ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 15 (2025) No. 1, pp. 69-84
DOI:10.33168/JSMS.2025.0105

A Systematic Review of Machine Learning-Based Approaches for Financial Fraud Detection

Teguh Wahyono, Felix David

Faculty of Information Technology, Satya Wacana Christian University
teguh.wahyono@uksw.edu (Corresponding author), felix@uksw.edu

Abstract. Financial fraud detection is a mounting challenge with rising digitsization of systems and data complexity. Although machine learning (ML) techniques have shown promise in identifying fraudulent patterns, substantial research gaps remain regarding algorithms, optimizations, and real-world effectiveness. This systematic literature review comprehensively synthesizes 51 recent studies on ML based fraud analysis in finance spanning supervised, unsupervised and reinforcement learning approaches. Findings reveal favored algorithms like random forests and key strategies around handling class imbalance, feature selection, model ensembles etc. The survey further unveils limitations in current literature such as data transparency and model interpretability. It also discusses future opportunities around hybrid models, reinforcement learning, and standardized evaluation metrics to enable nuanced fraud examinations without compromising privacy. Through its structured analysis, this review contributes a state-of-the-art ML perspective to advance fraud scholarship and practice.

Keywords: machine learning, financial fraud detection, systematic review, technological issues

1. Introduction

Fraud is a dishonest and harmful act carried out by irresponsible parties with the intention of gaining illegal profits (Zhu et al., 2021). In the financial sector, fraud can cause losses to companies, investors, and even society. A survey conducted by the Association of Certified Fraud Examiners (ACFE) on World Occupational Fraud reported that financial fraud in 2022 resulted in an average loss of USD 117,000 (ACFE, 2022). The most significant losses were in the real estate sector, followed by trade, transportation, construction, and utilities. The survey also revealed that fraud most commonly occurred in operational roles, followed by accounting and sales positions. Data from The Nilson Report indicates a significant increase in credit card fraud in recent years (Joshi et al., 2021). As shown in Figure 1, financial losses reached $31.67 million in 2020 due to credit card fraud.

Meanwhile, the rapid digital transformation following the COVID-19 pandemic has brought new challenges related to financial fraud. Although digitalization has made financial services more accessible to the public, it has also led to an increase in financial fraud incidents. The ACFE Fraud in the Wake of COVID-19 Benchmarking Report shows a substantial increase in fraud cases post-COVID-19 (ACFE, 2021). Additionally, Grant Thornton's report titled "The Next Normal: Preparing for a Post-Pandemic Fraud Landscape" states that more than half of the surveyed organizations (51%) in 2021 reported detecting more fraud cases since the beginning of the pandemic. In Indonesia, the National Cyber and Cryptography Agency (BSSN) reported an increase in cyberattacks in 2021, with 23% occurring in the financial sector. In the same year, the Financial Services Authority (OJK) received 7,087 reports of fraud cases in the banking industry through cybercrimes, resulting in total losses exceeding 246 billion Indonesian rupiahs (Kominfo, 2023).

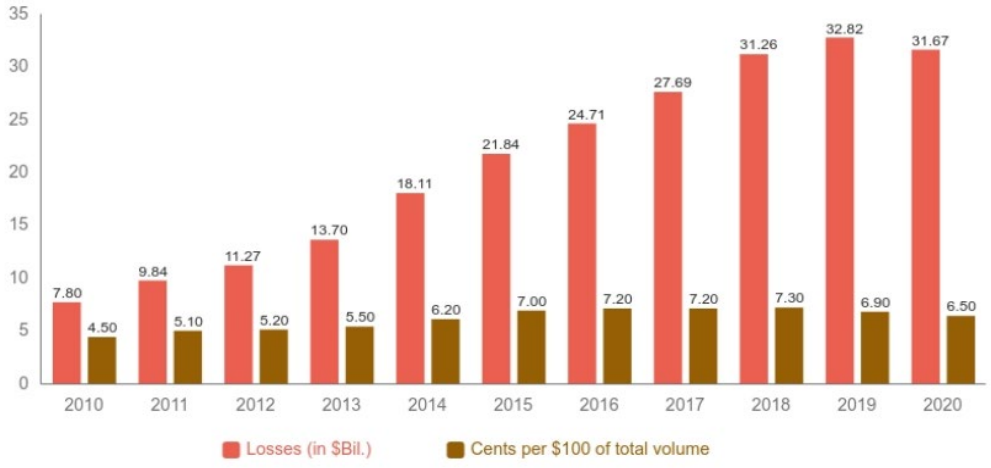

Fig.1: Card Fraud Worldwide (Joshi et al., 2021).

