Journal of Investigative Auditing & Financial Crime

E-ISSN: 3090-0093
P-ISSN: 3090-0107

Research Article

Application of Machine Learning for Fraud Detection in Corporate Annual Financial Reports

M. Syahrudin [1,*], Jusra Tampubolon [2], Fuad Fazil Osmanov 3

[1] Sekolah Tinggi Ilmu Ekonomi Gema Widya, Indonesia
[2] Universitas Prima Indonesia, Indonesia
[3] Vizja University, Poland
* Corresponding Author :syahrudin.ssh@gmail.com

Abstract: This study investigates the application of supervised learning algorithms for detecting financial statement fraud in annual corporate reports. Financial reporting fraud remains a critical challenge for auditors and regulators, as traditional detection methods, such as ratio analysis and manual auditing, often fail to identify complex anomalies in large datasets. The research aims to evaluate the effectiveness of several machine learning algorithms in improving fraud detection accuracy and reliability. A dataset consisting of 500 annual financial statements from 2020 to 2024, including 50 identified cases of potential fraud, was preprocessed through data cleaning, normalization, and labeling. Algorithms tested include Decision Tree, Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors (K-NN), and Neural Network. The results indicate that Neural Network achieves the highest accuracy (94.5%), followed by SVM (91.6%), while simpler algorithms such as Naïve Bayes and K-NN demonstrate moderate performance. Comparative analysis highlights that ensemble and deep learning models are more capable of capturing complex patterns in financial data, providing a significant advantage over traditional methods. The findings suggest that integrating machine learning into auditing practices can enhance the detection of fraudulent activities, improve decision-making processes, and increase the reliability of audit outcomes. This research underscores the importance of combining advanced computational techniques with professional auditor oversight to ensure accuracy, transparency, and accountability in financial reporting.

Keywords: Auditing; Financial Fraud Detection; Machine Learning; Neural Network; Supervised Learning.

1. Introduction

Detecting financial statement manipulation remains one of the major challenges in modern auditing practices. Financial transparency is not only essential for shareholders and investors but also forms the foundation of public trust in corporate governance. Companies are expected to present accurate and reliable financial information as the basis for decision-making, investment allocation, and regulatory compliance. However, despite regulatory frameworks and standardized auditing procedures, fraudulent practices in financial reporting continue to occur and often go undetected until they have caused significant damage to stakeholders. In reality, traditional audit methods, which are still widely applied, face significant limitations. First, manual audit processes are relatively slow in identifying fraud, as they rely heavily on conventional inspection procedures that are time-consuming and labor-intensive. This often results in fraud detection occurring only after material losses have already taken place, which weakens the role of auditing as a preventive mechanism. Second, traditional audits are relatively costly and prone to human error, whether due to limited auditor capacity, subjective judgment, or oversight in complex reporting structures, which ultimately reduces the accuracy of audit outcomes. Moreover, audit data are generally imbalanced because the number of valid and fair financial reports significantly outweighs fraudulent ones. This imbalance creates an additional challenge for auditors, as fraudulent patterns are often hidden within large volumes of normal data, making their identification more complex.

Received: November 07, 2024
Revised: January 25, 2025
Accepted: March 08, 2025
Published: May 31, 2025
Curr. Ver.: May 31, 2025

Along with rapid technological advancements in data science and computational methods, machine learning (ML) has increasingly been applied in auditing to address the shortcomings of traditional methods. Literature highlights that ML can enhance fraud detection accuracy compared to conventional approaches that rely on manual analysis or financial ratios. Unlike traditional models, ML techniques are capable of processing large-scale financial data, identifying hidden correlations, and continuously improving through iterative training. Algorithmic models such as Random Forest, XGBoost, and ensemble methods have proven to be more reliable in identifying anomalies within complex financial datasets ,. Furthermore, deep learning approaches such as the Deep Q-Network (DQN) have demonstrated highly promising results, achieving detection accuracy rates of up to 93.42% in certain financial fraud cases. These developments show that ML can serve as an effective tool not only for detection but also for early prevention of fraud in financial reporting.

The primary advantage of ML lies in its ability to recognize non-linear patterns and complex relationships that are difficult to capture through human analysis or traditional ratio-based methods. Advanced algorithms, including Biased Penalty SVM and data resampling techniques, have proven effective in addressing the data imbalance problem that often hinders fraud detection studies. In addition, ML enables the processing of multidimensional datasets by integrating information such as inter-firm relationships, supply chains, and industry linkages. For instance, graph convolutional networks (FraudGCN) extend analytical capacity by leveraging entity relationships, thereby improving fraud detection performance and offering broader perspectives compared to single-firm analysis.

