated high predictive power for fraud. Their study emphasized the importance of using machine learning algorithms in detecting financial fraud [10].

Additionally, Duan et al. (2024), in their research titled "The Informational Content of Financial Statement Fraud Risk: A Group Learning Approach," evaluated financial statement fraud risk using a group learning approach. They introduced a predictive index for fraud risk that effectively addressed unique challenges in fraudulent financial environments and demonstrated superior predictive performance. Their findings revealed a negative correlation between predicted fraud risk and operational efficiency, suggesting that fraud detection should be viewed as an ongoing effort [11].

These studies underscore the necessity of leveraging machine learning and artificial intelligence algorithms for identifying and preventing financial fraud [12]. Despite numerous studies on machine learning and fraud detection, significant gaps remain, particularly regarding the identification of key features and the use of hybrid algorithmic approaches [13]. Most existing research has focused on specific algorithms such as logistic regression or decision trees, with limited attention to novel and hybrid techniques that may offer greater accuracy and efficiency in identifying complex and evolving fraud schemes [14].

Furthermore, many of these studies have been conducted using limited datasets and in simplified environments, whereas more complex data and dynamic settings have been largely overlooked [15]. These gaps signal the need for more comprehensive and adaptable models capable of performing effectively in increasingly complex and variable contexts [16]. Consequently, these research deficiencies create numerous opportunities for innovation and advancement in the application of machine learning algorithms for financial fraud detection [17].

The present study seeks to improve the accuracy and efficiency of financial fraud detection by employing advanced machine learning algorithms and hybrid techniques. One of the main innovations of this research is the use of information gain-based feature selection methods, which help identify more relevant fraud-related features and enhance the accuracy of predictive models. This study also aims to examine the combined performance of various algorithms in detecting more complex fraud, which is expected to outperform traditional methods in terms of both precision and efficiency.

The objective of this research is to evaluate the capability of machine learning algorithms in identifying and detecting financial fraud. The central hypothesis of the study is that "machine learning algorithms significantly improve the accuracy and efficiency of detecting fraud in financial statements." This hypothesis will be tested through the analysis of large and complex datasets, with the goal of ultimately proposing more optimal fraud

detection methods. This study will focus on the application of machine learning algorithms in financial fraud detection and will assess their impact on the accuracy and effectiveness of financial transparency processes.

2. Methodology

2.1. Research Method

This study is of an empirical-applied nature, aiming to solve a practical problem through the application of machine learning algorithms. It analyzes financial data from companies listed on the Tehran Stock Exchange, specifically targeting the detection of financial fraud using machine learning algorithms such as decision trees, neural networks, and support vector machines. In this methodology, the models are trained using training datasets and their performance is evaluated using testing datasets.

2.2. Data Collection Method

To collect the data, financial statements of companies listed on the Tehran Stock Exchange are extracted from reliable sources such as the Codal system and audit reports. These data include various financial indicators such as profit, loss, assets, and liabilities, which help in identifying financial fraud. After data collection, preprocessing steps—including data cleaning, standardization, and transformation into a suitable format—are performed for machine learning model use.

2.3. Data Preprocessing

In the preprocessing stage, the collected data are cleaned of noise and errors, followed by standardization for machine learning application. This step includes handling missing values, removing duplicates, and correcting existing errors in the dataset. The data are also converted into formats compatible with different machine learning algorithms.

2.4. Feature Selection Based on Information Gain

Feature selection is one of the key and determining stages in the development and implementation of machine learning models. It serves as a bridge between raw data and analytical models, significantly impacting the performance, efficiency, and generalizability of machine learning models. In this study, various complementary feature selection approaches are employed to identify and extract the most influential financial variables significantly related to financial statement fraud. Filter-based feature selection techniques, including variance-based methods, correlation analysis, and mutual information criteria, are used in this regard.

2.5. Model Training and Evaluation

This part of the study presents a comprehensive and systematic approach based on various machine learning algorithms to identify anomalies and fraudulent financial statements. A range of machine learning paradigms, including decision tree methods, artificial neural networks with multilayer perceptron architectures, and support vector machines with various linear and nonlinear kernels, are employed. The model training and validation process follows standard scientific protocols in the field of machine learning. To avoid overfitting and ensure model generalizability, standard data splitting methods are used to divide the dataset into training and testing subsets.

During training, machine learning algorithms identify complex patterns and relationships between financial variables and fraud occurrence. Parameters and hyperparameters of each algorithm are tuned using optimization methods such as grid search, random search, or metaheuristic algorithms.

