DIGITAL TRANSFORMATION AND ADMINISTRATION INNOVATION

Citation: Menatpour, Z., & Farsad Amanollahi, G. (2025). The Impact of Artificial Intelligence Technology on Enhancing the Efficiency of Auditing Processes and Increasing Financial Transparency. *Digital Transformation and Administration Innovation*, 3(1), 1-13.

Received date: 2024-11-16 Revised date: 2025-02-23 Accepted date: 2025-02-25 Published date: 2025-03-01

The Impact of Artificial Intelligence Technology on Enhancing the Efficiency of Auditing Processes and Increasing Financial Transparency

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

1. Master's Degree, Department of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran
2. Department of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran

*Correspondence: g_farsad@iauctb.ac.ir

Abstract

This study was conducted with the aim of investigating the innovations of artificial intelligence (AI) technology in improving auditing processes and enhancing financial transparency. Given the increasing complexity of data and financial reports, the application of advanced machine learning algorithms and natural language processing enables more accurate analysis and the automation of repetitive auditing tasks. Utilizing an analytical-applied approach and relying on valid financial data, this study employed algorithms such as decision trees, random forests, artificial neural networks, and gradient boosting to detect financial fraud and assess the transparency of auditing processes. The principal innovation of this research lies in the integration of multiple AI algorithms and the application of advanced feature selection methods (based on variance, correlation, and mutual information), which led to increased precision and sensitivity in fraud detection and transparency analysis. The results indicate that the random forest algorithm achieved an accuracy rate of 92%, and the gradient boosting algorithm attained an accuracy of 95%, demonstrating outstanding performance in fraud identification. Meanwhile, artificial neural networks, with an accuracy of 90%, succeeded in detecting more complex patterns, although they were accompanied by computational challenges. This study emphasizes that AI, as an innovative technology, plays a key role in enhancing financial transparency and trust in audit reports by reducing human errors and increasing efficiency. However, it also requires high-quality data and the development of methods for interpreting results.

Keywords: artificial intelligence, financial auditing, financial transparency, financial fraud, machine learning algorithms.

1. Introduction

Today, financial processes have become increasingly complex and diverse, such that transparency and accuracy in financial reporting have become primary concerns for economic actors, auditors, and regulatory institutions (Ahmadian et al., 2024). Fraud in financial statements is a phenomenon that can severely damage the credibility of organizations and, in the long term, erode public trust in financial markets (Khorsheed et al., 2024). This issue has gained particular significance in the context of a constantly evolving global economy (Mirchi et al., 2019). In this regard, the use of traditional auditing methods—primarily reliant on manual analysis and human resources—is no longer capable of addressing all the complex and extensive aspects of

financial fraud (Sreseli, 2023). Therefore, the need to employ modern technologies, including artificial intelligence (AI) and machine learning, for the detection and prevention of financial fraud has become increasingly apparent (Malek Akhlagh & Rafieepour Chirani, 2024).

AI, with its capacity to analyze large volumes of data and detect hidden patterns, has emerged as a critical tool in auditing processes (Fedyk et al., 2022). These algorithms are capable of identifying anomalies and irregular behaviors that may go undetected using traditional methods (Stein Smith & Dargahi, 2021). The application of these technologies in the analysis of financial statements contributes to enhanced accuracy and transparency in reporting, thereby improving auditing processes. Ultimately, this study will explore how AI can be utilized to increase the precision in detecting financial fraud and enhancing financial transparency within organizations (Fidyah et al., 2024).

Although the use of AI in financial auditing and fraud detection promises significant advancements, it is also accompanied by challenges that must be addressed in its practical application (Rane et al., 2023). One of the key challenges is the selection of appropriate features for identifying financial fraud. Given the massive volume of financial data and the inherent complexities of fraud detection, AI models require precise and relevant features to analyze the data effectively (Sreseli, 2023). Furthermore, the substantial variation in types of financial fraud and the continuous evolution of fraudulent patterns necessitate the constant updating of AI models (Vaghefi & Darabi, 2019). Additionally, issues related to data quality, the lack of valid datasets, and privacy concerns—especially when analyzing sensitive financial data—present further barriers to the widespread adoption of AI in this domain (Javanmiri, 2024). These challenges underscore the need for further research aimed at improving and refining AI models to effectively combat complex financial fraud (Kuswara et al., 2024).

Numerous studies have been conducted on the application of AI and machine learning in the fields of finance and auditing, focusing on evaluating and enhancing financial transparency and fraud detection processes (Ghaemi Asl et al., 2023; Jiao et al., 2022). These studies highlight the importance of leveraging advanced technologies to improve accuracy, efficiency, and transparency in financial reporting, particularly through the use of machine learning algorithms for identifying and forecasting financial status and fraud. Haroonkalai and colleagues (2023) examined the financial variables influencing the recovery of distressed companies listed on the Tehran Stock Exchange. In this study, feature selection algorithms such as LARS and Relief were employed to identify ten significant financial variables. The results showed that the support vector machine algorithm outperformed decision trees in predicting the timeline for exiting financial distress (Haroonkolaei & Barzegar, 2023). Al-Omush and colleagues (2025) investigated the impact of AI and machine learning on financial auditing processes. Utilizing a mixed-methods approach, their study demonstrated that AI can enhance audit efficiency, reduce errors, and improve risk assessment capabilities. The findings indicated that AI could support improved decision-making in auditing and increase transparency in financial reports (Al-Omush et al., 2025).

Despite the positive outcomes reported in many studies concerning the use of AI in auditing and fraud detection, gaps still exist in this field that warrant further research (Matin Fard & Dastbaz, 2023). One such gap is the inconsistency of key variables across different fraud detection models, leading to conflicting findings that are not broadly applicable. Additionally, challenges related to the optimal selection of features in financial data and the need for continuous model updates due to the rapidly evolving nature of fraud have hindered the development of comprehensive and unified solutions. Therefore, research focused on identifying the best features and algorithms for detecting fraud in financial statements can significantly contribute to the expansion of AI applications in this domain (Taghipour et al., 2020).

