ISSN 2090-3359 (Print)
ISSN 2090-3367 (Online)

Advances in Decision Sciences

Volume 28
Issue 4
December 2024

Published by Asia University, Taiwan and Scientific and Business World

Identifying Fraudulent Financial Reports:
Verification Between the M-Score Model and the Auditor's Opinion

Nguyen Anh Phong
University of Economics and Law, Ho Chi Minh City, Vietnam, and
Vietnam National University, Ho Chi Minh City, Vietnam.
Corresponding Author Email: phongna@uel.edu.vn

Phan Huy Tam
University of Economics and Law, Ho Chi Minh City, Vietnam, and
Vietnam National University, Ho Chi Minh City, Vietnam.
Email: tamph@uel.edu.vn

Ngo Phu Thanh
University of Economics and Law, Ho Chi Minh City, Vietnam, and
Vietnam National University, Ho Chi Minh City, Vietnam.
Email: thanhnp@uel.edu.vn

Received: March 14, 2024; First Revision: April 5, 2024;
Last Revision: October 20, 2024; Accepted: November 26, 2024;
Published: December 31, 2024

Abstract

Purpose: This study examines and forecasts financial reporting fraud in listed enterprises using the M-score model and auditor opinions based on the fraud triangle model.

Methodology: Research data was collected from non-financial enterprises listed on the HSX and HNX exchanges from 2018-2022. This study uses today's popular machine learning methods to evaluate the performance of models to have a basis for recommendations (machine learning methods such as ANN, KNN, Decision Tree, and Random Forest) and gradient boosting algorithms (XGBoost and LightGBM). These methods help make decisions more accurately and help financial managers identify fraudulent financial reporting of companies early. This is consistent with requirements in management science and decision sciences.

Findings: The ANN model for the M-Score achieved the highest accuracy (97.9%) and F1-score (0.979). In comparison, the Decision Tree model was most effective for auditor opinions with an accuracy of 82.1% and an F1-score of 0.831. Additionally, the XGBoost algorithm consistently delivered strong results across both models, with an F1-score of 0.984 for M-Score and 0.942 for auditor opinions.

Originality/Value: In this article, this study relies on the fraud triangle theory, briefly finding the elements of the three factors from the fraud triangle model, combined with the auditor's opinion on all financial statements. From there, predict whether a company has fraudulent financial statements or not. This way, this study combines the financial statement fraud theory with reality based on auditors' comments. In addition, this study also compares the traditional forecasting method, M-score, to evaluate the performance of forecasting models.

Implications: The auditor opinion model holds practical value, integrating qualitative and quantitative insights for early fraud detection.

Limitations: Further empirical research is required to select indicators representing identifying signs in the fraud triangle model. The model based on auditors' opinions holds significant reference value as it integrates qualitative and quantitative aspects, thereby combining theory with practical application.

Keywords: Financial statement fraud, machine learning methods, Vietnamese listed companies, M-Score, Auditor opinion.

JEL Classifications: C11, C23, C53, E37.

1. Introduction

Financial reporting fraud is always a concern to researchers and investors because clean or honest financial statements are a significant source of information for investors, creditors, and others when making decisions. However, many fraudulent financial status declaration cases have recently become increasingly serious (Humpherys et al., 2011; Kamarudin et al., 2012; Yeh et al., 2010). The theory of the fraud triangle (Cressey, 1950) was the first study to highlight three critical issues that create fraudulent financial statements, including factors such as pressure, opportunity, and rationalization. Many researchers have developed methods or models to detect fraudulent corporate financial statements based on this theory. For example, researchers can use the Z-score model based on the pressure factor of debt default or bankruptcy (Altman, 2013). Alternatively, according to the fraud triangle model, one can utilize the M-Score model based on synthesizing factors contributing to fraudulent behavior (Beneish, et al., 2012). These methods are widely recognized for forecasting and rely on financial indicators computed manually.

Financial reporting fraud, especially in large companies, can have severe and far-reaching consequences, often leading to crises. A notable example occurred in 2001 when Enron and WorldCom corporations used accounting tricks, aided by auditors, to manipulate financial statements, inflating revenue and profits despite actual losses (Healy & Palepu, 2003). This misconduct triggered stock market collapses, severely impacting numerous investors. Similarly, in Vietnam, several instances of financial statement fraud have occurred in recent years, such as Northwest Mineral Investment Company in 2012 and Wood Industry Group in 2016. More recently, Tan Hoang Minh Group in 2023-2024 resorted to deceptive tactics to embellish financial statements, issuing multiple bond batches to attract capital. Likewise, Van Thinh Phat Group employed fraudulent methods, including fabricating customers with fake loans, falsifying documents, and providing collateral to create false loan documents, all with the assistance of auditing firms. Such fraudulent activities result in financial losses for investors, deplete state assets, and expose regulatory loopholes, prompting government regulatory adjustments. Given the increasing scale and sophistication of fraud cases, early detection is essential to alert investors and enable regulatory oversight.

Recently, the US accounting industry has directly addressed the responsibility of independent auditors in detecting fraud in financial statements in Statement on Auditing Standards (SAS) No. 82 titled “Considering Fraud in Financial Statement Audits” (Ramos & Lyons, 1997). The report requires the auditor to plan and perform the audit to obtain reasonable assurance that the financial statements are free from material misstatement, whether due to fraud or error. Based on auditing frameworks and standards, auditors can give opinions on partial, complete acceptance, or disagreement about compliance with information disclosed in the enterprise's financial reports. Previous studies have mainly relied on financial models to predict financial statement fraud and have not relied on the opinions of auditors commenting on the compliance of financial statements. In contrast, auditors are the people who assess and review actual financial data at the enterprise. Therefore, in this research, this study approaches financial statement fraud prediction based on the M-Score model and, at the same time, compares it with the results of comments from auditors to verify the authenticity and accuracy. To conduct research, this study applies algorithms and technology to forecast which popular methods this study uses in the article, such as

Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF), and KNN. This topic has been of current interest to many decision sciences and information systems researchers.

While traditional approaches, like the Beneish M-Score model (Beneish, et al., 2012), have been widely used to predict financial statement fraud, they often overlook the valuable insights provided by auditors, who have direct access to internal data and can better assess fraud risks. In this research, we extend fraud detection by combining the M-Score model with auditors' opinions, aiming to verify the authenticity and accuracy of the predictions. Utilizing advanced machine learning techniques such as artificial neural networks (ANN), decision trees (DT), and ensemble methods like XGBoost (Ali, et al., 2023) and Random Forest (El-Bannany, et al., 2021), we improve the predictive capability of fraud detection models. This approach reflects the growing interest among decision sciences and information systems researchers in leveraging machine learning for more accurate and comprehensive fraud detection (Hajek & Henriques, 2017; Kassem, 2022). This study is original in integrating the Beneish M-Score with auditors' qualitative opinions, offering a novel hybrid approach to verifying fraud predictions that has not been emphasized in prior literature. The study contributes to Decision Sciences by providing predictive models that support investors, auditors, and regulators in making more informed choices under uncertainty, thereby linking fraud detection with managerial and policy-related decision-making.

2. Literature Review

Fraud can be defined in many different ways. In the business environment, fraud is intentional deception, misappropriation of company assets, or manipulation of company financial data to benefit the perpetrator (Hall, 2007). According to International Standards on Auditing (ISA), fraud is an intentional act by one or more individuals in management, those responsible for administration, employees, or third parties. It involves deception to gain an unfair or illegal advantage. The Institute of Internal Auditors (IIA) defines fraud as any unlawful act characterized by deception, concealment, or breach of trust. These acts do not depend on threats of violence or force. Singleton, et al. (2006) state that no definite and immutable rule can be laid down as a general proposition in determining fraud, as it includes ways of surprise, trickery, and cunning, and it's not fair that others get cheated (Singleton, et al., 2006). Fraud always involves the intentional action of one person to gain an unfair advantage over another and can take many forms (Vaassen, 2004). According to the US Endowment Committee of the Treadway Commission (COSO), financial statement fraud is an intentional or reckless conduct based on erroneous information or omissions that result in a material misstatement of the financial statements (Beasley, 1996).