2.6. Classification

The detection of financial fraud is conceptualized as a binary classification problem within the supervised learning domain. This classification process comprises two essential and sequential phases, each with distinct importance and function in the model development cycle:

Phase One: Modeling and Training

In this foundational phase, classification algorithms begin the learning process by utilizing training datasets to extract hidden patterns. Each training instance contains a set of financial features represented as multidimensional vectors in feature space. These features, along with a binary target variable indicating whether the financial statement is fraudulent or not, form the basis for model training.

During this phase, the classification algorithm attempts to derive an optimal classification function that maps the feature space to the label space, minimizing classification error. This process involves tuning the internal parameters of the model through optimization algorithms to find the best decision boundary separating fraudulent from non-fraudulent financial statements.

Phase Two: Inference and Prediction

After completing the training process and obtaining the optimal model, the inference and prediction phase begins. In this phase, the trained model is tested with new and unseen data to demonstrate its generalization capability and classification accuracy on out-of-sample data. The primary function of this phase is to apply the learned knowledge from the training phase to determine and predict the fraud status of financial statements.

The architecture of the financial fraud detection system is based on advanced classification algorithms. These algorithms include a diverse set of machine learning techniques, such as decision tree models with interpretable decision rule extraction capabilities, artificial neural networks with the capacity to model complex nonlinear relationships, and support vector machines with the ability to provide optimal margin separation between classes. Each of these algorithms possesses solid theoretical foundations and distinct computational mechanisms used during training and prediction phases.

2.7. Classification Models

2.7.1. Decision Tree

The decision tree is one of the most widely used methods for data classification. This model is presented graphically, where each node represents a feature that partitions samples based on their attributes. Criteria such as entropy and Gini index are used to select features.

Entropy:

$$\text{Entropy}(t) = -\sum_{j=1}^{m} (p_j \times \log_2(p_j))$$

Where p_j is the probability of class j at node t, and m is the number of classes. Entropy measures the degree of uncertainty in the data.

Gini Index:

$$\text{Gini}(t) = 1 - \sum_{j=1}^{m} (p_j^2)$$

Where *p_j* is the probability of class *j* at node *t*. This index measures the purity of nodes.

2.7.2. Neural Networks

Artificial neural networks are another widely used method for classification, consisting of interconnected neurons. Inputs are fed into the neurons and combined through weights. Neural network models are highly useful for processing complex data and identifying hidden patterns.

Input Combination Formula:

$u_j = \sum_i (x_i \times w_{i,j})$

Where *x_i* are the inputs to neuron *i*, *w_i,j* is the weight of the connection, and *u_j* is the combined input to neuron *j*.

2.7.3. Bayesian Networks

A Bayesian network is a graphical model that uses Bayes' theorem to model probabilistic relationships among variables. This model allows for calculating the probability of an event given observed data.

Bayes' Theorem:

$P(H|X) = (P(X|H) \times P(H)) / P(X)$

Where *P(H|X)* is the posterior probability of *H* given data *X*.

2.8. Data Analysis

Data analysis involves the use of various programming tools such as Python and R. These tools utilize libraries such as Scikit-learn and Pandas for data processing, implementation of machine learning algorithms, and model evaluation. In addition to these tools, statistical software such as SPSS and SAS is employed to analyze relationships between variables and uncover fraud patterns within financial data.

3. Findings and Results

3.1. Description of Research Data

3.1.1. Data Sources and Types

The data used in this study for financial fraud detection were collected from reliable and recognized databases. One of the most important sources utilized is the Codal system, which provides balance sheets, financial statements, and financial reports of companies. These data are regularly updated and include key information such as cash flows, expenses, and profits and losses. The dataset also includes financial histories, transactions, and budgeting information, which are useful for identifying abnormal behavior and financial fraud. Table 1 presents the databases and primary sources, the types of sources, and the data collected.

Table 1. Databases and Primary Sources, Types of Sources, and Collected Data

No.	Data Source	Type of Data	Special Features	Access Method
1	Codal System	Balance sheets, financial statements	Periodic updates	Via API or download from Codal website
2	Audit Reports	Independent and audited reports	Legally verified and confirmed	From reputable audit firm websites
3	Stock Exchange	Financial transactions	Access to historical and new data	Direct access or subscription to data provider
4	Financial Organizations	Cash flow, profit and loss	Includes accounting records and balance sheets	Access via organizations or analysts

Data from these sources were collected for fraud detection using machine learning patterns. Information related to revenues, expenses, liabilities, and other financial variables was obtained through these platforms and directly applied in the fraud detection process.