The various incidents highlight the substantial financial losses that fraud can inflict on companies, investors, and society. Fraud can take on various forms, such as manipulating financial reports, misappropriating funds, and more. Within a company, fraud can also harm its reputation, negatively impacting its image in the eyes of investors and the public (Zhu, el at., 2021; ACFE, 2022). Damage to reputation can result in a loss of investor trust and affect the company's stock value. Furthermore, fraud constitutes illegal actions that can lead to legal sanctions for both the company and those involved. Given these backgrounds, fraud detection is of paramount importance to prevent greater financial losses and maintain the company's legal compliance (Wahyudi et al., 2022). Detecting fraud also plays a crucial role in safeguarding the interests of stakeholders like investors, employees, and consumers. Therefore, companies must prioritize security and monitoring in every transaction to prevent fraud and protect stakeholder interests.

Fraud detection becomes increasingly crucial due to the advancing digital transformation and the

growing complexity of financial transactions. Effective financial fraud detection necessitates rapid and accurate data processing to minimize risks to companies and society. Therefore, the application of Machine Learning is vital for the swift analysis of vast and diverse transaction data. Machine Learning has the capability to identify patterns and trends in fraud, enabling early detection of potential fraudulent activities (Mustika et al., 2021). Several prior research studies have focused on applying machine learning to detect financial fraud. For example, Husein et al. (2021) employed anomaly detection techniques with Support Vector Machine (SVM) and Random Forest algorithms to detect fraud in credit card transactions. Similarly, Zamachsari [10] conducted fraud detection in electronic financial transactions to prevent losses to customers and banks. This research recommended using the Synthetic Minority Over-sampling Technique (SMOTE) in combination with deep learning for the best detection model. Mohmad (2022) used Long Short Term Memory (LSTM) algorithms to detect fraud in credit risk assessment systems, showing that LSTM significantly improved the accuracy of risk assessment. These studies demonstrate that the application of Machine Learning produces more effective results in detecting various financial fraud incidents compared to traditional methods (Hussain et al., 2021; Zamachsari, 2021).

To obtain a comprehensive understanding of the application of machine learning in financial fraud detection, a systematic literature review is essential. Hence, this research aims to map the application of machine learning in the field of financial fraud detection, as well as answer several key questions about techniques, optimization and trends in machine learning based fraud detection. Some of the research questions addressed in this study include: what types of financial fraud cases can be handled through machine learning approaches, which machine learning algorithms are effective for financial fraud detection, what strategies or optimizations can be employed to enhance accuracy in financial fraud detection, and what are the research gaps, trends, and future directions in this area of study. The data collected are research articles that discuss at least 3 keywords, namely financial fraud, detection and machine learning, published in the last six years between 2018 and 2023. The results are expected to provide a systematic and comprehensive overview of the state of the art and prior research related to the application of Machine Learning for Financial Fraud Detection. This overview, along with the framework, is also expected to provide a roadmap for researchers and practitioners seeking a more comprehensive understanding of this field.

2. Research Methodology

The research was conducted using the Systematic Literature Review (SLR) method. Kitchenham (2007) defined SLR as a method for identifying, evaluating, and interpreting all findings on a research topic based on prior research to address predefined research questions. The SLR method is systematically performed, following stages and protocols that minimize bias and subjective interpretation by the researchers. This approach can be used to obtain valid and applicable findings from multiple previous studies related to a specific phenomenon (Ali et al., 2022).