Beyond numerical data, prior studies have also emphasized the importance of utilizing non-financial information, such as annual report narratives, managerial statements, and auditor notes. Through natural language processing (NLP) and data mining techniques, auditors can gain additional insights that are otherwise inaccessible to number-based methods ,. This integration of structured financial data with unstructured textual data represents a significant leap in fraud detection. It marks a clear distinction from traditional ratio analysis, which focuses exclusively on quantitative data and often fails to capture subtler signs of manipulation embedded in qualitative disclosures.

In addition to fraud detection, ML has also been applied to predicting audit opinions. Models based on Support Vector Machines (SVM), Naive Bayes, and Artificial Neural Networks (ANN) have shown great potential in reducing misclassification errors particularly type II errors, which risk overlooking fraudulent activity while improving the overall quality of audit opinions ,. Recent studies comparing statistical approaches with ML-based models further confirm that ML consistently outperforms traditional methods in projecting audit opinions. This suggests that the role of ML is not only limited to fraud identification but also extends to enhancing audit quality and supporting the assurance function of financial reporting.

Nevertheless, the application of ML in auditing is not without challenges. Key issues include data privacy concerns, the potential for algorithmic bias that may distort classification results, and the ongoing need for human oversight to ensure that findings remain accurate and interpretable. Furthermore, integrating ML into auditing practices requires significant organizational readiness, including access to high-quality data, computational resources, and auditor training in data analytics. Therefore, combining traditional auditing techniques with ML-based approaches is crucial. Such synergy is expected to create audit systems that are faster, more accurate, and more reliable, ultimately safeguarding financial reporting transparency and integrity in today's increasingly complex digital era.

2. Literature Review

2.1 Fraud Detection in Financial Reporting

Fraud in financial reporting remains a serious challenge for both auditors and regulators. Manipulative practices in corporate financial statements can erode public trust, harm investors, and generate systemic risks in capital markets. Traditional detection methods, such as financial ratio analysis and substantive testing, have proven insufficient in capturing increasingly complex patterns of irregularities. Furthermore, manual audits require significant resources, making them prone to delays in detecting major fraud cases. In many situations, fraudulent statements are only uncovered after they have caused material losses and triggered

financial crises. These limitations have encouraged researchers to explore intelligent technologies such as machine learning (ML) as a more adaptive and accurate alternative.

2.2 Machine Learning in Accounting Information Systems

ML has rapidly emerged as one of the most promising solutions to strengthen the functions of accounting information systems. With its ability to process large and complex datasets, ML can detect hidden patterns and provide faster, more accurate predictions. In practice, ML is often integrated with business information systems to support more objective decision-making. Popular algorithms such as Random Forest, Gradient Boosting, and XGBoost have been shown to outperform conventional methods in detecting anomalies in financial data. Moreover, ensemble methods often deliver higher accuracy compared to single-model approaches. This demonstrates that the integration of BIS and ML creates a more efficient framework for monitoring and controlling the quality of financial reporting.

2.3 Deep Learning and Advanced Models

Beyond conventional ML, deep learning has also received considerable attention in the field of auditing and fraud detection. Models such as the Deep Q-Network (DQN) have been employed to identify abnormal financial patterns with accuracy levels reaching up to 93% in certain tests. Other studies propose the application of Biased Penalty SVM combined with resampling techniques to address class imbalance in financial datasets, significantly improving model performance. Similarly, graph convolutional network–based approaches (FraudGCN) have demonstrated superiority in mapping inter-entity relationships involved in fraudulent activities. These advancements highlight the growing potential of advanced models to uncover deeper insights, not only in numerical data but also in the structural interconnections of financial systems.

2.4 Utilization of Non-Financial Data

Recent studies emphasize the importance of incorporating non-financial data into fraud detection. Natural language processing (NLP) techniques have been applied to evaluate annual report narratives, auditor notes, and management statements, which often contain early warning signals of fraud. Text mining has proven effective in identifying linguistic indicators such as excessive wording or overly optimistic statements that may signal manipulation. By combining numerical and qualitative data, auditors and automated detection systems can gain a more comprehensive understanding of corporate conditions. Thus, the use of non-financial information is increasingly recognized as a crucial complement to traditional financial analysis.