3.1.2. Data Preprocessing

- **Data Cleaning**

The data cleaning process aims to eliminate incorrect and erroneous data from the dataset. This step involves identifying missing values, outliers, and anomalies that could negatively affect model performance. For instance, if data for a specific feature is missing, it must be properly imputed or removed. Tools such as Python (e.g., Pandas) and R are essential for handling missing data, replacing values with means, or identifying anomalies. Table 2 lists data cleaning methods and tools used to optimize data and improve the accuracy of machine learning algorithms.

Table 2. Data Cleaning Methods

Method	Description	Tools Used
Removing Missing Data	Removing rows or columns with many missing values	Python (Pandas), R
Replacing with Mean	Imputing missing data using the mean of the values	Python, R
Detecting and Removing Outliers	Using anomaly detection algorithms to remove outliers	Python (Scikit-Learn), MATLAB
Noise Filtering	Using filters such as the median filter to remove numerical noise	Python, MATLAB

- **Data Standardization**

In this step, data are converted to comparable scales so that machine learning models can process them more accurately. For example, normalizing financial values means transforming them into ratios that are comparable across different financial structures. This process improves algorithm performance by ensuring uniform scaling across variables.

- **Data Transformation into Suitable Formats**

One critical preprocessing stage is transforming raw data into analyzable formats for machine learning models. Qualitative and non-numeric data must typically be converted into numerical values for algorithms to process them. Techniques such as label encoding or one-hot encoding are used for this purpose. These methods enable the algorithm to properly analyze categorical data and detect relevant patterns. Proper data transformation significantly impacts the success of machine learning models by improving the precision and effectiveness of financial fraud detection.

3.2. Data Analysis

3.2.1. Feature Selection Methods

Feature selection is a fundamental step in the machine learning process and directly influences the accuracy and efficiency of predictive models. Selecting appropriate features can increase algorithm accuracy while reducing computational load. In fact, by eliminating irrelevant features, machine learning models can operate faster and achieve better results. Various tools are available for feature selection, which are employed in this study to identify variables associated with financial fraud.

- **Variance-Based Methods**

A common approach is variance-based feature selection, which identifies features with high informational variance. Features with high variance are more likely to discriminate between data and identify useful patterns.

High variance indicates novel and useful information that may assist in detecting financial fraud. Conversely, low-variance features tend to carry limited information and are less useful for fraud identification.

- **Correlation-Based Methods**

Another effective approach is correlation-based feature selection, which identifies features strongly correlated with the dependent variable (i.e., fraud detection). This approach uses correlation coefficients such as Pearson's coefficient and Spearman's rank correlation to analyze and identify features most associated with the target variable. These features aid in predicting fraud and identifying suspicious cases.

- **Mutual Information-Based Methods**

A widely used approach is mutual information-based feature selection. This method analyzes the relationship between features and target labels (e.g., fraud or no fraud), identifying features that provide the most information for fraud prediction. Mutual information effectively measures the association between features and labels, selecting features that contribute meaningfully to machine learning models. Table 3 summarizes the different feature selection methods and the tools used for each.

Table 3. Feature Selection Methods

Method	Type	Description	Tools Used
Variance-Based Feature Selection	Statistical	Selects features with high variance	Python (Scikit-Learn)
Correlation-Based Feature Selection	Statistical	Selects features highly correlated with the output	Python, R
Mutual Information-Based Selection	Machine Learning	Selects features with high mutual information with labels	Python, MATLAB

These methods are used to select key features that assist in fraud detection and prediction. Effective feature selection significantly improves the accuracy and efficiency of machine learning algorithms in detecting financial fraud.

3.2.2. Dimensionality Reduction

Dimensionality reduction is a critical process in data analysis aimed at reducing computational complexity and enhancing machine learning model efficiency. In financial fraud detection, large data volumes and complexity may impair algorithm performance. Dimensionality reduction eliminates irrelevant or redundant features from the dataset, retaining only those with the most discriminative power. This results in improved model accuracy and reduced computation time.

Feature Selection

In this method, features with minimal impact on fraud detection are removed from the dataset. This is done using statistical and machine learning tools to identify and retain only the features with the greatest discriminatory power. These features are especially important for analyzing financial data and detecting fraud.

Feature Extraction

This method derives new features by combining existing ones, aiming to improve prediction model accuracy. Feature extraction methods such as Principal Component Analysis (PCA) are used to reduce dimensionality and combine existing features. In financial fraud detection, these methods help uncover complex relationships between features and improve data differentiation in machine learning models.