It can be said that the person who laid the foundation for research on financial fraud in general and financial reporting fraud in particular belongs to Cressey (1950), who discovered the fraud triangle theory. This theory is based on a survey of prisoners convicted of embezzlement and has become the basis of a large body of literature on crime (Cressey, 1953). According to Cressey (1950), pressure is the first cause of fraud because pressure often revolves around the financial situation at the individual or company level (Schuchter & Levi, 2016). Financial pressure forces managers to overstate revenue to artificially increase company performance to maintain operations and enjoy reward policies (Dorminey, et al., 2010). The second element of the fraud triangle is rationalization. Violators often do not consider themselves criminals and justify their actions with external factors. The third element of the fraud triangle involves knowledge and opportunity to commit fraud. Knowing internal control mechanisms and the possibility of detection plays a vital role in influencing the potential for individuals to commit fraud (Ramamoorti, et al., 2013; Rezaee & Riley, 2009). Regulators and standard setters have also evaluated the relevance of fraud theories in detecting financial reporting fraud. For example, the American Institute of Public Accountants has applied the fraud triangle in Statement on Auditing Standards No. 99-Consideration of Fraud in the Audit of Financial Statements. Cressey's Fraud Triangle (1950) is one of the most influential fraud theories. Its three essential elements of fraudulent behavior (pressure, rationalization, and opportunity) laid the foundations for a theoretical number based on the original structure. Based on the fraud triangle theory and its three main elements, researchers have developed models and criteria to evaluate whether a business exhibits signs of fraud in its financial statements. Besides the fundamental theories, some new research has also been developed based on the development of characteristics of fraudulent behavior. One cause of financial reporting fraud stems from information asymmetry. In business management, managers have better access to company information than outsiders, especially about the financial situation. Information asymmetry and fraud are closely linked, creating opportunities for them to commit fraud (Ndofor, et al., 2015).

Many recent studies have used the M-score model to identify businesses with signs of financial reporting fraud and confirm that this model is meaningful and reliable (Ashtiani & Raahemi, 2021; Craja, et al., 2020; Xiuguo & Shengyong, 2022). As Murdihardjo, et al. (2021) argue, the M-score method can effectively detect errors in financial reporting by creating three groups: manipulators, non-manipulators, and gray companies. Many other studies, such as Putri and Lestari (2021), Husnurrosyidah and Fatihah (2022), and Septiani, et al. (2020), also suggest that the M-score is worth using as a model to analyze financial statements about fraud in manufacturing companies. Today, the role of auditors is also vital in making comments on financial statements and helping stakeholders evaluate financial statements objectively. The auditor must detect fraud early to avoid fraud and prolonged scandals. The auditor can access the company's internal data and clearly understand important issues or errors. Based on auditing standards and experience, they can make objective and relatively accurate assessments of the reliability of financial statements. Auditing standards also provide frameworks to help auditors form opinions when auditing (SAS, 82). Auditors must be able to systematically and critically examine their tasks, specifically the financial statements prepared by the head of the company, along with accounting records and supporting evidence, to give a correct opinion on the completeness of the financial statements. As the importance of audit reports for companies becomes clear, auditors must be proficient in collecting and analyzing audit evidence (Hermawan, et al., 2021). The audit report is the auditor's statement about the accuracy of the presentation of the financial statements of the entity audited by the auditor (Pinto, et al., 2020). The audit report is the auditor's opinion on the truthfulness of presenting the audited organization/company's financial statements (Mazkiyani & Handoyo, 2017). However, recent studies have not seen any combination of considering the auditor's opinion with the factors in the fraud triangle model, nor combining it with machine learning algorithms to identify.

Today, with the development of technology, predicting or identifying enterprises with signs of financial reporting fraud is done more quickly and accurately. There are many methods used for forecasting, such as the logistic method (Beasley, 1996), KNN (El-Bannany, et al., 2021; Hajek & Henriques, 2017; Liu, et al., 2015), SVM (Cecchini, et al., 2010; El-Bannany, et al., 2021; Hajek & Henriques, 2017); Decision tree (El-Bannany, et al., 2021; Hajek & Henriques, 2017; Kirkos, et al., 2007; Lee, et al., 2022; Salehi & Fard, 2013) or artificial neural network-ANN (El-Bannany, et al., 2021; Green & Choi, 1997; Hajek & Henriques, 2017). Recently, there has been increasing interest in machine learning algorithms as more complex approaches tend to yield better results (West & Bhattacharya, 2016), typical of which are reinforcement learning methods, such as (Ali, et al., 2023) when identifying financial reporting fraud, the XGBoost algorithm was found to obtain the best result, with a final accuracy of 96.05%. Furthermore, most studies focus on one method and test many different tools (Abbasi, et al., 2012; Kassem, 2022; Liu, et al., 2015). Recent studies have mainly documented varying levels of evidence from both qualitative and quantitative sources on the effectiveness of some methods in detecting financial reporting fraud (Kapardis, et al., 2010; Kassem, 2022) while lacking an integrated assessment based on auditors' opinions, especially in comparison with previous methods (Makri & Neely, 2021). In this study, in addition to comparing the forecasting results based on the M-Score model with auditors' assessments of problematic financial statements, we also use supervised gradient boosting algorithms to examine the effectiveness of the evaluation.

The fraud triangle theory (Cressey, 1950; Schuchter & Levi, 2016) posits that financial reporting fraud arises from pressure, opportunity, and rationalization. Building on this foundation, researchers have developed models that operationalize these elements into measurable indicators. One of the most prominent is the Beneish M-Score model, which synthesizes accounting ratios to detect the likelihood of manipulation in financial statements (Beneish, et al., 2012). Numerous empirical studies confirm the effectiveness of the M-Score in identifying fraudulent reporting across different contexts (Ashtiani & Raahemi, 2021; Craja, et al., 2020; Husnurrosyidah & Fatihah, 2022; Murdihardjo, et al., 2021; Putri & Lestari, 2021; Septiani, et al., 2020; Xiuguo & Shengyong, 2022). Therefore, the first hypothesis is proposed:

> *H_1: The M-Score model is effective in detecting financial statement fraud among Vietnamese listed companies.*

Auditors play a central role in ensuring the credibility of financial reporting by issuing opinions based on direct access to firms' internal records and adherence to auditing standards (Mazkiyani & Handoyo, 2017; Pinto, et al., 2020; Ramos & Lyons, 1997). Their evaluations integrate both quantitative and qualitative aspects of financial statements, providing an alternative lens for identifying fraudulent practices. Despite their importance, prior research has rarely combined auditors' assessments with fraud triangle indicators or machine learning methods (Hermawan, et al., 2021; Kassem, 2022; Makri & Neely, 2021). This gap suggests a second hypothesis:

> *H_2: Auditor opinions provide significant predictive value in detecting fraudulent financial reporting.*

Recent advances in machine learning have further enhanced fraud detection capabilities, surpassing traditional statistical methods in predictive accuracy (Hajek & Henriques, 2017; West & Bhattacharya, 2016). Algorithms such as artificial neural networks, decision trees, and ensemble approaches like XGBoost and Random Forest have demonstrated high accuracy in identifying fraud (Ali, et al., 2023; El-Bannany, et al., 2021; Green & Choi, 1997; Liu, et al., 2015). By applying these methods to both the M-Score and auditors' opinions, the robustness of predictive models can be evaluated. This leads to the third hypothesis:

> *H_3: Machine learning techniques significantly improve the predictive accuracy of both the M-Score model and auditor opinion-based models in detecting financial statement fraud.*

3. Methodology

3.1 Data and Sample

This study collected data on variables that are required to calculate the M-score from financial reports of non-financial firms listed on the two stock exchanges, such as the Ho Chi Minh Stock Exchange (HSX) and the Hanoi Stock Exchange (HNX). In contrast, data on the variables required by the auditor to provide the auditor's opinion are collected from financial reports from 2018 to 2022. The financial firms, such as banks, insurance companies, and securities firms, are not considered in this study because their financial reporting characteristics differ significantly from those of non-financial firms. Financial institutions operate under distinct regulatory frameworks and accounting standards due to the nature of their business models, which are based on assets, liabilities, and regulatory capital requirements. These differences could affect the models' applicability for detecting fraud in financial reporting. By focusing on non-financial companies, we ensure the consistency and relevance of the applied fraud detection methods, such as the M-Score model, and use machine learning techniques, which are more suited to firms with traditional revenue, cost, and asset structures.