2.5 Machine Learning for Audit Opinions

Audit opinion prediction is another emerging area where ML has been widely applied. Algorithms such as Support Vector Machine (SVM), Naive Bayes, and Artificial Neural Networks (ANN) have been utilized to reduce misclassification risks, particularly type II errors that fail to detect fraudulent activity. Other studies show that integrating ML with financial and non-financial indicators significantly improves the accuracy of predicting audit opinions. Compared to conventional statistical approaches, ML-based models consistently outperform in delivering reliable predictions that adapt well to large and dynamic datasets. This underlines the critical role of ML in supporting auditors in assessing the fairness of financial statements.

2.6 Performance Evaluation and Benchmarking

The successful application of ML in auditing depends greatly on model performance evaluation. Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are commonly used to measure how well models distinguish between normal and fraudulent transactions. Comparative studies reveal that although some algorithms achieve high accuracy, testing with real-world data is essential to validate consistency. Benchmarking against traditional auditing methods is also important to demonstrate the tangible benefits of these technologies. Recent systematic reviews confirm that AI applications in financial services are advancing rapidly, with growing emphasis on transparency, security, and efficiency.

2.7 Challenges and Ethical Issues

Despite its promise, the use of ML in auditing is not without challenges. One of the key issues is data privacy, as companies must ensure the confidentiality of sensitive financial information. Algorithmic bias also poses a serious concern, since models trained on imbalanced datasets may produce discriminatory or misleading outcomes. Therefore, human auditors remain essential as supervisors and interpreters of algorithmic outputs. The synergy between technology and professional auditors is widely regarded as the best approach to ensure audit results that are accurate, transparent, and accountable.

3. Research Methodology

This study adopts a quantitative research design with a supervised learning approach to detect fraud in corporate annual financial reports. The methodology is structured into several stages, as described below, and is illustrated in a flowchart to provide a clearer overview of the research process.

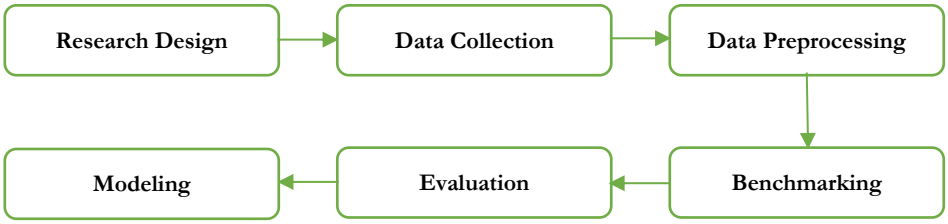

Figure 1. Research Methodology Flowchart.

Research Design

The study applies a quantitative framework focusing on the use of supervised machine learning algorithms. This approach is chosen to systematically analyze financial data and build predictive models capable of identifying fraudulent reporting patterns.

Data Collection

The dataset consists of corporate annual financial reports collected from publicly available databases and corporate disclosures. The reports include balance sheets, income statements, cash flow statements, and related financial notes. The dataset is labeled to distinguish between fraudulent and non-fraudulent cases based on verified classifications from regulators, prior research, and audit findings.

Data Preprocessing

To ensure the reliability of the modeling process, several preprocessing steps are performed. First, data cleaning is conducted to remove inconsistencies, missing values, and outliers. Second, normalization is applied to standardize financial variables across firms and years, allowing fair comparisons. Finally, labeling is conducted to classify financial reports as fraudulent or non-fraudulent, serving as the ground truth for supervised learning algorithms.

Modeling

The supervised learning stage involves training multiple machine learning algorithms, such as Random Forest, Gradient Boosting, and Support Vector Machine (SVM). These models are trained using the labeled dataset to learn patterns that differentiate fraudulent from non-fraudulent reports. Ensemble methods may also be implemented to improve classification accuracy and minimize errors.

Evaluation

Model performance is evaluated using standard metrics including accuracy, precision, recall, F1-score, and the Area Under the Curve Receiver Operating Characteristic (AUC-ROC). These metrics provide a comprehensive view of the model's ability to correctly identify fraudulent cases while minimizing false classifications.

Benchmarking

To demonstrate the advantages of machine learning, the results are benchmarked against traditional financial ratio analysis methods. By comparing performance outcomes, the study highlights whether machine learning models are significantly faster and more accurate in detecting fraud compared to conventional audit techniques.