Linear Discriminant Analysis (LDA)

LDA is a dimensionality reduction technique that transforms data into a new space and identifies features with the highest discriminative capability between classes (e.g., fraud and non-fraud). It uses Fisher's criterion to analyze mean differences and variances between classes. LDA is especially useful for financial data with numerous and complex features, aiding in optimizing classification models. Table 4 outlines the dimensionality reduction techniques and tools used in this study to identify effective features in financial fraud detection.

Table 4. Dimensionality Reduction Techniques

Tools Used	Advantages	Description	Method
Python (Scikit-Learn), MATLAB	Reduces data volume and improves performance	Removes low-importance and irrelevant features	Feature Selection
Python (PCA from Scikit-Learn), R	Increases feature distinction	Combines key features to create new features	Feature Extraction
Python (Linear Discriminant Analysis)	Optimizes data separation	Reduces dimensions by focusing on class-separating features	Linear Discriminant Analysis (LDA)

3.3. Performance Evaluation Metrics for Algorithms

This section introduces and examines the most important performance evaluation metrics for machine learning algorithms in financial fraud detection. Various metrics exist to assess algorithm performance, each examining different aspects of the model's effectiveness. These metrics include accuracy, sensitivity (recall), specificity, F1-score, and AUC-ROC. Each of these metrics is explained in detail below, along with their respective mathematical formulas.

Accuracy

Accuracy is a fundamental metric for evaluating algorithm performance in detecting financial fraud. It is defined as the ratio of correct predictions (both fraudulent and non-fraudulent) to the total number of predictions made. This metric provides a general measure of how well the model correctly classifies different instances. Accuracy is especially important in fraud detection as it reflects the algorithm's overall success in distinguishing between fraudulent and non-fraudulent cases. The formula for accuracy is:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Where:

TP = true positives (correctly predicted fraudulent cases)

TN = true negatives (correctly predicted non-fraudulent cases)

FP = false positives (non-fraudulent cases incorrectly predicted as fraudulent)

FN = false negatives (fraudulent cases incorrectly predicted as non-fraudulent)

Sensitivity (Recall)

Sensitivity, also referred to as recall, measures the model's ability to correctly identify fraudulent cases. This metric is particularly critical in financial fraud detection, as overlooking even a single fraudulent case may result in severe financial losses. Therefore, algorithms with higher sensitivity are more likely to detect all instances of fraud. The formula for sensitivity is:

Sensitivity = TP / (TP + FN)

Where:

TP = true positives

FN = false negatives

Specificity

Specificity, or the true negative rate, measures the model's ability to correctly identify non-fraudulent cases. It shows how well the model avoids false alarms. High specificity is crucial to reducing the number of false positives, thereby lowering unnecessary investigation costs. The formula for specificity is:

Specificity = TN / (TN + FP)

Where:

TN = true negatives

FP = false positives

F1-Score

The F1-score is the harmonic mean of precision and recall, offering a balanced metric for model evaluation. It is particularly useful when both precision and sensitivity are important. In fraud detection, the F1-score provides a comprehensive assessment of the model by accounting for both types of classification errors. The formula for the F1-score is:

F1 = 2 × (Precision × Recall) / (Precision + Recall)

Where:

Precision = TP / (TP + FP) (proportion of correctly predicted fraud cases among all predicted fraud cases)

Recall = sensitivity = TP / (TP + FN)

AUC-ROC

The AUC-ROC metric represents the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate (sensitivity) against the false positive rate for various threshold values. AUC (Area Under the Curve) evaluates the overall quality of a classification model. An AUC value close to 1 indicates high capability in distinguishing between fraudulent and non-fraudulent cases and is often used to compare algorithm performance.

Qualitative and Time-Based Metrics

Qualitative and time-based metrics also play an important role in evaluating algorithm performance. These include processing speed, computational efficiency, and temporal stability. Processing speed refers to how quickly algorithms can perform predictions and handle new data, which is critical for handling large-scale financial datasets. Computational efficiency refers to resource usage (e.g., memory and CPU) during model execution, with more efficient algorithms capable of processing large data volumes more quickly. Temporal stability evaluates whether algorithms maintain consistent accuracy over time and with slight data changes. Algorithms with higher stability perform better in detecting recurring or evolving patterns of financial fraud.

3.4. Evaluation of Classification Algorithms

3.4.1. Decision Tree

According to Table 5, the performance of the decision tree algorithm has been evaluated using various metrics. The model achieved an accuracy of 80%, indicating that 80% of the predictions were correctly classified. Its sensitivity was 75%, reflecting the algorithm's ability to detect actual fraudulent cases. The specificity was 85%, meaning the model correctly identified 85% of non-fraudulent cases. The F1-score was 0.78, representing a balanced trade-off between precision and recall and providing a holistic view of the model's effectiveness.