3.2 M-Score Model and the Auditor's Opinion for Fraud Detection

This study uses the M-Score model and the opinions of the auditors to identify the likelihood of manipulation in financial statements by the firms.

3.2.1 M-Score Model

The M-Score is a probabilistic model that identifies the firms that have manipulated their financial statements fraudulently (Beneish, et al., 2012). Firms with higher M-Scores are more likely to commit fraud. The mathematical model that estimates the M-Score is provided in Equation 1.

$$M-Score = -4.84 + 0.920 * \mathrm{DSRI} + 0.528 * \mathrm{GMI} + 0.404 * \mathrm{AQI} + 0.892 * \mathrm{SGI} + 0.115 * \mathrm{DEPI} - 0.172 * \mathrm{SGAI} + 4.679 * \mathrm{TATA} - 0.327 * \mathrm{LVGI}\,, \tag{1}$$

where DSRI is the Day Sales in Receivables Index, which shows the change in the average receivable period. An unusual increase in receivables may reflect discrepancies in how to record revenue. Increasing receivables helps improve business revenue and helps the executives of firms to achieve revenue targets. DSRI can be calculated using the following ratio:

$$DSRI = \frac{(\text{Accounts receivables}_t/\text{Sales}_t)}{(\text{Accounts receivables}_{t-1}/\text{Sales}_{t-1})}.$$

GMI is the Receivables Index, which indicates the decline in a company's profit margin over time. When a company is experiencing a decline in gross profit margins, more incentives are provided to the executives of the company to use such techniques that improve the financial statements. GMI can be calculated using the following ratio:

$$GMI = \frac{(Sales_{t-1} - COGS_{t-1})/Sales_{t-1}}{(Sales_t - COGS_t)/Sales_t}.$$

AQI is the Asset Quality Index, which measures the increase in long-term assets other than fixed assets essential for the business firm that may contribute to over-capitalization of the expenses. Suppose the company does not charge those costs in the current business period but capitalizes them as long-term assets and depreciates them over many years. In that case, it will significantly reduce the costs in that year. The profits may improve even though they do not come from the efficiency of its core business. AQI can be calculated using the following ratio:

$$AQI = \frac{(1 - [\text{Current Assets}_t + \text{PP\&E}_t + \text{Total Longterm Investments}_t])/\text{Total Assets}_t}{(1 - [\text{Current Assets}_{t-1} + \text{PP\&E}_{t-1} + \text{Total Longterm Investments}_{t-1}])/\text{Total Assets}_{t-1}}.$$

SGI is the Sales Index, which measures the rate of revenue growth. Although a company grows rapidly, it does not mean that it is likely to commit financial fraud. However, such companies' rapid growth and high capital requirements may put a lot of pressure on the operator. Therefore, the leader may be pressured to maintain a sufficiently attractive growth rate, which is difficult. SGI can be calculated using the following ratio:

$$SGI = \frac{Sales_t}{Sales_{t-1}}.$$

DEPI is the **Depreciation Index;** the company may seek to extend the depreciation period or use other financial tricks to record a reduction in depreciation expense during the period. When depreciation expense is reduced, the profits improve. DEPI can be calculated using the following ratio:

$$DEPI = \frac{(Depreciation_{t-1}/(PP\&E_{t-1} + Depreciation_{t-1}))}{(Depreciation_t/(PP\&E_t + Depreciation_t))}.$$

SGAI is the Selling, General, & Admin Expenses Index. Similar to the impact of declining gross profit margins, rising selling and administrative costs present a negative picture for investors. As a result, managers start committing financial fraud. SGAI can be calculated using the following ratio:

$$SGAI = \frac{(SG\&Q\ Expense_t / Sales_t)}{(SG\&A\ Expense_{t-1} / Sales_{t-1})}.$$

TATA is the Total Accruals to Total Assets Index, which is the Accumulated Return on Total Assets Ratio. In an accrual accounting system, the recognition of revenue and expenses differs from when cash flows into the business. The significant difference between pre-tax profit and net cash flow from operations raises many doubts among investors about the company's ability to use some unreasonable accounting policies. TATA can be calculated using the following ratio:

$$TATA = \frac{(\text{Income from Continuing Operations} - \text{Cash Flow from Operations})}{Total\ Asset}.$$

LVGI is the Leverage Index, which compares a company's total debt to its assets over time. Increasing debt to total assets worsens a company's financial ratios; therefore, companies try to cheat on their financial statements to ensure good credit metrics.

$$LVGI = \frac{[\text{Current Liabilities}_t + (\text{Total Long} - \text{term Debt})_t]/T\text{otal Assets}_t}{[\text{Current Liabilities}_{t-1} + (\text{Total Long} - \text{term Debt})_{t-1}]/T\text{otal Assets}_{t-1}}.$$

An M-score value greater than (-2.22) provides an indication of financial fraud in companies. Based on Equation 1, we assign the value = 1 (when M> -2.22) for firms with fraud and = 0 for firms without fraud (when M<= -2.22). The M-score that classifies whether a firm has fraudulent financial statements or not is called **Fraud1**.

3.2.2 Auditor's Opinion for Fraud Detection

An audit report is a document in which auditors share their views on the organization's financial performance and whether it complies with financial reporting regulations. Auditors must follow the format defined by generally accepted auditing standards (GAAS), with some exceptions depending on the nature of the audit. The audit report typically includes a description of the auditor's role, management's role, the scope of the audit, and the auditor's opinion. There are four types of auditor opinion reports. The auditor's opinion report is a letter that the auditors attach to the statutory audit report to reflect their opinion on the audit. These four types of audit opinions are Unqualified (Clean report), Qualified (Qualified report), Disclaimer (Disclaimer report), and Adverse (Adverse audit report). Financial reports include component reports; if a company has these component reports, the auditors evaluate them as Unqualified (Clean report), which means the financial statements are entirely accepted and informed, reflecting reliable information (assigned as 0). Conversely, when the auditor evaluates any component reports as not being

Unqualified, the auditor assigns the value as 1 (indicating problems with the financial statements). In this study, this variable is referred to as **Fraud2**.

Based on the auditor's opinion on the component financial statements, the variable **Fraud2** is constructed as follows:

$$\text{Fraud2} = \text{Balance Sheet} + \text{Profit Report} + \text{Cash Flows Statements} + \text{Financial Statements Footnotes}. \quad (2)$$

The variables used in Equation 2 are explained as follows: First, if the auditor assesses all component reports as Unqualified (entirely accepted), the company's financial statements are considered "clean" and assigned a zero value. On the other hand, if there is any component report that the auditor does not evaluate as Unqualified, the company's financial statements are considered "unclean" and are assigned a value of 1.