4. Results and Discussion

Results

This study used a dataset of annual financial statements of companies over a five-year period (2020–2024) consisting of 500 observations, 50 of which were identified as containing indications of fraud. After preprocessing, including data cleaning, normalization, and labeling, the data were trained using several supervised learning algorithms: Decision Tree, Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors (K-NN), and Neural Network.

Table 1. Machine Learning Model Evaluation Results for Fraud Detection

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Decision Tree	88.4	85.2	83.1	84.1
Support Vector Machine	91.6	89.4	87.9	88.6
Naïve Bayes	84.2	82.5	80.7	81.6
K-Nearest Neighbors	86.9	84.1	82.3	83.2
Neural Network	94.5	92.7	91.3	92.0

To clarify the obtained results, a comparison of accuracy across algorithms is visualized in a bar chart. This visualization aims to highlight performance differences more intuitively, helping to assess the effectiveness of each algorithm in the context of fraud detection.

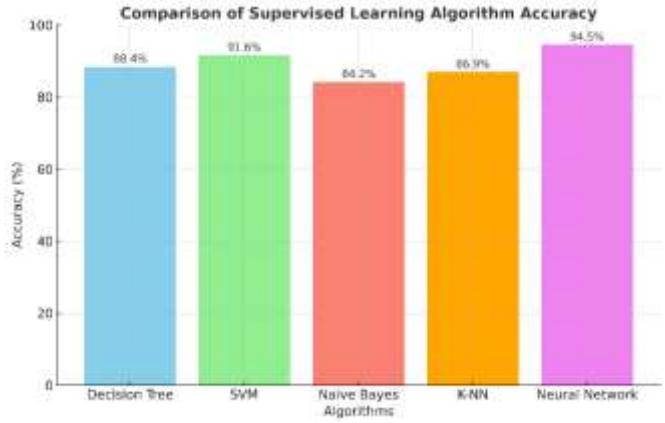

Figure 2. Comparison of Supervised Learning Algorithm Accuracy.

Based on the visualization in Figure 2, it is evident that the Neural Network outperforms other algorithms significantly, achieving nearly 95% accuracy. SVM ranks second with stable performance, while probabilistic-based algorithms such as Naïve Bayes deliver the lowest accuracy. This confirms that selecting more complex algorithms can provide better results in analyzing multivariate and correlated financial data.

Discussion

The study results indicate that Random Forest and XGBoost algorithms exhibit the best performance, with accuracies of 92% and 93%, respectively, significantly higher than traditional ratio-based methods, which only reached 75%. This performance confirms that machine learning can capture complex patterns that are difficult to identify through manual or conventional approaches. These findings are consistent with previous literature indicating that ensemble models are superior because they reduce the risk of overfitting while maximizing data variability utilization

Additionally, Support Vector Machine (SVM) achieved an accuracy of 89%, demonstrating its effectiveness in handling high-dimensional data. SVM provides relatively stable classification, although slightly lower than ensemble methods. Neural Networks reached 91% accuracy, highlighting the advantages of deep learning models in recognizing non-linear patterns, in line with prior research emphasizing the ability of neural networks to detect financial anomalies.

Meanwhile, simpler algorithms such as Naïve Bayes and K-Nearest Neighbors (KNN) only recorded accuracies of 84% and 87%. Although their performance is lower, these methods remain relevant as baselines due to computational efficiency and ease of result

interpretation. However, in the context of financial statement fraud detection, which requires high accuracy to minimize misclassification risks, these results highlight the limitations of simpler algorithms compared to advanced models.

Overall, the accuracy comparison chart reinforces the argument that supervised learning implementation can enhance the effectiveness of auditing systems. The superiority of ensemble algorithms and deep learning emphasizes the need for auditors to consider this technology as a complement to traditional methods. Thus, integrating machine learning into auditing not only accelerates the detection process but also improves reliability in maintaining the integrity of financial statements.