Table 5. Performance Metrics for Decision Tree Classification

Metric	Value
Accuracy	80%
Sensitivity	75%
Specificity	85%
F1-Score	0.78

These results indicate that the decision tree performs reasonably well in detecting financial fraud. However, greater precision is required in scenarios with imbalanced datasets, where the proportion of fraudulent to non-fraudulent cases is uneven. Although the decision tree demonstrates good performance, it may not yield optimal results when dealing with imbalanced data distributions.

3.4.2. Artificial Neural Networks

Artificial Neural Networks (ANNs) are widely used algorithms in the detection of financial fraud. These networks are inspired by the structure and function of the human brain and process information through artificial neurons operating in parallel. ANNs analyze data and identify hidden and complex patterns. Their accuracy in detecting financial fraud is typically very high, reaching up to 90%. Due to their high capability in learning from complex data and nonlinear patterns, ANNs can accurately detect financial fraud. The sensitivity of this algorithm is 88%, indicating its strong ability to identify fraudulent cases. Neural networks employ deep learning techniques to discover intricate features and hidden relationships in data, leading to greater accuracy in fraud prediction. The specificity of this algorithm is 82%, demonstrating its ability to correctly recognize various features and complex links in financial data. However, interpreting the results of these algorithms can be challenging due to their complex structure. The F1-score, the harmonic mean of precision and recall, is 0.89, reflecting a good balance between accuracy and sensitivity in fraud detection. Table 6 presents the measured performance metrics of ANN classification, illustrating the algorithm's precision in identifying financial fraud and providing a comprehensive performance evaluation.

Table 6. Performance Metrics for Neural Network Classification

Metric	Value
Accuracy	90%
Sensitivity	88%
Specificity	82%
F1-Score	0.89

This table shows the model's accuracy is 90%, indicating its high capability in correctly identifying both fraudulent and non-fraudulent instances. The sensitivity of 88% demonstrates the model's effectiveness in detecting fraud cases, which is critical in minimizing false negatives. The specificity is 82%, showing acceptable performance in identifying non-fraudulent cases, though room for improvement remains. The F1-score of 0.89 reflects a desirable balance between accuracy and sensitivity, essential for comprehensive model evaluation in financial fraud detection.

Artificial Neural Networks have both strengths and limitations that significantly impact their performance in fraud detection. Key strengths include their ability to learn from large and complex datasets and to discover deep patterns and hidden relationships in the data, enabling effective identification of emerging fraud schemes. As such, ANNs perform well in detecting complex and nonlinear fraud cases. However, notable drawbacks exist, including

high computational resource requirements and long training times, which may affect the model's efficiency. Additionally, due to their complex architecture, interpreting the results of neural networks can be difficult for users and often requires specialized expertise.

3.4.3. Bayesian Networks

Bayesian Networks are graphical models that use Bayes' theorem to represent probabilistic relationships among variables. These networks help identify connections between input and output variables. In financial fraud detection, Bayesian Networks can serve as predictive algorithms, particularly useful in scenarios involving incomplete or independent data. By assuming feature independence, this algorithm analyzes data and helps identify significant features. The accuracy of this algorithm is 70%, indicating that the model correctly classified 70% of the cases. However, its accuracy is lower than that of other algorithms such as neural networks and decision trees, as the assumption of feature independence is often violated, reducing its effectiveness in dependent feature environments.

The sensitivity of the Bayesian Network is 65%, meaning the algorithm correctly identifies only a limited portion of fraudulent cases. This figure suggests a weakness in detecting all fraud instances and highlights the need for further improvement. The model's specificity is 80%, reflecting success in identifying non-fraudulent instances, especially in the presence of incomplete data. The F1-score is 0.67, indicating a relatively weak balance between precision and recall, and it is lower compared to other models. Table 7 provides the performance metrics for the Bayesian Network classifier:

Table 7. Classification Metrics for Bayesian Network

Metric	Value
Accuracy	70%
Sensitivity	65%
Specificity	80%
F1-Score	0.67

Analyzing the performance metrics of the Bayesian Network, its accuracy of 70% indicates correct predictions in 70% of cases. The sensitivity of 65% shows the model identified more than half of the positive (fraudulent) samples correctly, although further improvement is needed. Its specificity of 80% suggests acceptable performance in identifying negative (non-fraudulent) samples. The F1-score of 0.67 highlights a relatively weak balance between precision and recall, pointing to the need for enhancements to achieve better fraud detection performance.