The selection of variables for Fraud2 — Balance Sheet, Profit Report, Cash Flow Statements, and Financial Statements Footnotes is based on their fundamental role in reflecting a company's financial position, performance, and cash flows, and providing essential disclosures. These components are critical in auditing, as they offer a comprehensive view of the financial statements' reliability and compliance with accounting standards. While additional variables may contribute to a more nuanced detection of fraudulent activities, these core elements are widely acknowledged as primary indicators for assessing financial reporting integrity. Future studies may consider incorporating further variables to enhance predictive accuracy.

In this study, Fraud2 is defined as a binary variable representing the auditor's overall opinion across the key financial statements. Specifically, if any component report (Balance Sheet, Profit and Loss Report, Cash Flow Statement, or Footnotes to Financial Statements) is assessed as anything other than "Unqualified," the company's financial statements are classified as fraudulent. Formally, this can be expressed as:

- **Fraud2 = 1** (the company is classified as having problematic financial statements) if **any one** of the following is true:
    - The auditor's opinion on the **Balance Sheet** is not "Unqualified".
    - OR the auditor's opinion on the **Profit Report** is not "Unqualified".
    - OR the auditor's opinion on the **Cash Flow Statement** is not "Unqualified".
    - OR the auditor's opinion on the **Footnotes** is not "Unqualified".
- **Fraud2 = 0** (the company is classified as having clean financial statements) only when **all four** opinions are "Unqualified."

This logical "OR" structure reflects the fact that even a single qualified or adverse component indicates potential problems with the company's financial reporting integrity.

3.3 Model

3.3.1 Machine Learning Techniques

ANN method: Artificial Neural Network (ANN) is a type of Artificial Intelligence (AI) that uses data mining techniques. ANN is a mathematical model based on biological neural networks. Unlike traditional systems, where knowledge is always subject to rules, an ANN creates its own rules by learning and practicing from given examples, implying that the method has been trained to pattern activity (Omar, et al., 2017). ANN is a tool that uses the same model, structure, and processing techniques as the human brain to analyze sample data repeatedly. This study has used this method to predict financial statement fraud (El-Bannany, et al., 2021; Green & Choi, 1997; Hajek & Henriques, 2017).

KNN method: K-nearest neighbor is one of the simplest (and effective in some cases) supervised learning algorithms in Machine Learning. When training, this algorithm does not learn anything from the training data (this is also why it is classified as lazy learning); all calculations are performed when it needs to predict new data results**.** K-nearest neighbor can be applied to both Supervised learning problems: Classification and Regression. KNN is also known as an Instance-based or Memory-based learning algorithm. KNN, an adaptable algorithm, determines the output of a new data point by utilizing the information of the K data points in the training set closest to it (K-neighbors), regardless of the noise in some of these closest data points. This method has been effectively used in predicting financial statement fraud, as evidenced by (El-Bannany, et al., 2021; Hajek & Henriques, 2017; Liu, et al., 2015).

Random Forest: One popular algorithm for supervised learning is the Random Forest algorithm. This can be applied to classification and regression problems. However, classification problems are the primary application for this algorithm. Like a forest typically composed of trees, the Random Forest algorithm builds decision trees using sample data and derives predictions from each data point. Thus, an ensemble approach is the Random Forest algorithm. Because this algorithm averages the result, it reduces over-fitting, making it superior to single decision trees (Sailusha, et al., 2020). This study has used this method in predicting financial statement fraud, such as (El-Bannany, et al., 2021; Hajek & Henriques, 2017; Liu, et al., 2015).

Decision Tree: The supervised machine learning family includes decision tree induction, which learns class-labeled training tuples and attempts to predict a class for a given input vector. A decision tree is a type of tree structure that resembles a flowchart. Usually, it begins with a single top node and branches out to form other nodes. Every one of those branched nodes branches off into different nodes, leading to more internal nodes that test an attribute. It eventually takes on the shape of a tree due to its continuous branching into new nodes. The decision tree will keep growing until each branch ends, indicating no more requirements to consider. The final labels or choices in the Decision Tree are represented by the endpoint, also known as the lead node (Lim, et al., 2021). This study has used this method to predict financial statement fraud (Liu, et al., 2015; West & Bhattacharya, 2016).

3.3.2 Gradient Boosting Algorithms

XGBoost: XGBoost, or eXtreme Gradient Boosting, is a tree-based algorithm (Chen & Guestrin, 2016). XGBoost shows promising results in both performance and speed due to its bias and variance reduction strategy. XGBoost sequentially builds weak trees, misclassifying data from previous trees into new trees based on adjusted data weights. XGBoost uses a "weighted percentile sketching algorithm" that helps the classifier focus on incorrectly classified data.

LightGBM: LightGBM trains a gradient-boosted decision tree (GBDT), but it also supports random forests, Dropout Multiple Additive Regression Trees (DART), and Gradient One-Side Sampling (Goss). In some cases, LightGBM can produce more accurate and faster results than XGBoost. Both LightGBM and XGBoost are widely used and provide highly optimized, scalable, and fast implementations of the gradient-boosted machine learning (GBM) algorithm (Quinto, 2020)

3.4 Preliminary Diagnostics

Table 1 provides a detailed explanation of the number of firms selected from the HSX and HNX stock exchanges of Vietnam for the period of five years from 2018 to 2022, for which the consolidated financial statements have been audited. It also provides details on auditors' opinions on the financial statements of these firms.

Table 1. Number of companies studied

Stock exchanges	**2018**	**2019**	**2020**	**2021**	**2022**
HSX	304	316	332	343	349
HNX	288	294	298	304	309
Total	592	610	630	647	658
Total listed companies	729	731	756	761	758
Sampling rate	81.2%	83.4%	83.3%	85.02%	86.8%
Proportion of positive samples according to Fraud1	28.71% (170/592)	31.63% (193/610)	29.36% (185/630)	31.37% (203/647)	28.87% (190/658)
Proportion of positive samples according to Fraud2	7.23% (43/592)	7.21% (44/610)	6.98% (44/630)	6.96% (45/647)	7.15% (47/658)

Note: "Positive samples" represent companies classified as having signs of financial reporting fraud. For Fraud1, a positive sample is defined as a firm with an M-Score greater than –2.22 (Model 1). For Fraud2, a positive sample is defined as a firm whose financial statements received at least one auditor opinion other than "Unqualified" (Model 2). Source: Author's compilation

The listed firms in HSX and HNX stock exchanges are (729 in 2018), (731 in 2019), (756 in 2020), (761 in 2021), and (758 in 2022). The percentage of firms selected as a sample to study in each year varies from 81.2% in 2018 (the lowest) to 86.8% in 2022 (the highest), which is representative of the overall population. Additionally, the table exhibits variations of companies with and without fraudulent financial statements, particularly with classification based on the auditor's assessment (Fraud2). The proportion of companies with financial statements flagged by auditors as problematic only represents a small percentage (6.96% to 7.23%). This significant variation could lead to falsely positive evaluation results. To address this issue, this study employs the oversampling technique. This involves increasing the sample size in the low group until the classification samples achieve a balanced 50:50 ratio; the process is discontinued at this point.

4. Results and Discussion

To get the best forecasting results, this study first tests the overfitting problem of the dataset; the test results are shown in several methods, including M-Score and the Auditor's opinion.