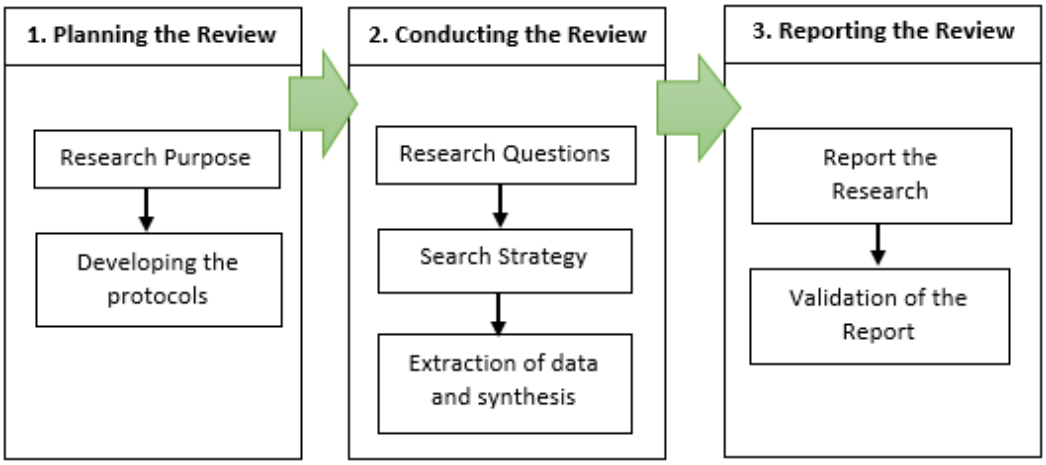

Fig.2: Metode Penelitian

2.1. Planning the Review

The planning phase involves identifying the research objectives and establishing a protocol. The protocol, in this context, is a document that details the design and methodology to be used in the literature review (Kitchenham, 2007). This document includes the research objectives, research questions, inclusion and exclusion criteria, literature search strategy, article selection method, data collection and analysis, as well as the evaluation of the quality of selected articles.

2.2. Conducting the Review

Following the planning phase, the next step is conducting the review. In this section, the formulation of the Research Question (RQ) is carried out, indicating the main issues to be discussed and analyzed in the review. This is followed by the formulation of a search strategy and data extraction. Thus, the RQ serves as the initial guide for the literature search and extraction process. The analysis and synthesis of data resulting from the SLR provide the answers to the predetermined RQ. The following is a further explanation of these stages (Ali, 2022). The formulation of the RQ is based on five elements consisting of Population and issue, Intervention, Comparison, Outcomes, and Context, known as the PICOC criteria (Kitchenham, 2007). Table 1 presents the PICOC criteria in this research.

Tabel 1. PICOC Criteria

Criteria	Information
Population	Machine Learning, Financial Fraud
Intervention	Detection Systems, Algorithm, Methods, Strategies
Comparison	Comparations between methods
Outcomes	Accuracy of computational methods
Context	Academic Research

After determining the contents of the PICOC Criteria, the next step is to formulate the research questions (RQ) for this study as follows.

Tabel 2. Research Question

ID	Research Question	Motivation
RQ1	What types of financial fraud cases can be addressed using a machine learning approach?	To understand and map the types of cases handled with a machine learning approach
RQ2	Which machine learning algorithms are effective for financial fraud detection?	To identify effective algorithms used for financial fraud detection

RQ3	What strategies or optimizations can be employed to enhance the accuracy of financial fraud detection?	To identify optimization strategies aimed at improving machine learning accuracy
RQ4	What are the research gaps, trends, and future directions in this research area?	To identify research gaps and provide recommendations for future research

The search strategy involves identifying the sources of literature to be targeted. The literature screening process was carried out using the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) method. PRISMA is the accepted standard for presenting evidence in systematic literature reviews. The stages of the article selection and review process using the PRISMA method can be seen in Figure 3. In this study, seven sources of database were chosen: ACM Digital Library, IEEE Explore Digital Library, EBSCO Host, Google Scholar, Science Direct, Semantic Scholar, and Springer. The next step is determining the search string for the literature based on the predefined PICOC criteria. Selecting the appropriate keywords will determine the accuracy of the literature found. The following search string was used in this research: ("financial fraud" OR "anomaly transaction") AND ("detection" OR "prediction" OR "forecasting") AND ("machine learning" OR "artificial intelligence").