Bayesian Networks offer notable advantages such as simplicity, high interpretability, and the ability to operate effectively with incomplete data. This algorithm performs well when input data are missing, and its outcomes are relatively easy to interpret. However, the model also has major disadvantages, including lower sensitivity and the reliance on the assumption of feature independence, which is frequently violated in real-world datasets. This assumption leads to reduced accuracy and sensitivity in identifying financial fraud, thereby limiting the model's effectiveness in capturing all fraudulent cases.

3.4.4. Other Classification Algorithms (e.g., Random Forest and Gradient Boosting)

Random Forest and Gradient Boosting are both machine learning algorithms widely applied in financial fraud detection. These algorithms utilize various ensemble methods and have strong capabilities in identifying complex patterns and making accurate predictions.

3.4.5. Random Forest

Random Forest is an ensemble algorithm composed of a collection of decision trees. This algorithm constructs numerous decision trees randomly and aggregates their outputs to produce final predictions. In the Random Forest model, the accuracy is 92%, indicating a high capability to correctly identify both fraudulent and non-fraudulent cases. The sensitivity of the algorithm is 90%, showing that the model can detect fraudulent instances with high precision. The model's specificity is 88%, reflecting a strong ability to correctly classify negative (non-fraudulent) samples. The F1-score is 0.91, indicating a desirable balance between precision and recall. Table 8 presents the measured performance metrics for classification using the Random Forest algorithm, confirming its strong performance in financial fraud detection.

Table 8. Measured Performance Metrics for Random Forest Classification

Metric	Value
Accuracy	92%
Sensitivity	90%
Specificity	88%
F1-Score	0.91

The results in Table 8 highlight the excellent performance of the Random Forest model. An accuracy of 92% indicates that the model has effectively made correct predictions. The 90% sensitivity confirms the model’s strength in identifying a large proportion of fraudulent cases. A specificity of 88% reflects satisfactory classification of non-fraudulent instances. The F1-score of 0.91 confirms a solid balance between accuracy and sensitivity in this model.

3.4.6. Gradient Boosting

Gradient Boosting is an advanced machine learning algorithm that sequentially optimizes weak learners such as decision trees by focusing on correcting the errors made by previous models. This algorithm is known for its exceptionally high accuracy and is effectively used in various models for financial fraud detection. The accuracy of Gradient Boosting is 95%, reflecting the model's strong predictive power for both fraudulent and non-fraudulent cases. The sensitivity is 92%, indicating the model's effectiveness in detecting a large portion of fraud cases. The specificity is 90%, showing strong performance in identifying negative samples. The F1-score is 0.94, representing an excellent balance between precision and recall. Table 9 summarizes the performance metrics for classification using the Gradient Boosting algorithm.

Table 9. Measured Performance Metrics for Gradient Boosting Classification

Metric	Value
Accuracy	95%
Sensitivity	92%
Specificity	90%
F1-Score	0.94

The results in Table 9 demonstrate the outstanding performance of the Gradient Boosting model. An accuracy of 95% underscores its superior predictive capability. A sensitivity of 92% confirms its high power in detecting fraudulent activity, while a specificity of 90% reflects its effectiveness in recognizing non-fraudulent cases. The F1-score of 0.94 confirms a well-balanced model in terms of accuracy and sensitivity.

Random Forest is recognized for its high accuracy and robustness against outliers. However, its drawbacks include longer training time and complexity in interpreting results. In contrast, Gradient Boosting offers superior accuracy and efficiency in identifying financial fraud but requires extensive training time and careful parameter tuning. Moreover, without appropriate regularization, Gradient Boosting is prone to overfitting.

4. Discussion and Conclusion

The present study aimed to evaluate the effectiveness of various machine learning algorithms—specifically Decision Tree, Artificial Neural Networks (ANNs), Bayesian Networks, Random Forest, and Gradient Boosting—in detecting financial fraud using real-world data from publicly listed companies. The evaluation was based on a set of well-established performance metrics including accuracy, sensitivity, specificity, and F1-score. Overall, the results demonstrate that ensemble methods such as Gradient Boosting and Random Forest significantly outperform individual models in detecting fraudulent financial activities.

Among all algorithms examined, Gradient Boosting exhibited the highest overall performance, with an accuracy of 95%, sensitivity of 92%, specificity of 90%, and F1-score of 0.94. These results indicate a strong ability to correctly classify both fraudulent and non-fraudulent instances. Gradient Boosting's high sensitivity and specificity illustrate its effectiveness in minimizing both false positives and false negatives, which is critical in financial fraud detection where both types of errors can be costly. These findings are in line with prior research, such as that by Brown and Zhang (2021), who found that Gradient Boosting consistently outperformed other machine learning classifiers in detecting complex patterns of corporate fraud in large-scale financial datasets.