Figure 1. Overfit test results of simpler models (Decision Tree, Random Forest, and KNN)

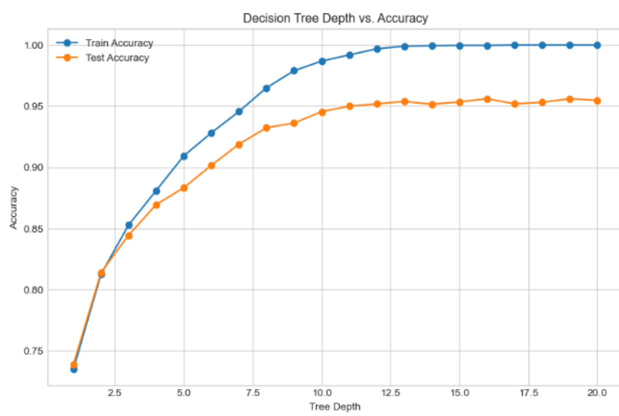
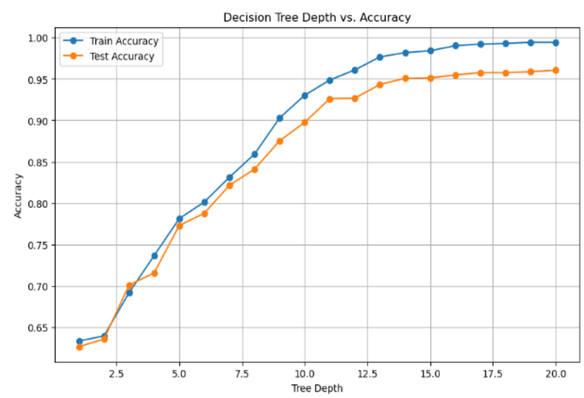
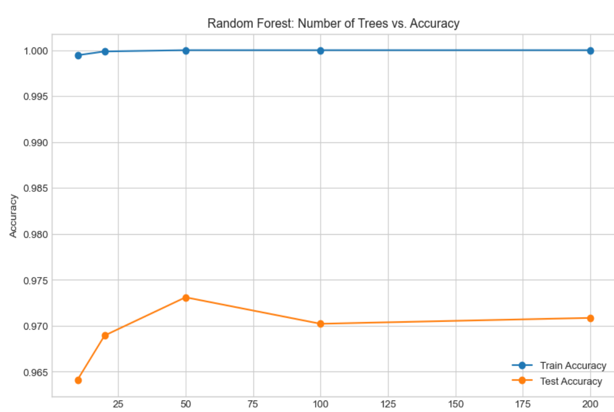
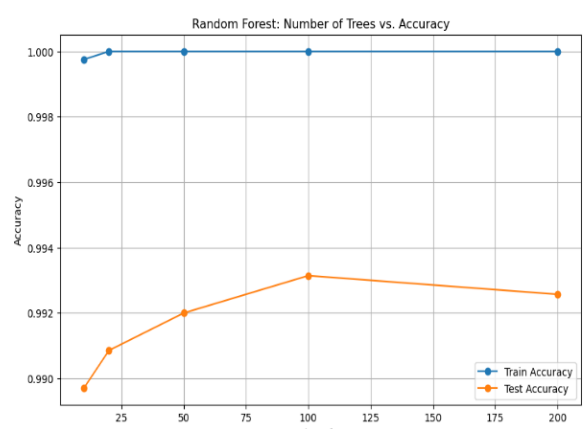
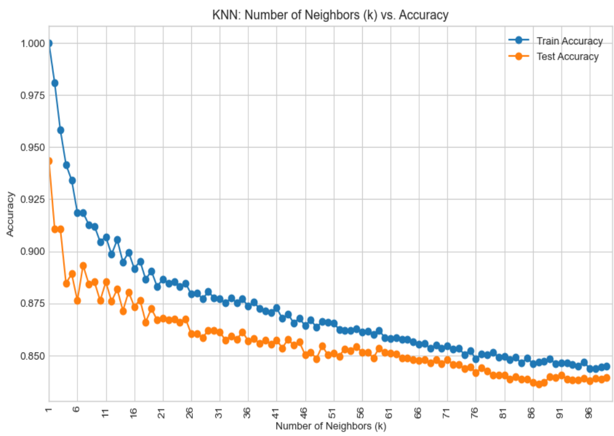
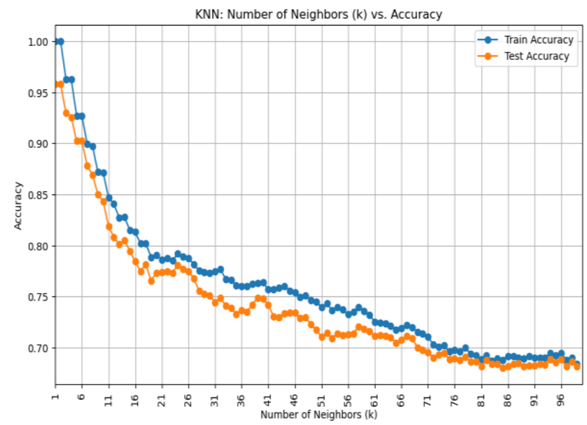
Method	**Test Overfit with Model (1)-Fraud1 (M-Score)**	**Test Overfit with Model (2)-Fraud2 (Auditor's opinion**
Decision Tree		
Random Forest		

Figure 1. Overfit test results of simpler models (Decision Tree, Random Forest, and KNN)

Method	Test Overfit with Model (1)-Fraud1 (M-Score)	Test Overfit with Model (2)-Fraud2 (Auditor's opinion
KNN		