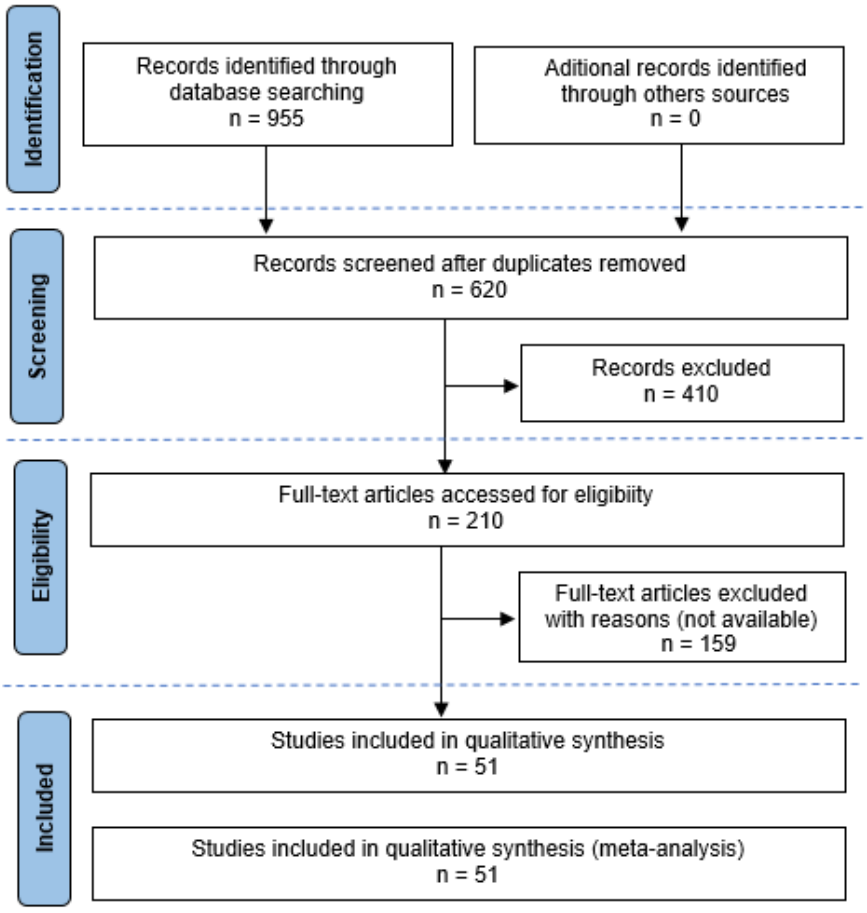

Fig.3: PRISMA Flow Diagram.

The data search resulted in 955 records identified according to the keywords provided. Only articles from peer reviewed journals and articles from reputable conferences were selected, to ensure the scientific integrity of this research. Several articles such as books, magazines, short surveys, correspondence, discussions, book reviews, product reviews, editorials and publisher notes are not included in the selected paper criteria. Then the year of publication is filtered from 2018 to 2023.

Thematic analysis was conducted to identify themes related to our research patterns and trends. In the end, 51 articles were selected to be included in the qualitative synthesis (meta-analysis).

After obtaining the literature, the next step is data extraction, followed by the synthesis of various findings from the literature. Synthesis can be done narratively or quantitatively (meta-analysis). The primary purpose of data synthesis is to analyze and evaluate the various research findings from the literature and to select the most appropriate method for integrating explanations and interpretations of various findings (Wahono, 2015).

2.3. Reporting The Review

Reporting refers to writing the results of the SLR in a scientific article format. The structure of the article comprises three main sections: Introduction, Main Body, and Conclusion. The Introduction section provides the background and rationale for the SLR. The main body section includes the protocol, results of the analysis and synthesis of findings, and ends with a discussion of the implications of the results. The Conclusion section summarizes the findings obtained in accordance with the predefined RQs.

3. Results and Discussion

3.1. Description of Studies

Employing the defined search strategy and selection criteria, this research identified 51 relevant articles. The articles are as follows: 3 articles from ACM Digital Library (6%), 3 articles from EBSCO (6%), 9 articles from Google Scholar (17%), 14 articles from IEEE Explore (27%), 8 articles from Science Direct (16%), 6 articles from Semantic Scholar (12%), and 8 articles from Springer (16%). Figure 4 below illustrates the composition of these article sources.