The objective of this study is to evaluate the capability of AI in detecting financial fraud and improving the transparency and accuracy of auditing processes through financial data analysis and the identification of suspicious patterns. The innovation of this research lies in the use of advanced AI classification algorithms for fraud detection and the application of novel feature selection methods to enhance the accuracy and performance of predictive models. This study also emphasizes the development of adaptive and flexible models that can adjust to changing fraud patterns and yield greater efficiency in fraud detection. In the following sections, the research methodology, employed algorithms, and the results of the analyses will be discussed in detail.

2. Methods and Materials

2.1. Research Approach

This study employs an analytical-applied research approach, which integrates both theoretical analysis and practical application. In this approach, financial data are systematically and precisely analyzed to identify complex relationships and hidden patterns in auditing processes and financial transparency. Subsequently, the results of these analyses are practically applied to improve auditing processes and detect financial fraud. This approach enables an in-depth examination of the impact of artificial intelligence (AI) technology on enhancing the accuracy, speed, and efficiency of auditing and specifically evaluates the capabilities of machine learning algorithms in detecting financial fraud. Moreover, due to its combination of quantitative and qualitative data analysis, this method allows for the provision of practical and executable solutions in real-world financial environments.

2.2. Data Collection Method

The data used in this study were collected from official, credible, and up-to-date sources to ensure the validity and reliability of the results. Financial information of companies was extracted from the Comprehensive Database of Publishers' Information (Codal) and the database of the Securities and Exchange Organization of Iran. These data include annual and interim financial statements, independent audit reports, explanatory notes, and other relevant documents that are available in structured and standardized formats. Additionally, data related to financial fraud cases were collected from official audit reports, judicial case files, and relevant legal documents. The data collection process adhered to research ethics principles, maintained confidentiality, and utilized valid and reliable data. Furthermore, to enhance data comprehensiveness, a diverse sample of companies active in various industries was selected to increase the generalizability of the findings.

2.3. Data Preprocessing

The data preprocessing stage is one of the most critical phases of the research, ensuring the quality and accuracy of subsequent analyses. At this stage, the data were cleaned of potential errors, missing values, and existing noise. Then, incomplete or inconsistent data were removed or corrected to prevent deviations in the results. Data were standardized to unify scales and formats, enabling AI algorithms to process the data effectively and unambiguously. Additionally, textual data, such as audit descriptions, were converted into analyzable numerical formats (e.g., using natural language processing techniques). Finally, the data were divided into training and testing sets and prepared for model training and evaluation. This preprocessing process provides a solid foundation for the successful implementation of machine learning algorithms.

2.4. Feature Selection

Selecting relevant and effective features plays a crucial role in improving the performance of AI models. In this study, advanced feature selection methods were used, which prioritize and identify the features most strongly associated with detecting financial fraud based on statistical and informational criteria. These methods include variance-based filters (eliminating low-variance features), correlation analysis (identifying features correlated with the target variable), and mutual information measures, which evaluate the amount of shared information between each feature and the fraud label. This process reduces data dimensionality, eliminates irrelevant features, and decreases the computational complexity of the models, thereby improving the accuracy, speed, and generalizability of AI models. Additionally, feature selection helps reduce the likelihood of overfitting, enabling models to perform better on new data.

2.5. Classification Algorithms

In this study, several advanced machine learning algorithms were used for classifying and detecting financial fraud. These algorithms include decision trees, random forests, and artificial neural networks, each with specific advantages and limitations. The models were first trained using training data and then used to predict and classify new samples. The purpose of employing multiple algorithms was to compare performance and select the best model based on criteria such as accuracy, sensitivity, and generalizability.

2.5.1. Decision Tree

The decision tree, as an interpretable and comprehensible algorithm, hierarchically divides data based on the features with the highest discriminatory power. Feature selection criteria such as entropy and the Gini index were used to optimize the tree structure. Due to its simplicity, high speed, and interpretability of results, this algorithm is effectively used in detecting anomalies and financial fraud and can serve as an auxiliary tool for auditors in data analysis.

2.5.2. Artificial Neural Networks

Artificial neural networks with multilayer structures and the ability to learn complex and nonlinear patterns were employed for financial data analysis. These networks, using a training process based on sample data, are capable of predicting and detecting financial fraud in new data. However, the need for high computational resources, long training times, and challenges in interpreting results are among the limitations of this method. Nevertheless, due to their powerful learning capabilities, neural networks are highly effective in identifying hidden and complex fraud patterns.

2.5.3. Random Forest

The random forest, as an algorithm based on the ensemble of multiple decision trees, is designed to increase accuracy, reduce variance, and prevent overfitting. By generating a set of independent decision trees and combining their results, this algorithm has demonstrated strong performance in identifying complex and nonlinear financial fraud patterns. Due to its high resistance to noise and superior generalizability, the random forest is considered one of the most popular algorithms in the field of fraud detection.

2.6. Results Analysis

After the implementation of machine learning algorithms, the results of the models were analyzed using standard evaluation metrics, including accuracy, recall, precision, and the F1-score. These metrics comprehensively assess the models' ability to correctly identify fraudulent cases (true positives) and reduce Type I and Type II errors. In addition, supplementary statistical analyses were conducted to examine the influence of each financial feature on model performance, allowing the identification of key factors in fraud detection. The results were interpreted both quantitatively and qualitatively, and the strengths, limitations, and improvement suggestions for each algorithm were discussed. Moreover, sensitivity analysis was performed to evaluate the models' robustness and generalizability with respect to variations in input parameters and training data.

2.7. Tools Used

For implementing AI models and analyzing data, the programming languages Python and R were utilized. In Python, specialized libraries such as Pandas for data management and processing, NumPy for numerical computations, Scikit-learn for executing machine learning algorithms, and TensorFlow and Keras for developing and training neural networks were used. In the R environment, statistical and machine learning packages were used for complementary analyses. Additionally, statistical software such as SPSS and SAS was employed for advanced statistical analyses, hypothesis testing, and fraud pattern identification. The use of these tools enabled precise preprocessing, the training of complex models, comprehensive evaluation, and valid statistical analysis, contributing significantly to the accuracy and credibility of the research findings.