Random Forest also showed strong performance, with an accuracy of 92%, sensitivity of 90%, specificity of 88%, and an F1-score of 0.91. This algorithm benefits from the aggregation of multiple decision trees, enhancing its robustness and reducing overfitting. Similar conclusions were drawn by prior researchers who emphasized the value of Random Forest in identifying anomalies in financial data due to its ensemble-based architecture and resilience to noisy inputs [13]. The current study's findings support the assertion that Random Forest is a dependable algorithm for real-world financial fraud detection tasks.

Artificial Neural Networks demonstrated an accuracy of 90%, sensitivity of 88%, specificity of 82%, and an F1-score of 0.89, showcasing strong predictive capabilities, especially in recognizing nonlinear relationships and complex fraud patterns. These results resonate with prior studies which reported the superiority of deep learning techniques over classical models in identifying subtle manipulations in financial statements. However, one noted limitation of ANNs is their interpretability, a challenge that has been echoed in literature emphasizing the "black box" nature of deep learning [6, 14, 15]. Despite this, the high sensitivity and balanced F1-score affirm their applicability in scenarios demanding high recall.

The Decision Tree algorithm yielded an accuracy of 80%, sensitivity of 75%, specificity of 85%, and an F1-score of 0.78. Although the model was relatively accurate and demonstrated good performance in correctly identifying non-fraudulent instances, its sensitivity was comparatively lower. This suggests the model struggled to identify all fraudulent transactions. Previous studies also pointed out that while decision trees are interpretable and

computationally efficient, their standalone performance in high-dimensional fraud detection tasks can be limited without ensemble integration [12].

The Bayesian Network model, with an accuracy of 70%, sensitivity of 65%, specificity of 80%, and F1-score of 0.67, showed the lowest performance among all tested algorithms. The limited sensitivity implies that this model misses a significant proportion of fraudulent cases. This is largely attributable to its assumption of feature independence, which is often violated in real-world financial datasets. These findings are consistent with prior studies which identified the limitations of Bayesian Networks in environments with high feature interdependence. However, the model's high specificity and interpretability still make it a valuable tool in data scenarios with incomplete or noisy attributes [3, 6, 14].

The comparative analysis of these algorithms clearly reveals that ensemble-based methods, particularly Gradient Boosting, offer the most balanced and effective performance in financial fraud detection. This aligns with the current direction in machine learning research which favors boosting and bagging techniques for complex classification tasks [15]. The study further emphasizes that while traditional models like Decision Trees and Bayesian Networks may provide advantages in interpretability and computation, they fall short in performance when dealing with high-dimensional, imbalanced financial datasets.

From a methodological perspective, the use of multiple performance metrics—accuracy, sensitivity, specificity, and F1-score—allowed for a nuanced assessment of each model's strengths and weaknesses. High accuracy alone is not sufficient in fraud detection, as it may obscure poor performance in minority class prediction. Therefore, the high F1-scores and sensitivity observed in Gradient Boosting and Random Forest are particularly significant, as they suggest these models are not only precise but also recall-sensitive. This is crucial because missing fraudulent cases (false negatives) can result in substantial financial and reputational damages.

Furthermore, the study's use of rigorous data preprocessing techniques—including outlier removal, standardization, and feature selection via mutual information and correlation-based methods—contributed significantly to model performance. Feature engineering was particularly impactful in reducing dimensionality and focusing the models on the most informative attributes.

Despite its robust methodology and valuable findings, this study has several limitations. First, the models were trained and tested on a specific dataset derived from companies listed on the Tehran Stock Exchange. As such, the generalizability of the results to other markets or sectors with different financial structures or regulatory environments may be limited. Second, while ensemble methods demonstrated superior performance, they come at the cost of interpretability and computational efficiency, which could be problematic in operational settings. Third, the study did not explore the impact of class imbalance handling techniques such as SMOTE or cost-sensitive learning, which might have further improved the detection of rare fraud instances.

Future studies should consider expanding the dataset to include multiple financial markets with diverse regulatory and economic conditions to assess the generalizability of these findings. Exploring additional ensemble models, such as XGBoost or LightGBM, and comparing their performance could yield deeper insights. Moreover, future research should investigate the impact of hybrid models that combine the interpretability of decision trees with the predictive power of neural networks or boosting algorithms. The integration of temporal data and dynamic feature updates may also enhance model responsiveness to evolving fraud patterns.