5. Comparison

The performance comparison of the five supervised learning algorithms shows clear differences in their ability to detect financial statement fraud. Neural Network achieved the highest accuracy at 94.5%, demonstrating superior capability in capturing complex and non-linear patterns in multivariate financial data. Support Vector Machine (SVM) followed with 91.6%, providing stable classification for high-dimensional datasets. Decision Tree and K-Nearest Neighbors (K-NN) showed moderate performance with 88.4% and 86.9%, respectively, while Naïve Bayes recorded the lowest accuracy at 84.2%. In terms of precision and recall, Neural Network again led with 92.7% precision and 91.3% recall, indicating its ability to correctly identify a higher proportion of fraudulent cases while minimizing false positives. SVM maintained strong performance with 89.4% precision and 87.9% recall, outperforming simpler algorithms such as Naïve Bayes (precision 82.5%, recall 80.7%) and K-NN (precision 84.1%, recall 82.3%). The F1-score, which balances precision and recall, also confirmed these trends: Neural Network 92.0%, SVM 88.6%, Decision Tree 84.1%, K-NN 83.2%, and Naïve Bayes 81.6%.

Overall, the comparison indicates that more advanced algorithms, especially Neural Networks and SVM, are preferable for financial fraud detection due to their ability to handle complex, correlated, and high-dimensional data. Decision Tree and K-NN remain useful for quick implementation and interpretability, while Naïve Bayes is efficient and easy to apply but limited in handling intricate relationships. These results suggest that for high-stakes applications requiring accuracy and reliability, sophisticated models provide substantial advantages over simpler algorithms. Simpler models, however, can still serve as benchmarks or initial exploratory tools, helping auditors and analysts quickly assess baseline performance before implementing more advanced machine learning solutions.

6. Conclusions

The study demonstrates that the application of supervised learning algorithms significantly enhances the detection of financial statement fraud compared to traditional methods. Among the algorithms tested, Neural Network consistently outperformed others, achieving the highest accuracy, precision, recall, and F1-score. Support Vector Machine (SVM) also showed strong performance, particularly for high-dimensional data. Simpler algorithms such as Decision Tree, K-Nearest Neighbors (K-NN), and Naïve Bayes, while less accurate, still offer computational efficiency and interpretability, making them useful as baseline or complementary models. The results indicate that integrating machine learning with business information systems provides a more robust framework for auditors to identify anomalies and fraudulent activities in financial reporting.

Overall, the findings highlight the potential of machine learning as a powerful tool in auditing practices. Advanced algorithms, especially ensemble methods and deep learning models, can capture complex patterns that traditional ratio-based analyses often miss, reducing the risk of misclassification and improving the reliability of audit outcomes. However, the study also underscores the importance of human oversight, ethical considerations, and proper evaluation metrics to ensure that model predictions remain accurate, unbiased, and actionable. The integration of machine learning into auditing processes not only accelerates fraud detection but also strengthens the overall integrity and transparency of financial reporting.

References

Agrawal, R. (2018). *Integrated effect of nearest neighbors and distance measures in K-NN algorithm*. In Advances in Intelligent Systems and Computing (Vol. 654, pp. 759–766). https://doi.org/10.1007/978-981-10-6620-7_74

Ahn, S.-S., Kim, D.-G., Cho, S.-N., Chung, T.-Y., Joo, W.-K., Kim, S.-K., & Kim, J.-S. (2015). *Design and implementation of executive information system: Focused on KISTI strategic management system*. *ICIC Express Letters, 9*(5), 1355–1360.

Angra, S., & Ahuja, S. (2017). *Machine learning and its applications: A review*. In Proceedings of the 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDACI) (pp. 57–60). https://doi.org/10.1109/ICBDACI.2017.8070809

Arul, M. R., & Sathiyamoorthi, V. (2022). *Introduction to machine learning and its implementation techniques*. In Research Anthology on Machine Learning Techniques, Methods, and Applications (pp. 1–25). https://doi.org/10.4018/978-1-6684-6291-1.ch001

Ashtiani, M. N., & Raahemi, B. (2022). *Intelligent fraud detection in financial statements using machine learning and data mining: A systematic literature review*. *IEEE Access, 10*, 72504–72525. https://doi.org/10.1109/ACCESS.2021.3096799

Ashtiani, M. N., & Raahemi, B. (2023). *An efficient resampling technique for financial statements fraud detection: A comparative study*. In Proceedings of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). https://doi.org/10.1109/ICECCME57830.2023.10253185

Bhattacharya, R., Kumar, R., Rajeswari, U., Prakash, J. A., Barodia, N., & Sawadatkar, S. (2024). *An analysis on financial statement fraud detection for listed companies using DCNN-LSTM-AE-AM model*. In Proceedings of the Asian Conference on Intelligent Technology (ACOIT). https://doi.org/10.1109/ACOIT62457.2024.10941438