3. Findings and Results

3.1. Description of the Research Data

The data used in this study to examine artificial intelligence (AI) in auditing processes and financial transparency were collected from credible sources and well-known databases. One of the primary sources of these data was the Codal system, which provides companies' balance sheets, financial statements, and financial reports. These data are regularly updated and include key information such as cash flows, expenses, and profit and loss figures. In addition, data such as financial histories,

transactions, and budgeting information were collected from various sources. These data were directly utilized in the process of financial transparency and in enhancing audit accuracy using AI algorithms. In this research, particular emphasis was placed on corporate financial statement data to assess the impact of AI on clarifying and improving audit process accuracy.

3.2. Data Preprocessing

The data preprocessing procedure comprises several essential steps to prepare the data for use in AI models. These stages include data cleansing, data standardization, and transformation into appropriate formats. Each of these steps contributes to optimizing the data so that AI algorithms can analyze and detect financial fraud with greater precision and efficiency. Data cleansing is essential for ensuring the accuracy of AI models in detecting fraud. This step focuses on eliminating incorrect data, missing values, and anomalies, which could negatively impact model performance. In financial datasets, missing values may result from human error or incomplete reports. To address this issue, missing values are either imputed with the mean or removed. Furthermore, algorithms such as Isolation Forest and DBSCAN are used to detect anomalies, ensuring the data are properly prepared for model training.

The next step in preprocessing is standardization or normalization of the data. In financial data, variables often have different scales. For instance, a company's revenue may be significantly larger than its costs or profit. These scale differences can cause certain features to appear more influential during algorithm learning, even if they have weaker correlations with fraud. To resolve this, the data must be rescaled to comparable levels. At this stage, financial features are converted into a standard scale so AI algorithms can process them more accurately. Standardization is typically performed using Z-score normalization or Min-Max scaling. For example, in Z-score normalization, each feature is standardized using its mean and standard deviation, ensuring that all features lie on a similar scale.

A critical step in preprocessing is transforming data into a suitable format for AI models. Qualitative and non-numeric data, common in many financial systems, must be converted into numerical values for processing by AI algorithms. This is especially important for non-numeric variables such as categorical variables or features like transaction type (e.g., purchase, sale, or loan). Various methods are used for this purpose, including Label Encoding and One-Hot Encoding. In Label Encoding, each category is assigned a number (e.g., "purchase" is converted to 0 and "sale" to 1). One-Hot Encoding, on the other hand, creates a new binary column for each category, indicating its presence or absence. These methods enable AI models to accurately process categorical data and detect patterns associated with financial fraud.

3.3. Data Analysis

3.3.1. Feature Selection and Tools Used in Fraud Detection

Based on Table 1, three main methods were employed in this study for feature selection: variance-based feature selection, correlation-based feature selection, and mutual information-based feature selection. These methods were used to identify features related to financial fraud and to help AI models make more accurate predictions while avoiding misleading patterns. Initially, variance-based feature selection was applied. This method identifies features with high variability and data discrimination capability. Features with greater variance typically provide more useful information for detecting financial fraud. It selects features with wide value ranges, which may indicate hidden patterns associated with fraud.

The next method used was correlation analysis. This technique selects features that exhibit strong relationships with the target variable (fraud or non-fraud). Pearson and Spearman correlation coefficients were employed to analyze these relationships. These indices help identify features that are most strongly correlated with fraudulent occurrences and are thus suitable for model input.

Finally, mutual information was used to select features most relevant to the labels. This method identifies features that share the highest amount of information with the target variable. Features with high mutual information are typically more important in predicting and detecting fraud and assist models in simulating more complex patterns. Using these three methods, key features that support the detection and prediction of fraud were selected, thereby improving the accuracy and performance of AI models in financial fraud detection.

Table 1. Feature Selection Methods and Tools Used

Method	Description	Tools Used
Variance-based Feature Selection	Identifies features with greater variability and better discrimination power	Python (Scikit-Learn)
Correlation-based Feature Selection	Identifies features more strongly correlated with the target variable	Python, R
Mutual Information-based Selection	Identifies features with highest shared information with labels	Python, MATLAB

3.3.2. Performance Evaluation Metrics of Algorithms

This section evaluates the most important performance metrics of AI algorithms in detecting financial fraud. The evaluation of AI models using various metrics captures different dimensions of model performance. Table 2 presents metrics such as Accuracy, Sensitivity, Specificity, F1-score, and AUC-ROC, along with full explanations and corresponding formulas. This table provides comprehensive information for comparing and selecting the best model for financial fraud detection.

- **Accuracy**
- Accuracy is one of the primary metrics for evaluating AI models, indicating the model's ability to correctly predict different cases. In other words, accuracy is the ratio of correct predictions to total predictions made and is especially relevant in financial fraud detection. It reflects how well the algorithm correctly identifies both fraudulent and non-fraudulent instances.
- **Sensitivity (Recall)**
- Sensitivity, also known as recall, refers to the model's ability to correctly identify fraud cases. In financial fraud detection, sensitivity is critically important, as failing to detect even a single case can result in significant financial losses. This metric reveals how many fraudulent cases the model successfully identifies.
- **Specificity**
- Specificity, or the true negative rate, refers to the model's ability to correctly identify non-fraud cases. This metric is essential for reducing false alarms and avoiding unnecessary costs caused by misclassifying non-fraudulent data. High specificity indicates the model's success in accurately recognizing non-fraud cases.
- **F1-Score**
- The F1-score is the harmonic mean of precision and recall and is a balanced metric for evaluating overall model performance. It is especially useful when a trade-off between accuracy and sensitivity is required, particularly in scenarios where overlooking or misidentifying fraud could have serious consequences. The F1-score provides a general indication of the model's performance in fraud detection.
- **AUC-ROC Metric**
- AUC-ROC (Area Under the Receiver Operating Characteristic Curve) plots sensitivity versus false positive rate across all classification thresholds. The AUC value represents the area under this curve and is a measure of the overall quality of the classification model. An AUC close to 1 indicates excellent model performance in distinguishing between fraudulent and non-fraudulent cases.