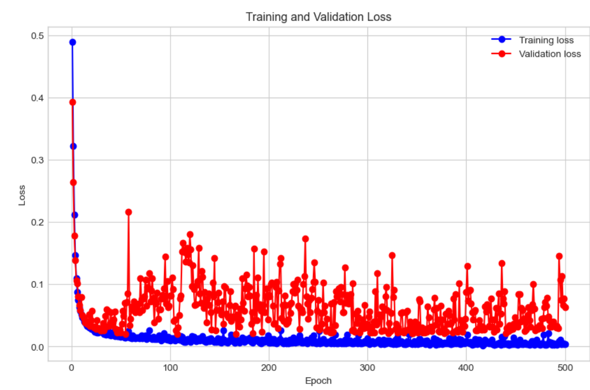
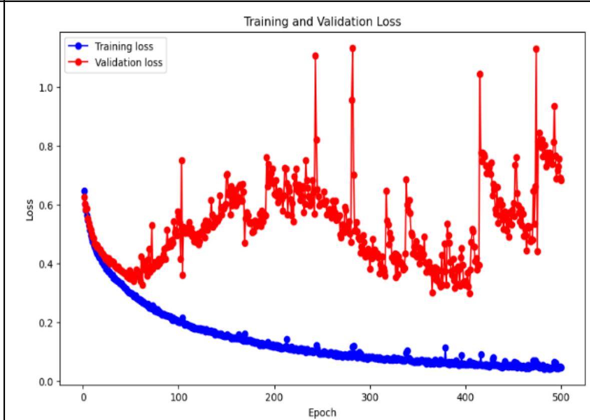
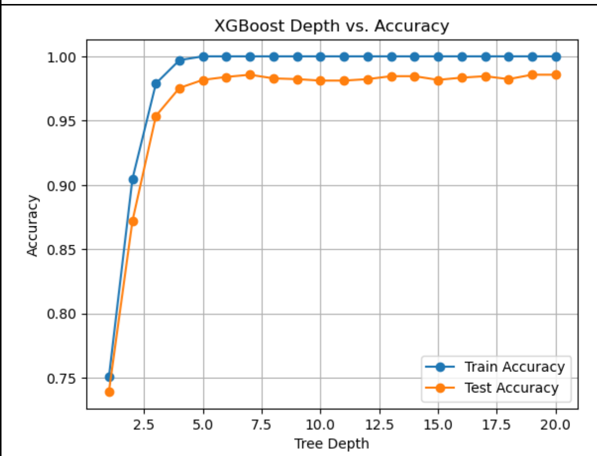
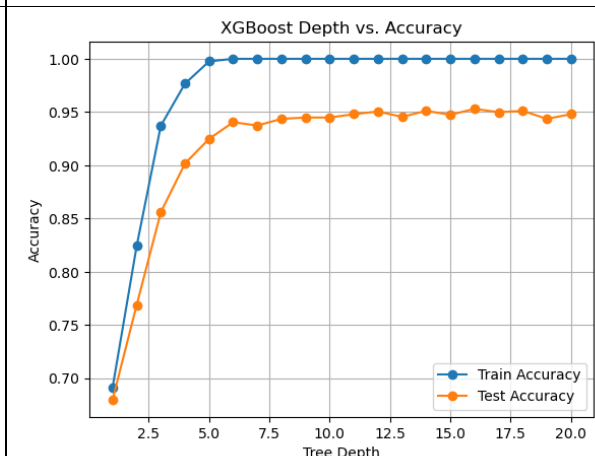
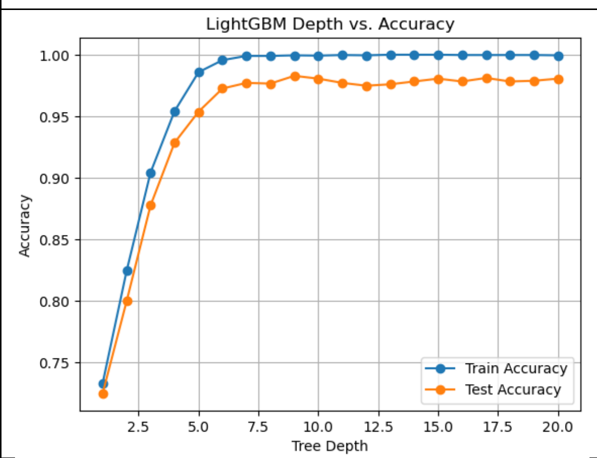
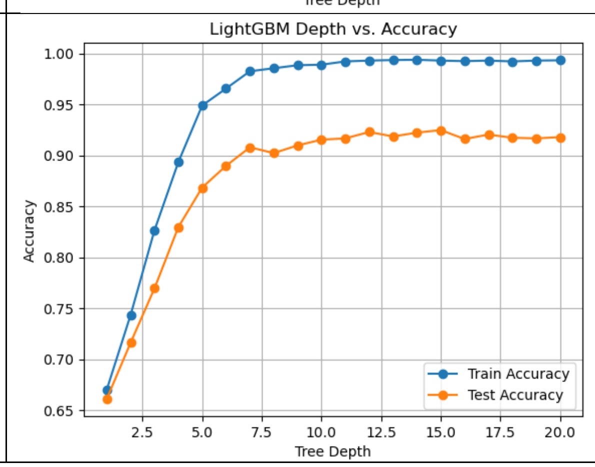
Source: Authors' calculations

In Figure 1, for the Decision Tree method, when forecasting with Fraud1, there is a convergence between train and test accuracy at a Tree Depth of five. This suggests that selecting a Tree Depth of five prevents overfitting. Therefore, calibrating the Decision tree model at a tree depth of five is necessary. Similarly, when forecasting with Fraud2, convergence between train and test accuracy is observed for Tree Depth values around five and seven. This indicates that selecting a Tree Depth within this range prevents overfitting. Therefore, adjusting the Decision tree method for the Fraud2 variable at a tree depth of seven is needed. For two supervised gradient boosting algorithms (**XGBoost and LightGBM)**, choosing a tree depth of five is also optimal.

For forecasting using Random Forest in both models, it shows that there is no convergence between train and test, so severe Overfitting occurs. This method is not suitable for forecasting for either model. Although the Random Forest model reports very high average accuracy and F1-scores for both Fraud1 and Fraud2, further diagnostic analysis reveals clear signs of overfitting. Specifically, the training and validation accuracy curves do not converge across different tree depths and numbers of estimators, indicating unstable generalization behavior. This phenomenon is likely amplified by the use of oversampling techniques, which can lead Random Forest to memorize resampled minority observations rather than learn robust decision boundaries. As a result, the high accuracy values reported in Table 3 should be interpreted with caution, as they may reflect model complexity and data resampling effects rather than genuine predictive reliability. Given the objective of fraud detection as an early-warning and decision-support tool, this study prioritizes models that demonstrate both strong performance and stable generalization. Therefore, despite its high apparent accuracy, Random Forest is not selected as the preferred forecasting model in this context.

For the KNN method, when forecasting using Fraud1, which shows that the value k=20 is reasonable, it is necessary to adjust KNN with Fraud1 to the value k-20. As for predicting using Fraud2, the k value is from 20 to 23, showing that KNN is stable and converges between train & test accuracy. Therefore, the test adjusts the model to the value k=20.

Figure 2. Overfit test results of complex models (ANN, XGBoost, and LightGBM)

Method	Test Overfit with Model (1)-Fraud1 (M-Score)	Test Overfit with Model (2)-Fraud2 (Auditor's opinion
ANN		
XGBoost		
LightGBM		

Source: Authors' calculations

Figure 2 shows the results of testing the Overfit phenomenon of the methods through the two models (1) and (2) above. For the ANN method, when forecasting with Fraud1 (Model 1), it is observed that at epoch 50, the validation loss starts to fluctuate strongly. This indicates that overfitting will occur if the epoch exceeds 50, resulting in poor forecasting results. Therefore, it is necessary to adjust the ANN method in Model 1 at epoch 50. Similarly, when forecasting with Fraud2 using ANN, it is noted that from around epoch 80, the overfitting phenomenon occurs as the validation loss tends to increase and fluctuate strongly. Thus, it is essential to fine-tune the ANN method with the Fraud2 variable at epoch 80.

After correcting the overfitting problem, we reprocess the methods according to two prediction models and give the results in Table 3.

Table 3. Summary of the results of forecasting indicators according to different methods

Method		Average accuracy	Average precision	Average recall	Average F1-score	Average time
KNN	Fraud2	0.781	0.732	0.870	0.795	1.345
	Fraud1	0.885	0.873	0.896	0.885	2.095
Decision Tree	**Fraud2**	**0.821**	**0.773**	**0.897**	**0.831**	**0.610**
	Fraud1	0.883	0.871	0.895	0.883	0.516
Random Forest	Fraud2	0.991	0.983	1.000	0.991	26.417
	Fraud1	0.970	0.946	0.996	0.970	24.713
ANN	Fraud2	0.718	0.698	0.745	0.721	12.274
	Fraud1	**0.979**	**0.966**	**0.992**	**0.979**	**34.173**
Naive Bayes	Fraud2	0.511	0.500	0.951	0.655	0.157
	Fraud1	0.606	0.937	0.213	0.346	0.094
SVM (sigmoid kernel)	Fraud2	0.579	0.568	0.575	0.572	10.444
	Fraud1	0.710	0.702	0.713	0.707	19.357
XGBoost	Fraud2	**0.940**	**0.906**	**0.981**	**0.942**	**0.682**
	Fraud1	**0.984**	**0.968**	**1.000**	**0.984**	**0.689**
LightGBM	Fraud2	0.915	0.877	0.962	0.917	0.540
	Fraud1	0.978	0.957	1.000	0.978	0.650

Note: Accuracy measures the proportion of correctly predicted cases (both fraudulent and non-fraudulent). Precision measures the proportion of predicted fraudulent cases that are truly fraudulent. Recall measures the proportion of actual fraudulent cases that were correctly identified. The F1-score is the harmonic mean of Precision and Recall, providing a balanced evaluation metric in the presence of imbalanced data. Each row of Table 3 reports performance results for either Fraud1 (M-Score-based model, Model 1) or Fraud2 (Auditor opinion-based model, Model 2), as indicated in the first column. Source: Authors' calculations

Accuracy and F1 score are two reliability indicators used to evaluate the performance of a classification model. When constructing a classification model, assessing the proportion of correctly predicted cases relative to the total number of cases is crucial, termed accuracy. Accuracy serves as a metric for evaluating the predictive performance of a model on a given dataset, where higher accuracy signifies a more precise model. While the F1-score also measures the model's accuracy, it computes the average between precision and recall, making it more representative in assessing accuracy across both precision and recall. Additionally, the F1-score only calculates accuracy within the positive sample group, typically the group of greater interest. Thus, in the case of imbalanced data, it's more appropriate for accuracy to be computed on both positive and negative samples. Given the imbalanced nature of the data and the focus on predicting positive values (as indicated in Table 1), this study evaluates the methods using Accuracy and F1-score.