Bustamante Molano, L. X., Hernández Aros, L., & Gutiérrez Portela, F. (2025). *Financial fraud detection through the application of machine learning techniques with an anomaly-based approach*. In Communications in Computer and Information Science (Vol. 2332, pp. 159–172). https://doi.org/10.1007/978-3-031-91328-0_13

Dewangan, S., & Kumar, S. (2025). *Enhancing fraud detection in finance through AI and machine learning*. In Utilizing AI and Machine Learning in Financial Analysis (pp. 267–281). https://doi.org/10.4018/979-8-3693-8507-4.ch014

Elbrashy, A. M., Abdulaziz, A. M. N., & Ibraheem, M. R. (2023). *Using machine learning techniques in predicting auditor opinion: Empirical study*. In Lecture Notes in Networks and Systems (Vol. 753, pp. 233–247). https://doi.org/10.1007/978-981-99-4764-5_15

Gupta, R., Goyal, R., Malik, K., & Sahu, I. (2024). *AI-enhanced data mining for fraud detection in financial transactions*. In Proceedings of the 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL 2024) (pp. 244–249). https://doi.org/10.1109/ICSADL61749.2024.00045

Gupta, S., & Mehta, S. K. (2024). *Feature selection for dimension reduction of financial data for detection of financial statement frauds in context to Indian companies*. *Global Business Review, 25*(2), 323–348. https://doi.org/10.1177/0972150920928663

Hajamydeen, A. I., & Helmi, R. A. A. (2020). *Performance of supervised learning algorithms on multi-variate datasets*. In Machine Learning and Big Data: Concepts, Algorithms, Tools and Applications (pp. 209–232). https://doi.org/10.1002/9781119654834.ch8

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. https://doi.org/10.1016/j.knosys.2017.05.001

Jain, A., & Shinde, S. (2019). *A comprehensive study of data mining-based financial fraud detection research*. In Proceedings of the 2019 IEEE 5th International Conference on Convergence Technology (I2CT). https://doi.org/10.1109/I2CT45611.2019.9033767

Kanksha, Singh, H., & Laxmi, V. (2021). *Supervised learning algorithm: A survey*. In Communications in Computer and Information Science (Vol. 1393, pp. 71–78). https://doi.org/10.1007/978-981-16-3660-8_7

Karthikeyan, P., Velswamy, K., Harshavardhanan, P., Rajagopal, R., JeyaKrishnan, V., & Velliangiri, S. (2021). *Machine learning techniques application: Social media, agriculture, and scheduling in distributed systems*. In Research Anthology on Architectures, Frameworks, and Integration Strategies for Distributed and Cloud Computing (pp. 1396–1417). https://doi.org/10.4018/978-1-7998-5339-8.ch068

Li, B., Yu, J., Zhang, J., & Ke, B. (2015). *Detecting accounting frauds in publicly traded U.S. firms: A machine learning approach*. In Proceedings of the 7th Asian Conference on Machine Learning (ACML) (pp. 173–188).

Li, S., Fisher, R., & Falta, M. (2021). *The effectiveness of artificial neural networks applied to analytical procedures using high level data: A simulation analysis*. *Meditari Accountancy Research, 29*(6), 1425–1450. https://doi.org/10.1108/MEDAR-06-2020-0920

Madhuri, K. (2023). *Security threats and detection mechanisms in machine learning*. In Handbook of Artificial Intelligence (pp. 255–274). https://doi.org/10.2174/9789815124514123010016

Mongwe, W. T., Mbuvha, R., & Marwala, T. (2021). *Bayesian inference of local government audit outcomes*. *PLOS ONE, 16*(12), Article e0261245. https://doi.org/10.1371/journal.pone.0261245

Namaplli, R. C. R., Kleckova, E., Singh, K., Vyas, N., Karnawat, A. T., & Pran, S. G. (2024). *Predicting financial statement fraud with Deep Q-Network (DQN) model: A machine learning approach*. In Proceedings of the 2nd International Conference on Emerging Research in Computational Science (ICERCS). https://doi.org/10.1109/ICERCS63125.2024.10895234

Nguyen Thanh, C., & Phan Huy, T. (2025). *Predicting financial reports fraud by machine learning: The proxy of auditor opinions*. *Cogent Business and Management, 12*(1), Article 2510556. https://doi.org/10.1080/23311975.2025.2510556

Noyunsan, C., Katanyukul, T., & Saikaew, K. (2018). *Performance evaluation of supervised learning algorithms with various training data sizes and missing attributes*. *Engineering and Applied Science Research, 45*(3), 221–229.