Qualitative and temporal metrics, alongside quantitative metrics, are also important for evaluating algorithms. These metrics include processing speed, computational efficiency, and stability over time. Processing speed is critical when dealing with large volumes of financial data, while computational efficiency refers to the algorithm's ability to optimize system resource usage. Model stability in the face of data fluctuations also plays a key role in ensuring the consistent and accurate performance of algorithms in fraud detection.

Table 2. Performance Evaluation Metrics for Artificial Intelligence Algorithms in Financial Fraud Detection

Metric	Description	Formula	Variables
Accuracy	The ratio of correct predictions to the total number of predictions. It indicates the model's ability to correctly identify various cases.	(TP + TN) / (TP + TN + FP + FN)	TP: True positives – correctly predicted fraud cases. TN: True negatives – correctly predicted non-fraud cases. FP: False positives. FN: False negatives.
Sensitivity (Recall)	The model's ability to correctly identify fraud cases. This metric is crucial in financial fraud detection.	TP / (TP + FN)	TP: True positives. FN: False negatives.
Specificity	The model's ability to correctly identify non-fraud cases, essential for reducing false alarms.	TN / (TN + FP)	TN: True negatives. FP: False positives.
F1 Score	The weighted harmonic mean of precision and recall, balancing the two in fraud detection tasks.	(2 × Precision × Recall) / (Precision + Recall)	Precision: TP / (TP + FP) Recall: TP / (TP + FN)
AUC-ROC	A metric for evaluating overall model quality using the ROC curve, which shows sensitivity versus false positive rate at various thresholds.	–	–
Qualitative & Temporal	Includes processing speed, computational efficiency, and model stability against data fluctuations for ensuring consistent and precise performance.	–	–

3.4. Evaluation of Classification Algorithms

3.4.1. Decision Tree

The decision tree is one of the key and widely used algorithms in classification tasks, playing a significant role in detecting financial fraud. The algorithm's performance based on various metrics is illustrated in Figure 1. The accuracy of this algorithm is 80 percent, indicating that 80 percent of its predictions were correctly classified into fraudulent and non-fraudulent categories. The model's sensitivity is 75 percent, which demonstrates the algorithm's success in correctly identifying 75 percent of the fraudulent cases. The specificity of the algorithm is 85 percent, reflecting strong performance in distinguishing non-fraudulent instances. This means the model correctly identified 85 percent of non-fraud cases. The F1 score of this algorithm is 0.78, which serves as a balanced measure between precision and recall. This value represents a relative balance between these two metrics and is particularly important in fraud detection scenarios where such balance is critical.

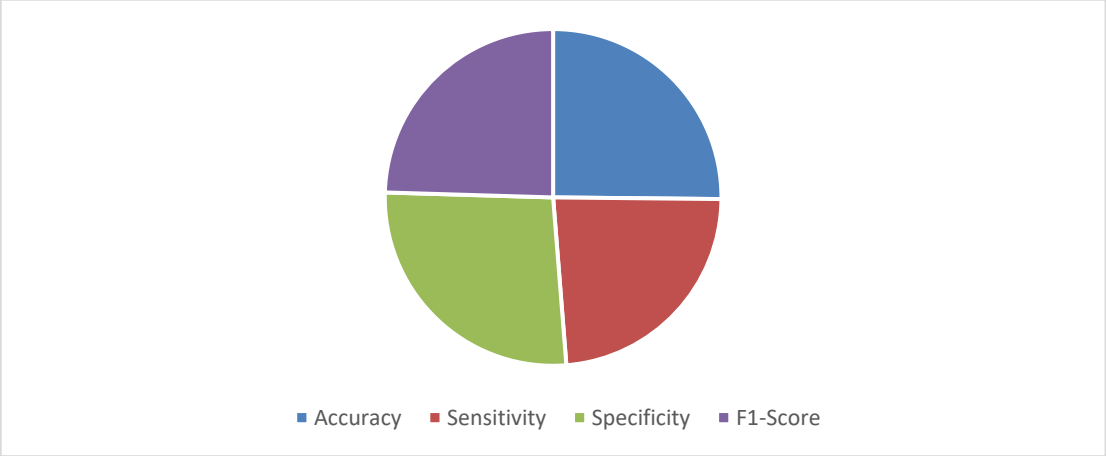

Figure 1. Performance Evaluation of the Decision Tree Algorithm

The evaluation results of the decision tree algorithm indicate relatively good performance in identifying financial fraud. However, in situations where the data distribution is imbalanced between fraud and non-fraud cases, the algorithm may require improvement. Despite its relative efficiency, the decision tree may not yield optimal results in more complex and imbalanced data conditions.

3.4.2. Artificial Neural Networks (ANN)

Artificial neural networks (ANN) are among the most widely used algorithms for detecting financial fraud. These algorithms process information in parallel through artificial neurons, making them capable of identifying complex and hidden patterns in the data. In Figure 2, the performance of ANNs is shown across various evaluation metrics. The accuracy of this model in detecting financial fraud reaches 90 percent, indicating a high ability to correctly predict fraudulent activities. Given the complex nature of financial data, neural networks exhibit strong capabilities in recognizing nonlinear patterns. The model's sensitivity is 88 percent, indicating it successfully detects 88 percent of fraudulent cases. This is especially important in minimizing false negatives and reducing the risk of missing fraud cases.

On the other hand, the specificity of this model is 82 percent, reflecting its ability to identify non-fraud cases accurately, though there is still room for improvement in this area. The F1 score of this model is 0.89, indicating a well-balanced performance between precision and recall. This value highlights the model's capacity to maintain equilibrium between two critical metrics in fraud detection and demonstrates that the ANN provided a balanced performance in identifying both fraudulent and non-fraudulent transactions.