Organizations aiming to implement automated financial fraud detection systems should prioritize Gradient Boosting and Random Forest algorithms due to their high sensitivity and balanced performance across metrics. Practitioners should also invest in robust data preprocessing pipelines, including normalization and feature

selection, to maximize model effectiveness. While using advanced algorithms, it is critical to accompany them with explainability tools (such as SHAP values) to improve interpretability and regulatory compliance. Finally, institutions should consider continuous model monitoring and retraining mechanisms to adapt to emerging fraud tactics and maintain long-term effectiveness.

Authors' Contributions

Authors equally contributed to this article.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all participants who participate in this study.

Conflict of Interest

The authors report no conflict of interest.

Funding/Financial Support

According to the authors, this article has no financial support.

References

[1] S. Mehrani and A. Rahimipour, "Optimizing the Beneish Fraud Model in Predicting Financial Statement Restatements Using a Combination of Neural Networks and Genetic Algorithms," *Journal of Accounting and Management Auditing,* vol. 54, pp. 73-87, 2025.

[2] H. Lak, "Application of Artificial Intelligence and Machine Learning in Continuous Auditing and Detecting Financial Fraud," 2024.

[3] A. A. S. Alsuwailem, E. Salem, and A. K. J. Saudagar, "Performance of different machine learning algorithms in detecting financial fraud," *Computational Economics,* vol. 62, no. 4, pp. 1631-1667, 2023, doi: 10.1007/s10614-022-10314-x.

[4] E. A. Minastireanu and G. Mesnita, "An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection," *Informatica Economica,* vol. 23, no. 1, 2019, doi: 10.12948/issn14531305/23.1.2019.01.

[5] M. N. Nezami, F. M. Nodeh, and S. Khordyar, "Modeling Bankruptcy Prediction with Emphasis on Modern Measurement Methods Using Neural Networks and Support Vector Machines," *Journal of Accounting and Management Auditing,* vol. 55, pp. 265-278, 2025.

[6] M. Asadi and A. Rad, "Fraud Detection in Banking Transactions Using Hyperparameter Optimization of the Support Vector Machine Algorithm," 2023.

[7] Z. Zhao and T. Bai, "Financial fraud detection and prediction in listed companies using SMOTE and machine learning algorithms," *Entropy,* vol. 24, no. 8, p. 1157, 2022, doi: 10.3390/e24081157.

[8] M. Ali, S. A. Mirarab Baygi, and N. Farjian, "Presenting a Model for Predicting Financial Bankruptcy Risk in Listed and Over-the-Counter Companies Using Machine Learning Algorithms," *Capital Market Analysis Journal,* vol. 2, pp. 79-99, 2022.

[9] M. Mahmoudi and M. Shahrokh, "Machine Learning in Fraud Detection," 2024.

[10] H. Kamrani and B. Abedini, "Developing a Model for Detecting Financial Statement Fraud Using Artificial Neural Networks and Support Vector Machines in Companies Listed on the Tehran Stock Exchange," *Journal of Accounting and Management Auditing,* vol. 41, pp. 285-314, 2022.

[11] W. Duan, N. Hu, and F. Xue, "The information content of financial statement fraud risk: An ensemble learning approach," *Decision Support Systems,* vol. 182, p. 114231, 2024, doi: 10.1016/j.dss.2024.114231.

[12] S. Alizadeh Fard, "Presenting a Novel Framework Based on Collective Learning for Detecting Fraud in Financial Data," 2023.

[13] S. Ghorbani, "Analysis and Explanation of Financial Accounting Theory Based on the Conceptual Framework of the Financial Accounting Standards Board," *Journal of Accounting and Management Auditing,* vol. 53, pp. 37-53, 2025.
[14] M. S. Anari and M. K. Yazdi, "Application of Data Envelopment Analysis and Machine Learning in Detecting Accounting Fraud," 2024.
[15] Y. Chen and Z. Wu, "Financial fraud detection of listed companies in China: A machine learning approach," *Sustainability,* vol. 15, no. 1, p. 105, 2022, doi: 10.3390/su15010105.
[16] Y. Alghofaili, A. Albattah, and M. A. Rassam, "A financial fraud detection model based on LSTM deep learning technique," *Journal of Applied Security Research,* vol. 15, no. 4, pp. 498-516, 2020, doi: 10.1080/19361610.2020.1815491.
[17] T. Sadeghi and A. Nodehi, "Fraud Detection in Bank Cards Based on Image Processing Using Machine Learning Algorithms," 2023.