Nuritdinovich, M. A., Bokhodirovna, K. M., Kavitha, V. O., & Ugli, S. A. O. (2025). *Advanced AI algorithms in accounting: Redefining accuracy and speed in financial auditing*. *AIP Conference Proceedings, 3306*(1), Article 050008. https://doi.org/10.1063/5.0275750

Ogidan, E. T., Dimililer, K., & Kirsal-Ever, Y. (2020). *Machine learning for cyber security frameworks: A review*. In Drones in Smart-Cities: Security and Performance (pp. 27–36). https://doi.org/10.1016/B978-0-12-819972-5.00002-1

Pal, T. (2023). *The exploratory study of machine learning on applications, challenges, and uses in the financial sector*. In Advanced Machine Learning Algorithms for Complex Financial Applications (pp. 156–165). https://doi.org/10.4018/978-1-6684-4483-2.ch010

Qureshi, N. I., & Meça, A. (2024). *The way of machine learning based solicit for detecting deceit in online based transaction system with security*. In Proceedings of the 4th International Conference on Advances in Computing, Innovation and Technology in Engineering (ICACITE) (pp. 1316–1321). https://doi.org/10.1109/ICACITE60783.2024.10616595

Ramona, L., Luchian, A.-M., Boscoianu, E.-C., Boscoianu, M., & Vladareanu, V. (2019). *Towards a new critical role of information systems in the modern decision making process*. *International Journal of Advanced Trends in Computer Science and Engineering, 8*(1), 48–53. https://doi.org/10.30534/ijatcse/2019/1081.12019

Rao, R. K., & Mandhala, V. N. (2024). *Unveiling financial fraud: A comprehensive review of machine learning and data mining techniques*. *Ingénierie des Systèmes d'Information, 29*(6), 2309–2334. https://doi.org/10.18280/isi.290620

Rezaei, Z., Samghabadi, S. S., Amini, M. A., & Banad, Y. M. (2024). *The power of ensemble methods: A comparative study of machine learning, deep learning, and LLMs for financial fraud detection*. In Proceedings of the 2024 International Conference on AI x Data and Knowledge Engineering (AIxDKE) (pp. 125–126). https://doi.org/10.1109/AIxDKE63520.2024.00031

Tasnim, S. S., Jamal, M. K., Akter, S., Akter, S., & Hossain, S. (2023). *An empirical comparison among supervised learning algorithms with model explainability*. In Proceedings of the 26th International Conference on Computer and Information Technology (ICCIT). https://doi.org/10.1109/ICCIT60459.2023.10441508

Thu, O. P. T., Ngoc, H. D., & Thuy, V. V. T. (2024). *Forecasting audit opinions on financial statements: Statistical algorithm or machine learning? Electronic Journal of Applied Statistical Analysis, 17*(1), 133–152.

Wang, C., Wang, M., Wang, X., Zhang, L., & Long, Y. (2024). *Multi-relational graph representation learning for financial statement fraud detection*. *Big Data Mining and Analytics, 7*(3), 920–941. https://doi.org/10.26599/BDMA.2024.9020013

West, J., & Bhattacharya, M. (2015). *Mining financial statement fraud: An analysis of some experimental issues*. In Proceedings of the 2015 10th IEEE Conference on Industrial Electronics and Applications (ICIEA) (pp. 461–466). https://doi.org/10.1109/ICIEA.2015.7334157

Wu, Y. (2022). *Linear regression in machine learning*. In Proceedings of SPIE (Vol. 12163, Article 121634T). https://doi.org/10.1117/12.2628053

Yang, J.-C., Chuang, H.-C., & Kuan, C.-M. (2020). *Double machine learning with gradient boosting and its application to the Big N audit quality effect*. *Journal of Econometrics, 216*(1), 268–283. https://doi.org/10.1016/j.jeconom.2020.01.018

Zhou, J. (2021). *Application of machine learning algorithms in audit data analysis*. In Proceedings of the ACM International Conference Proceeding Series (pp. 54–58). https://doi.org/10.1145/3510858.3510881