As stated above, the Random Forest method is unsuitable for this study because it violates the overfitting phenomenon, but it cannot be overcome. All three remaining methods have the potential to be used to predict financial statement fraud. This result is consistent with research results using ANN (Green & Choi, 1997); KNN (Liu, et al., 2015); and Decision Tree (Kirkos, et al., 2007; Salehi & Fard, 2013) to predict whether businesses have signs of financial reporting fraud.

For forecasts based on Fraud1 (according to the M-Score model), the highest Accuracy and F1-score results are 97.9%, occurring in the ANN method. This shows that if predicting financial statement fraud according to the M-score, you should use ANN; this result is also consistent with the research of Husnurrosyidah and Fatihah (2022), Murdihardjo, et al. (2021), and Phong, et al. (2022). In the forecasting model based on the auditor's opinion for classification, the forecasting indexes are also over 80%, showing that the forecasting ability is also quite good. Still, the values are lower than the M-model Score. Specifically, the accuracy value reached 0.821, while the F1-score value reached 0.831, and these were the highest values that occurred in the decision tree method. These results agree with using the Decision Tree method in forecasting (El-Bannany, et al., 2021; Hajek & Henriques, 2017; Lee, et al., 2022). The results of supervised gradient boosting algorithms show that the XGBoost method gives the best results and has higher performance than conventional learning methods. The supervised gradient boosting algorithms' results also support the M-Score fraud prediction model (F1-score index reaches 0.984). This result is quite similar to the study of Ali, et al. (2023). Still, the prediction based on auditors' opinions also gives reliable results worthy of attention for verification (F1-score index reaches 0.942).

These values are lower than the M-Score because the results depend heavily on input variables when using Cressey's (1950) fraud triangle model. Accordingly, this model is based on three pillars, which are three factors causing financial reporting fraud: pressure, opportunity, and rationality. However, choosing which indices to represent the above three groups of factors is still controversial and inconsistent. It depends heavily on regulations on the financial reporting system and the characteristics of the financial markets of countries. In recent years, the rise of virtual assets and digital financial instruments has introduced new fraud schemes that are not fully captured by traditional models, such as Cressey’s (1950) fraud triangle. While this study relies on the established variables of pressure, opportunity, and rationalization, future research could benefit from incorporating more contemporary factors related to virtual assets and digital fraud. This would allow for a more comprehensive approach, reflecting the evolving nature of financial crimes in today’s digital economy.

To identify effective models for predicting financial statement fraud with high accuracy, this study presents two forecasting methods that demonstrate notable accuracy. By conducting empirical surveys to explore additional input variables and establishing a robust database as a foundation for forecasting, the study seeks to enhance prediction accuracy further. Continuous reporting is an effective tool to provide investors with early warnings, aiding in the prevention and mitigation of risks while enabling state management agencies to closely grasp information to monitor businesses effectively and mitigate risks to the market.

The results provide strong support for H_1. The M-Score model demonstrates high predictive effectiveness in detecting financial statement fraud among Vietnamese listed firms, particularly when combined with advanced machine learning techniques such as ANN and XGBoost. The consistently high accuracy and F1-scores indicate that the financial ratios embedded in the M-Score successfully capture fraud-related patterns in the data. Also, H_2 is supported, as auditor opinions exhibit significant predictive value in identifying problematic financial statements. Although the predictive performance of the auditor-opinion-based model is lower than that of the M-Score-based model, the results remain robust, particularly when decision tree and gradient boosting algorithms are applied. This finding highlights the value of auditors' qualitative assessments as an alternative and complementary fraud detection mechanism. The empirical results strongly support H_3. Machine learning techniques substantially enhance the predictive accuracy of both M-Score-based and auditor-opinion-based models. In particular, gradient boosting algorithms such as XGBoost and LightGBM consistently outperform traditional classifiers, confirming the advantages of ensemble learning in fraud detection tasks characterized by nonlinearity and data imbalance.

5. Conclusions and Recommendations

This research aims to consider identifying businesses with signs of fraudulent financial statements by comparing the M-Score measure through the M-Score model of Beneish, et al. (2012), based on the auditors' opinion in evaluating the classification of financial statements, combined with input factors that can cause financial statement fraud through the fraud triangle model. The results indicate that the M-Score model continues to provide better predictive value, with the best forecasting outcomes achieved through the ANN method. However, the auditor's opinion also has good results in identifying businesses with signs of financial statement fraud using the Decision Tree method. Suppose further in-depth research is conducted to explore various indicators and identify those that better represent the three groups of factors in the fraud triangle model. In that case, the prediction model will rely on expert opinions. The auditor's opinion will serve as a valuable forecasting model, merging qualitative insights with quantitative analysis through the fraud triangle model enhanced by modern identification methods using machine learning. To optimize the problem of predicting financial statement fraud in future studies, we recommend using auditors' opinions as a reference variable in such forecasting models that will ensure more authenticity in practice. If supervised gradient boosting algorithms such as the XGBoost technique are used, the forecasting results are effective on both the M-Score and based on the auditor's opinion.

Based on the research findings, this study aims to develop this outcome into a platform for early and timely fraud detection in financial reporting. This platform is not just a tool but a powerful assistant, empowering investors to formulate effective stock selections. It also enables policy management agencies to implement timely intervention measures to mitigate the risk of market collapse and prevent domino effects that could lead to crises. This is a decision-making platform based on reliable financial reporting information for managers and investors to refer to.

This study has several limitations that should be acknowledged. First, the identification of indicators representing the three components of the fraud triangle model, as pressure, opportunity, and rationalization, remains a challenge. The selection of variables is not fully standardized and may vary across regulatory environments and market characteristics, which can affect the consistency of results (Cressey, 1950; Schuchter & Levi, 2016). Second, the research is limited to non-financial companies listed on the Ho Chi Minh City (HSX) and Hanoi (HNX) stock exchanges from 2018 to 2022. While this provides valuable insights into the Vietnamese context, the generalizability of findings to financial institutions or other emerging markets may be constrained.

Future research could address these limitations in several ways. One promising direction is to expand the range of input variables by incorporating emerging risk factors related to digital assets and virtual financial instruments, which are increasingly linked to sophisticated fraud schemes and are not captured by traditional fraud models. In addition, researchers may further refine the operationalization of the fraud triangle by empirically testing alternative indicators and validating them across different industries and institutional settings. Finally, the study's findings provide a foundation for developing an integrated decision-making platform that combines M-Score metrics, auditor opinions, and machine learning algorithms. Such a platform could serve as a real-time early warning system to support investors, managers, auditors, and regulators in detecting fraud more effectively and in making timely decisions under uncertainty.

Acknowledgment

This research is funded by Vietnam National University, Ho Chi Minh City (VNU-HCM) under grant number DS2022-34-03/HĐ-KHCN

References

Abbasi, A., Albrecht, C. C., Vance, A., & Hansen, J. (2012). MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud. *MIS Quarterly*, 36(4), 1293–1327

Ali, A. A., Khedr, A. M., El-Bannany, M., & Kanakkayil, S. (2023). A powerful predicting model for financial statement fraud based on optimized XGBoost ensemble learning technique. *Applied Sciences*, 13(4), 2272.