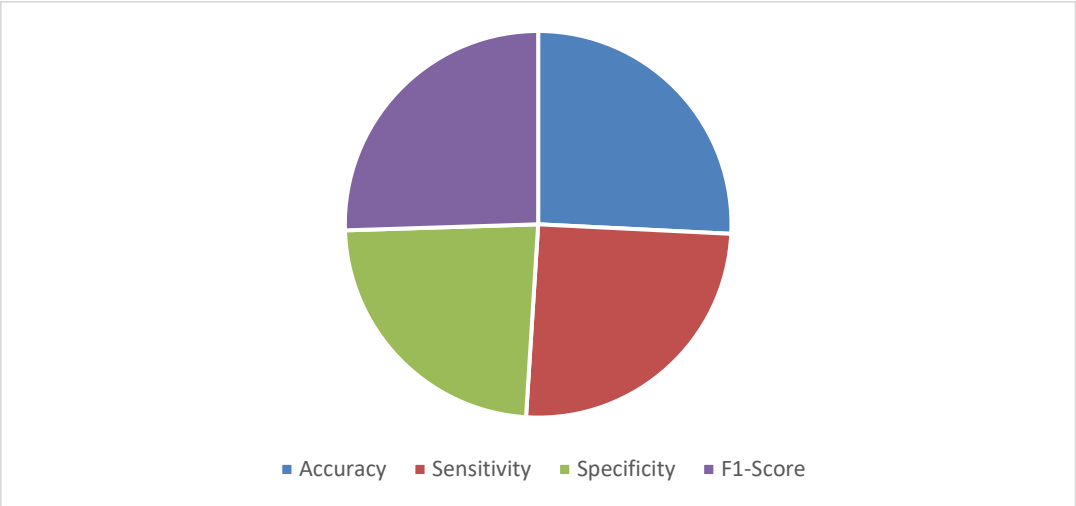

Figure 2. Performance Evaluation of Artificial Neural Networks in Detecting Financial Fraud

The results of evaluating artificial neural networks reveal high accuracy in fraud detection. This algorithm is capable of identifying complex features and hidden relationships within financial data, significantly enhancing prediction accuracy. However, key challenges of this algorithm include its high computational resource demands and long training time. Additionally, the complexity of neural network architectures can make result interpretation difficult for users, requiring high expertise for proper analysis.

3.4.3. Bayesian Network

The Bayesian Network is a graphical model that utilizes probability theory to analyze probabilistic relationships among variables. This algorithm is particularly effective in detecting financial fraud in cases where data are incomplete or variables operate independently. The assumption of feature independence is a core characteristic of this algorithm, aiding in data analysis and the identification of key features. As shown in Figure 3, the evaluation results indicate that the model's accuracy is 70%, and its sensitivity is 65%, reflecting the algorithm's ability to correctly detect 65% of fraud cases. This highlights the algorithm's limitation in identifying all fraudulent instances. The specificity of the model is 80%, indicating that 80% of non-fraud cases were accurately classified. The F1 score of the model is 0.67, revealing a weak balance between accuracy and sensitivity and suggesting the need for performance improvement to achieve more accurate fraud detection.

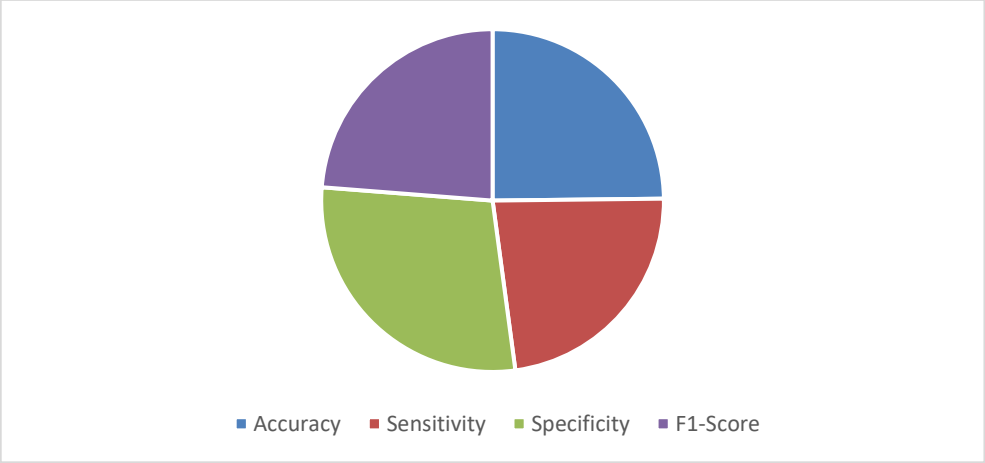

Figure 3. Performance of the Bayesian Network Based on Evaluation Metrics

3.4.4. Random Forest

In the realm of artificial intelligence, models such as the Random Forest are widely used for detecting complex patterns and accurately predicting financial fraud. These models have become effective tools for simulating complex decision-making in real-world environments due to their strong capabilities in processing large and intricate datasets. The Random Forest is an ensemble model that operates based on a collection of decision trees. It generates multiple random trees and combines their outcomes for the final prediction, achieving high performance in fraud detection. As illustrated in Figure 4, the model's accuracy is 92%, meaning that 92% of predictions were correct. This high level of accuracy, compared to other models, demonstrates the strong ability of Random Forest to distinguish between fraudulent and non-fraudulent cases. The model's sensitivity is 90%, indicating excellent performance in identifying fraud cases. Furthermore, the model's specificity is 88%, showing it is also highly capable of accurately detecting negative (non-fraudulent) instances. The F1 score is 0.91, demonstrating a well-balanced performance between accuracy and sensitivity in financial fraud detection.