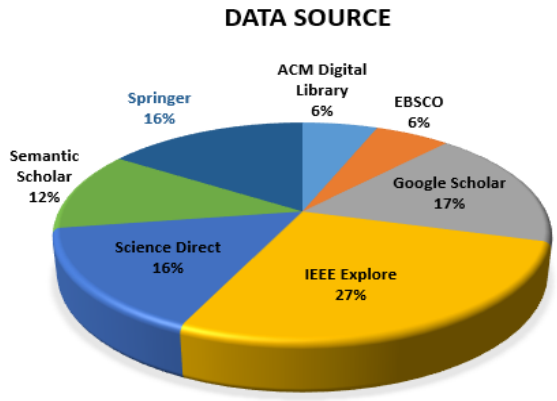

Fig.4: The composition of the article based on a data source.

In terms of publication years, 4 articles were published in 2018, 7 articles in 2019, 6 articles in 2020, 13 articles in 2021, 10 articles in 2022, and 11 articles in 2023. The distribution of articles per year can be seen in Figure 5 below.

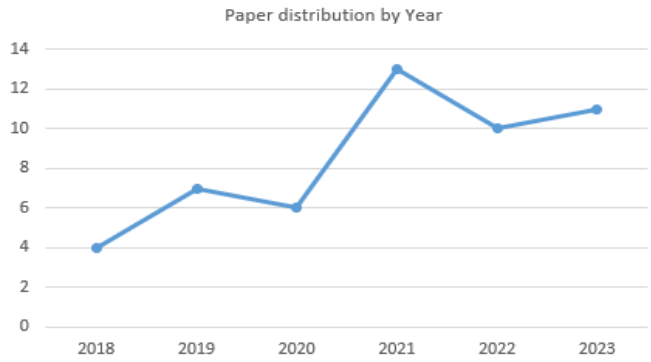

Fig.5: Distribution of Selected Paper over the Year

3.2. Synthesis Results

This section discusses the answers to the predefined research questions based on the findings of the systematic literature review.

3.2.1. Financial Fraud Areas

This section addresses RQ1, which pertains to the types of financial fraud cases that can be addressed using a machine learning approach. Based on the review conducted, there are eight common areas of financial fraud frequently handled with machine learning as can be seen in Figure 6. These eight areas are: tax, banking, investment and capital markets, fintech and online transactions, healthcare and insurance, credit cards, financial statements, and e-commerce. Among these areas, credit card fraud is the most commonly occurring (15 studies), e-commerce fraud (9 studies), followed by financial statement fraud (8 studies), and healthcare and insurance fraud (e-commerce fraud (7 studies).

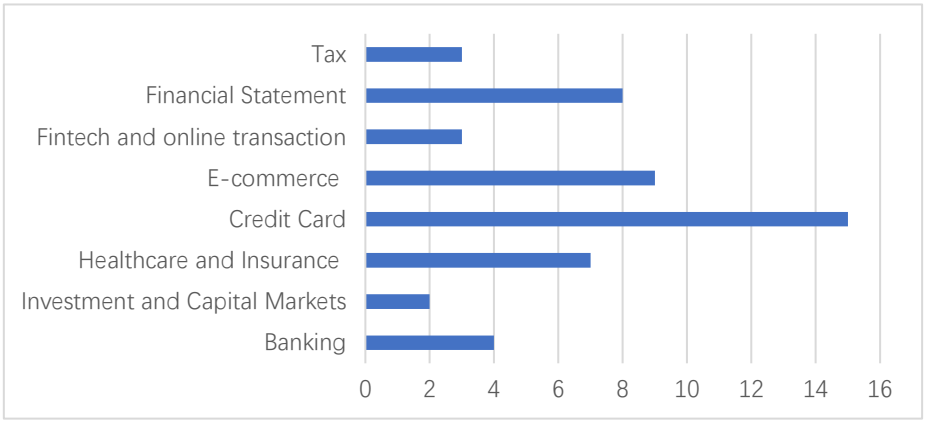

Fig.6: Area Financial Fraud Detection dengan Machine Learning

3.2.2. Effective Machine Learning Algorithms

This section addresses RQ2, which focuses on the machine learning algorithms that are effective for financial fraud detection. Based on the data obtained in this research, 72% of financial fraud detection methods utilize supervised learning, 16% employ unsupervised learning, 8% use hybrid approaches, and 4% leverage reinforcement learning. The findings are presented in Figure 7 below.