Altman, E. I. (2013). Predicting financial distress of companies: revisiting the Z-score and ZETA® models. In Handbook of research methods and applications in empirical finance (pp. 428-456). Edward Elgar Publishing.

Ashtiani, M. N., & Raahemi, B. (2021). Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review. *IEEE Access*, 10, 72504-72525.

Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *Accounting Review*, 443–465.

Beneish, M. D., Lee, C., & Nichols, D. C. (2012). Fraud detection and expected returns. *Available at SSRN 1998387*.

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Making words work: Using financial text as a predictor of financial events. *Decision support systems*, *50*(1), 164–175.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).

Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decision Support Systems*, 139, 113421.

Cressey, D. R. (1950). The Criminal Violation of Financial Trust. *American Sociological Review*, 15(6), 738–743.

Cressey, D. R. (1953). Other people's money: a study of the social psychology of embezzlement.

Dorminey, J. W., Fleming, A. S., Kranacher, M. J., & Riley Jr, R. A. (2010). Beyond the fraud triangle. *The CPA Journal*, *80*(7), 17.

El-Bannany, M., Dehghan, A. H., & Khedr, A. M. (2021, March). Prediction of financial statement fraud using machine learning techniques in the UAE. In *2021, 18th International Multi-Conference on Systems, Signals & Devices (SSD)* (pp. 649–654). IEEE.

Green, B. P., & Choi, J. H. (1997). Assessing the Risk of Management Fraud Through Neural Network Technology. *Auditing: A Journal of Practice & Theory*, 16(1), 14–28

Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud–A comparative study of machine learning methods. *Knowledge-Based Systems*, *128*, 139-152.

Hall, J. A. (2007). *Accounting Information Systems. Fifth Edition.* Cincinnati: Thomson Southwestern College Publishing.

Healy, P. M., & Palepu, K. G. (2003). The Fall of Enron: *Journal of Economics Perspectives*. *Volume*, *17*, 13.

Hermawan, S., Rahayu, D., Biduri, S., Rahayu, R. A., & Salisa, N. A. N. (2021). Determining Audit Quality in the Accounting Profession with Audit Ethics as a Moderating Variable. *Indonesian Journal of Sustainability Accounting and Management*, *5*(1), 11-22.

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision support systems*, 50(3), 585–594.

Husnurrosyidah, H., & Fatihah, I. (2022). Fraud Detecting Using Beneish M-Score and F-Score: Which is More Effective. *EQUILIBRIUM: Jurnal Ekonomi Syariah*, 10(1), 137-151.

Kamarudin, K. A., Ismail, W. A. W., & Mustapha, W. A. H. W. (2012). Aggressive financial reporting and corporate fraud. *Procedia-Social and Behavioral Sciences*, 65, 638–643.

Kapardis, M. K., Christodoulou, C., & Agathocleous, M. (2010). Neural networks: The panacea in fraud detection? *Managerial Auditing Journal*, 25 (7): 659-678.

Kassem, R. (2022). How could external auditors assess the rationalization of fraud?. *Journal of Financial Crime*, *29*(4), 1458–1467.

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert systems with applications*, 32(4), 995–1003.

Lee, C. S., Cheang, P. Y. S., & Moslehpour, M. (2022). Predictive analytics in business analytics: decision tree. *Advances in Decision Sciences*, 26(1), 1-29.

Lim, K. S., Lee, L. H., & Sim, Y. W. (2021). A review of machine learning algorithms for fraud detection in credit card transaction. *International Journal of Computer Science & Network Security*, *21*(9), 31–40.

Liu, C., Chan, Y., Alam Kazmi, S. H., & Fu, H. (2015). Financial fraud detection model: Based on random forest. *International journal of economics and finance*, *7*(7), 178-188.

Makri, C., & Neely, A. (2021). Grounded theory: A guide for exploratory studies in management research. *International Journal of Qualitative Methods*, *20*, 16094069211013654.

Mazkiyani, N., & Handoyo, S. (2017). Audit report lag of listed companies in Indonesia stock exchange. *Jurnal Aplikasi Bisnis*, 77-95.

Murdihardjo, L., Nurjanah, Y., & Sari, F. I. (2021). Penggunaan Metode Beneish Ratio Dalam Pendeteksian Kecurangan Laporan Keuangan. *Jurnal Akuntansi*, *10*(1), 179-194.

Ndofor, H. A., Wesley, C., & Priem, R. L. (2015). Providing CEOs with opportunities to cheat: The effects of complexity-based information asymmetries on financial reporting fraud. *Journal of Management*, *41*(6), 1774-1797.

Omar, N., Johari, Z. A., & Smith, M. (2017). Predicting fraudulent financial reporting using artificial neural network. *Journal of Financial Crime*, *24*(2), 362–387.

Phong, N. A., Tam, P. H., & Thanh, N. P. (2022, January). Fraud identification of financial statements by machine learning technology: case of listed companies in Vietnam. In *International Econometric Conference of Vietnam* (pp. 425-436). Cham: Springer International Publishing.

Pinto, I., Morais, A. I., & Quick, R. (2020). The impact of the precision of accounting standards on the expanded auditor's report in the European Union. *Journal of International Accounting, Auditing and Taxation*, *40*, 100333.

Putri, N., & Lestari, I. P. (2021). Analisis Determinan Financial Statement Fraudulent Dengan Model Beneish M-Score (Studi Pada Perusahaan Manufaktur Yang Terdaftar Di Bursa Efek Indonesia Tahun 2016–2018). *Jurnal Ilmiah Ekonomi Bisnis*, *26*(1), 69-85.

Quinto, B. (2020). *Next-generation machine learning with spark: Covers XGBoost, LightGBM, Spark NLP, distributed deep learning with keras, and more*. Apress.

Ramamoorti, S., Morrison III, D. E., Koletar, J. W., & Pope, K. R. (2013). *ABC's of behavioral forensics: applying psychology to financial fraud prevention and detection*. John Wiley & Sons.

Ramos, M. J., & Lyons, A. M. (1997). Considering fraud in a financial statement audit: practical guidance for applying SAS No. 82.

Rezaee, Z., & Riley, R. (2009). *Financial Statement Fraud: Prevention and Detection (2nd edition)*. Hoboken, NJ: John Wiley & Sons

Sailusha, R., Gnaneswar, V., Ramesh, R., & Rao, G. R. (2020, May). Credit card fraud detection using machine learning. In *2020, the 4th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1264–1270). IEEE.

Salehi, M., & Fard, F. Z. (2013). Data mining approach to prediction of going concern using classification and regression tree (CART). *Global Journal of Management and Business Research*, *13*(3), 25–29.

Schuchter, A., & Levi, M. (2016). The Fraud Triangle revisited. Security Journal, 29(2), 107–121.

Septiani, R., Musyarofah, S., & Yuliana, R. (2020). Beneish M-Score Reliability as a Tool For Detecting Financial Statements Fraud. In *International Colloquium Forensics Accounting and Governance (ICFAG)* (Vol. 1, No. 1, pp. 140-149).

Singleton, T. W., Singleton, A. J., Bologna, G. J., & Lindquist, R. J. (2006). *Fraud auditing and forensic accounting*. John Wiley & Sons.

Vaassen, E. H. J. (2004). Accounting Information systems, a Managerial Approach, John Wiley & Sons

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. *Computers & security*, *57*, 47–66.

Xiuguo, W., & Shengyong, D. (2022). An analysis on financial statement fraud detection for Chinese listed companies using deep learning. *IEEE Access*, 10, 22516-22532.

Yeh, C. C., Chi, D. J., & Hsu, M. F. (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37(2), 1535–1541.