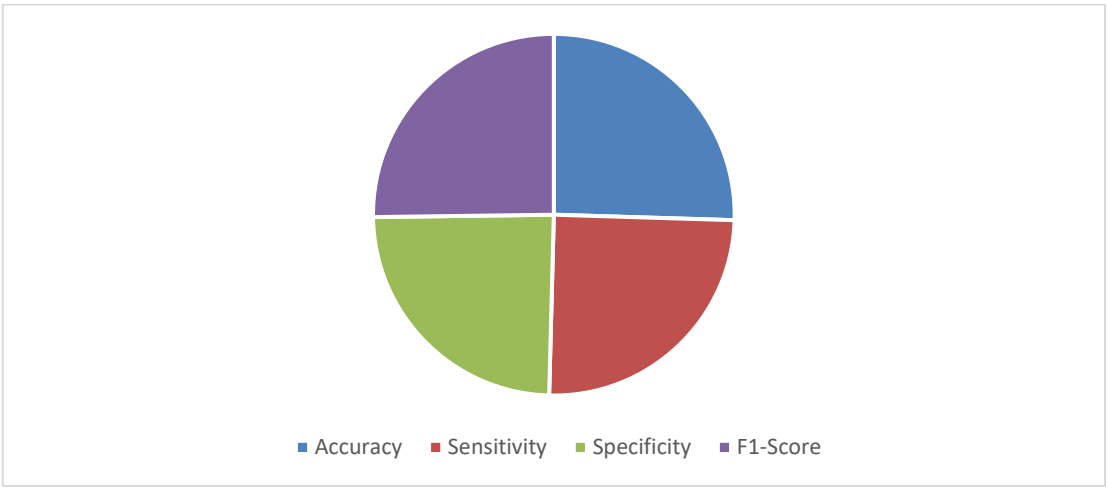

Figure 4. Performance Evaluation Results of the Random Forest Model in Detecting Financial Fraud

These results indicate that the Random Forest not only performs with high precision in identifying financial fraud, but also effectively detects non-fraud cases. The desirable balance between accuracy and sensitivity optimizes the model's performance in complex scenarios. Given its ability to analyze large and complex datasets, this algorithm proves highly useful in real-world fraud detection environments.

3.4.5. Gradient Boosting

In Figure 5, the performance evaluation results of the Gradient Boosting model in detecting financial fraud are presented in detail. This algorithm achieved a very high accuracy of 95% in detecting both fraudulent and non-fraudulent cases. The model's

high accuracy demonstrates its capacity to provide precise and effective predictions. Additionally, its sensitivity is 92%, indicating the model's strong ability to identify a high proportion of fraud cases. Moreover, the model's specificity is 90%, showing its effectiveness in correctly classifying non-fraud cases. The F1 score, a balanced metric for accuracy and sensitivity, is 0.94, reflecting an excellent balance between these two measures in the model's performance.

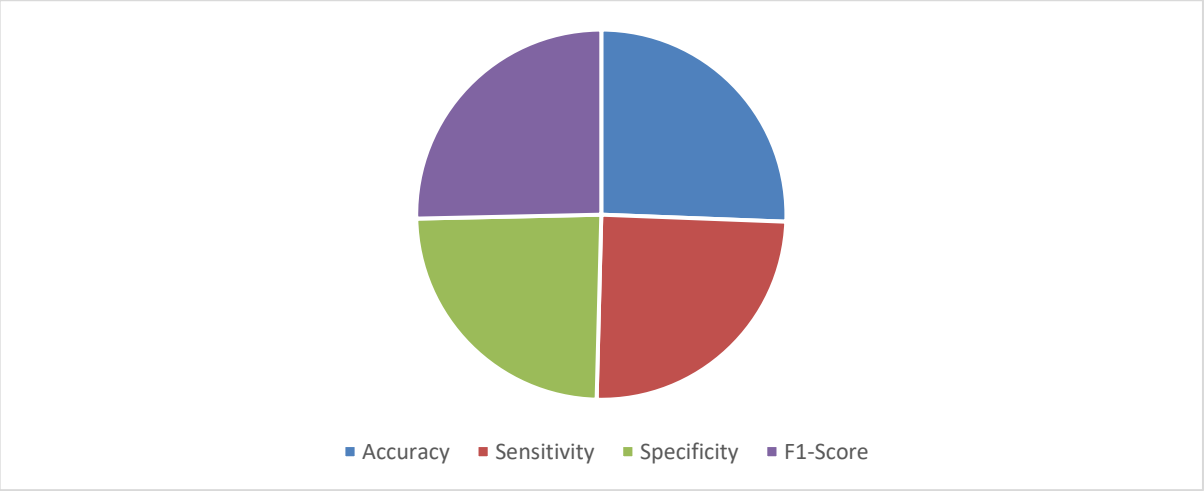

Figure 5. Performance Evaluation Results of the Gradient Boosting Model in Detecting Financial Fraud

These results show that Gradient Boosting not only excels in identifying fraudulent cases but also performs well in classifying non-fraudulent instances. The F1 score of 0.94, which indicates a desirable balance between accuracy and sensitivity, underscores this algorithm's high efficiency in accurately predicting financial fraud.

Among the strengths of the Gradient Boosting algorithm is its precision and capacity for fraud detection. By focusing on the errors of previous models, this algorithm continuously improves in identifying complex fraud. However, its drawbacks include longer training times and the need for parameter tuning, which may reduce model performance in rapidly changing data environments. Furthermore, without careful parameter adjustment, there is a risk of overfitting, which can reduce prediction accuracy on new data.

4. Discussion and Conclusion

The results of the present study highlight the remarkable efficacy of artificial intelligence (AI) algorithms in the detection of financial fraud and the enhancement of auditing accuracy. Among the evaluated models, Gradient Boosting and Random Forest demonstrated the highest levels of precision, sensitivity, specificity, and F1 scores. The Gradient Boosting model achieved an overall accuracy of 95%, a sensitivity of 92%, a specificity of 90%, and an F1 score of 0.94, positioning it as the most effective algorithm in the study. The Random Forest model followed closely, with an accuracy of 92%, sensitivity of 90%, specificity of 88%, and an F1 score of 0.91. These results suggest that ensemble methods outperform individual classifiers in terms of predictive power and generalizability, particularly in the context of large and complex financial datasets.

The superior performance of ensemble methods aligns with prior research emphasizing their robustness in dealing with high-dimensional and imbalanced data. For instance, Al-Omush et al. (2025) demonstrated that the implementation of AI-based models in auditing environments enhances fraud detection capabilities and improves overall risk assessment by minimizing false positives and negatives (Al-Omush et al., 2025). Their findings underscore the importance of integrating AI tools into auditing practices to ensure more precise and scalable decision-making processes. Similarly, Rane (2023) confirmed that Gradient Boosting and Random Forest models have a notable advantage over simpler classifiers, especially in contexts requiring the identification of subtle fraud patterns across dynamic financial transactions (Rane et al., 2023).

Artificial Neural Networks (ANNs) also performed admirably in this study, achieving 90% accuracy, 88% sensitivity, 82% specificity, and an F1 score of 0.89. These findings reinforce those of Fidyah (2024), who documented the effectiveness of ANNs in capturing nonlinear relationships and hidden dependencies within complex financial datasets. ANNs are particularly advantageous in identifying atypical or outlier behaviors often associated with fraudulent activities. However, as also noted by Fidyah (2024), this modeling approach comes with computational costs and interpretability challenges, which limit its

widespread adoption in resource-constrained auditing environments (Fidyah et al., 2024). Despite these constraints, the performance metrics observed in this research validate the ANN model as a viable tool for institutions seeking a balance between complexity and predictive power.

In contrast, the Bayesian Network algorithm yielded relatively weaker performance results, with an accuracy of 70%, sensitivity of 65%, specificity of 80%, and an F1 score of 0.67. These findings suggest a diminished capacity of the Bayesian model to handle dependencies among financial variables, especially in cases where the assumption of feature independence is violated. As supported by Sreseli (2023), Bayesian classifiers often underperform in real-world financial settings due to their structural simplicity and reliance on probabilistic assumptions (Sreseli, 2023). These limitations reduce their suitability for detecting fraudulent behavior, which often involves intricate relationships and non-linear interdependencies across multiple financial indicators.

The Decision Tree model showed moderate performance with 80% accuracy, 75% sensitivity, 85% specificity, and an F1 score of 0.78. Although interpretable and computationally efficient, its lower sensitivity suggests a reduced capability in capturing all fraud instances, particularly in skewed datasets where fraudulent transactions are rare compared to legitimate ones. This limitation is echoed by Haroonkalai et al. (2023), who found that Support Vector Machines (SVMs) and other advanced classifiers consistently outperform Decision Trees in identifying financial anomalies (Haroonkolaei & Barzegar, 2023). The relatively inferior results of the Decision Tree algorithm in this study, compared to ensemble and deep learning methods, emphasize the necessity of using more sophisticated tools in modern auditing.

A significant takeaway from the comparison between AI-based methods and traditional auditing techniques is the transformative potential of AI in detecting complex fraud patterns. Traditional methods, such as manual inspection and rule-based systems, are time-intensive and prone to human error. As highlighted by Khorsheed (2024), the reliance on human judgment in traditional auditing increases the risk of oversight, especially in large-scale financial systems (Khorsheed et al., 2024). In contrast, AI-based systems process vast amounts of data with consistent accuracy and efficiency, providing auditors with timely insights. Al-Omush et al. (2025) further noted that AI improves transparency and reduces operational risks by standardizing auditing procedures, thereby promoting regulatory compliance (Al-Omush et al., 2025).

Moreover, AI tools offer adaptability and scalability, features that are particularly beneficial in evolving fraud landscapes. Fraudulent tactics are increasingly sophisticated, requiring dynamic detection mechanisms capable of continuous learning and adjustment. As demonstrated in this study, Gradient Boosting and Random Forest models excelled in capturing these evolving fraud indicators, outperforming static rule-based approaches. These findings corroborate the conclusions of Rane (2023), who emphasized that ensemble-based AI methods are well-suited for adaptive fraud detection due to their ability to aggregate insights from multiple classifiers and adjust to changing patterns in financial data (Rane et al., 2023).

Despite their superior performance, AI algorithms are not without limitations. The current study identified key challenges that align with those noted in prior literature. Most notably, the success of AI models is contingent on the availability of high-quality, error-free data. As emphasized by Al-Omush et al. (2025), data incompleteness, inconsistency, and inaccuracies can significantly impair model accuracy (Al-Omush et al., 2025). This concern is especially pertinent in financial auditing, where missing values and inconsistently formatted records are common. Moreover, the interpretability of complex models, such as ANNs and Gradient Boosting, remains a barrier to their practical implementation. While these models offer high predictive accuracy, they often lack transparency in their decision-making processes, which complicates their use in legal and regulatory contexts where auditors must justify their findings.

The computational demands associated with complex models also warrant consideration. As shown in this study, models such as Gradient Boosting and ANNs require longer training times and greater processing power, which may not be feasible for smaller firms with limited IT infrastructure. Additionally, the risk of overfitting—particularly in models that rely heavily on parameter tuning—can reduce model performance when applied to new datasets. This issue was observed in the performance stability of the Gradient Boosting model, which, despite high accuracy, necessitated careful calibration to avoid excessive sensitivity to training data. These practical limitations must be addressed to ensure the scalable and responsible deployment of AI in auditing environments.

This study is not without its limitations. First, the research relied on secondary financial datasets extracted from structured repositories such as Codal, which may not fully represent the breadth of real-world data inconsistencies. Additionally, while the study evaluated several leading classification algorithms, it did not include models such as Support Vector Machines (SVMs) or XGBoost, which have shown promising results in other studies. The generalizability of the findings may also be limited by the specific configuration of the models and the feature selection techniques applied. Another limitation is the absence of real-time fraud detection scenarios, which would be more representative of actual auditing environments. Finally, the study did not conduct a cost-benefit analysis of implementing these AI tools in various organizational contexts, which could influence practical adoption.

Future research should focus on expanding the range of algorithms to include additional ensemble and deep learning models, such as XGBoost, LightGBM, and recurrent neural networks. Incorporating real-time data streams would enhance the practical relevance of the findings and support the development of dynamic auditing systems. Moreover, studies should explore hybrid models that combine rule-based systems with AI to leverage both human judgment and machine precision. Further investigations into explainable AI (XAI) frameworks could help address interpretability issues, allowing auditors and regulators to better understand and trust AI-generated results. Future research should also consider conducting cross-industry comparisons to assess how AI performance varies across financial sectors with different regulatory requirements.

Organizations seeking to integrate AI into their auditing processes should prioritize data quality management, including consistent formatting, validation, and cleansing protocols. Investment in computational infrastructure and cloud-based solutions can help mitigate the resource limitations associated with advanced AI models. Training programs should be implemented to equip auditors with the skills needed to interpret and leverage AI-generated insights. Collaborative frameworks between data scientists and financial professionals are also essential to ensure that AI tools align with auditing objectives. Finally, companies should establish clear ethical guidelines and governance protocols for AI use to maintain transparency, accountability, and compliance with legal standards.

Ethical Considerations

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

Acknowledgments

Authors thank all who helped us through 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

Ahmadian, V., Fazelzadeh, A., Naghdi, S., & Yahya Salman, I. (2024). The Effect of Artificial Intelligence and Blockchain Technologies on Improving the Quality of Financial Reports. The Twenty-First International Conference on Research in Management, Economics and Development, https://civilica.com/doc/2055226/

Al-Omush, A., Almasarwah, A., & Al-Wreikat, A. (2025). Artificial intelligence in financial auditing: redefining accuracy and transparency in assurance services. *EDPACS*, 1-20. https://doi.org/10.1080/07366981.2025.2459490

Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies*, *27*(3), 938-985. https://doi.org/10.1007/s11142-022-09697-x

Fidyah, F., Usman, S., Pradita, A. E., & Setyawati, D. M. (2024). The Impact of Artificial Intelligence on Auditing Processes and Accuracy: A Future Outlook. *Dinasti International Journal of Economics, Finance & Accounting (DIJEFA)*, *5*(4). https://doi.org/10.38035/dijefa.v5i4.3224

Ghaemi Asl, M., Nasr Esfahani, M., & Kashani, L. (2023). *Artificial Intelligence and Economics: (Functions and Perspectives in Business Space, Economic Policy, Financial Industry and Islamic Finance)*. Kharazmi University. https://www.gisoom.com/book/44829972/

Haroonkolaei, K., & Barzegar, G. (2023). Explaining the Effective Financial Variables in Predicting Financial Recovery Using Artificial Intelligence Approach. *Economic Modeling*, *1*(1), 89-103. https://journals.iau.ir/article_703693.html

Jiao, D., Zhang, W., Rahnama Roodposhti, F., & Poureaskari Jorshari, F. (2022). *Artificial Financial Intelligence: FinTech and Financial Mathematics*. Hooshmand Tadbir. https://torob.com/p/6ae0fc20-3ddc-4136-89f1-456a3bd8e903/%D8%AF%D8%A7%D9%86%D9%84%D9%88%D8%AF-%DA%A9%D8%AA%D8%A7%D8%A8-artificial-financial-intelligence-in-china-financial-mathematics-and-fintech-1st-ed-2021/

Khorsheed, H. S., Ismael, N. B., & Mahmod, S. H. O. (2024). The impact of artificial intelligence and machine learning on financial reporting and auditing practices. *International Journal of Advanced Engineering, Management and Science*, *10*(6), 30-37. https://doi.org/10.22161/ijaems.106.4

Kuswara, Z., Pasaribu, M., Fitriana, F., & Santoso, R. A. (2024). Artificial Intelligence in Financial Reports: How it Affects the Process's Effectiveness and Efficiency. *Jurnal ilmu keuangan dan perbankan (jika)*, *13*(2), 257-272. https://doi.org/10.34010/jika.v13i2.12730

Malek Akhlagh, E., & Rafieepour Chirani, I. (2024). The Impact of Artificial Intelligence and Business Intelligence on Improving the Financial Performance of Companies. The First National Conference on Management in the Era of Transformation with Emphasis on Technology, Science and Practice, https://civilica.com/doc/2114653/

Matin Fard, M., & Dastbaz, S. H. (2023). The Impact of Artificial Intelligence Training of Auditors on the Timeliness of the Audit Report. The First International Conference on Modern Studies in Management, Economics and Accounting, https://civilica.com/doc/2055038/

Mirchi, A. R., Maleki, M., & Shams, K. (2019). Modeling the Effect of Brand Equity, Financial Transaction and LRFM Pattern on Banking Customers Loyalty Using Artificial Intelligence. The Second International Conference on Modern Research Solutions in Management, Accounting and Economics, https://civilica.com/doc/954952/

Rane, N., Choudhary, S., & Rane, J. (2023). Blockchain and Artificial Intelligence (AI) integration for revolutionizing security and transparency in finance. https://doi.org/10.2139/ssrn.4644253

Sreseli, N. (2023). Use of Artificial Intelligence for Accounting and Financial Reporting Purposes: A Review of the Key Issues. *American International Journal of Business Management (AIJBM)*, *6*(8), 72-83. https://www.aijbm.com/wp-content/uploads/2023/08/I687283.pdf

Stein Smith, S., & Dargahi, I. (2021). *Blockchain, Artificial Intelligence and Financial Services: Implications and Applications for Financial and Accounting Experts*. Simorgh Aseman Azargan Publishing Institute. https://cir.nii.ac.jp/crid/1130285377013618945

Taghipour, S., Goodarzi, A., & Bobillier, T. (2020). *Future Banking: Transformation of Banking Through Emerging Financial Technologies of Artificial Intelligence and Machine Learning (Theoretical and Applied Foundations)*. Virast. https://www.gisoom.com/book/11673797/

Vaghefi, S. H., & Darabi, R. (2019). Validation of Artificial Intelligence Algorithms in Predicting Financial Distress in the Industry and Mining Sector with Emphasis on the Role of Macroeconomic, Financial, Managerial and Risk Variables. *Journal of Business Research*, *21*(111), 213-243. https://pajooheshnameh.itsr.ir/article_36983.html?lang=fa