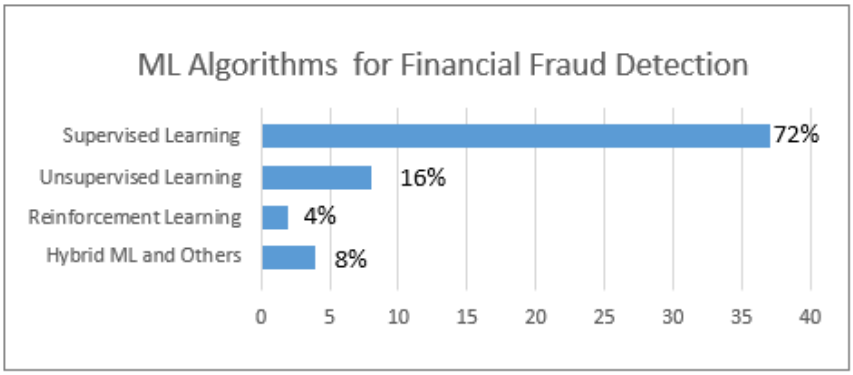

Fig.7: ML Algorithms for Financial Fraud Detection

Supervised Learning

Supervised learning is a machine learning approach that uses labeled data or data whose characteristics are known to the designer. These data then serve as training data for an algorithm to perform actions as needed, such as accurate classification or prediction. The most commonly used methods include Random Forest, followed by Support Vector Machine, Logistic Regression, Neural Network, and LSTM.

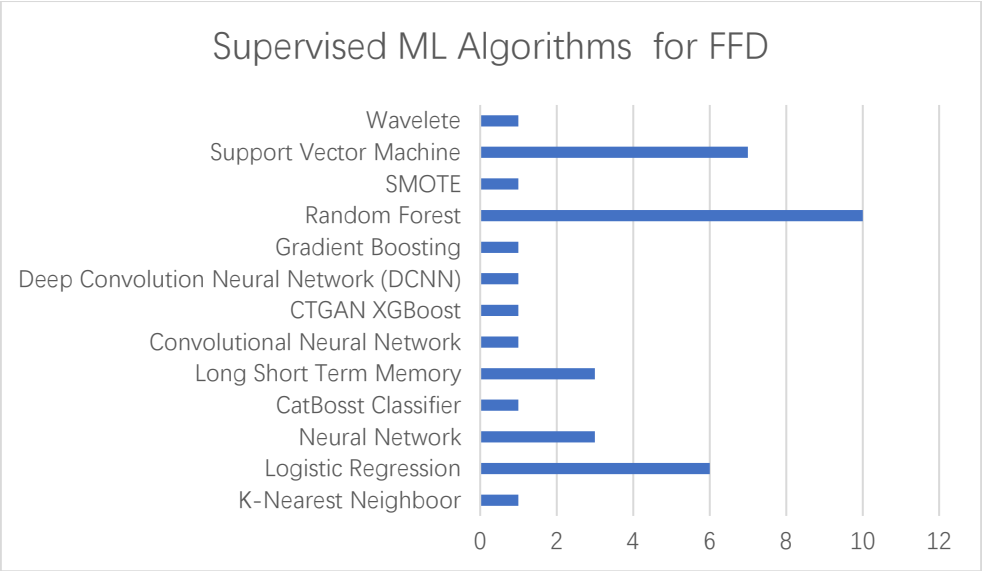

Fig.8: Supervised Algorithms for Financial Fraud Detection

In addition to these top five methods, other algorithms used include Wavelet, SMOTE, Gradient Boosting, Deep Convolutional Neural Network (DCNN), Catboost Classifier, and K-Nearest Neighbor. The summary of article counts with supervised learning methods and relevant references can be seen in Table 2 below.

Tabel 2. Article counts with supervised learning methods.

Machine Learning Algorithm	Article Count	References
Random Forest	10	Dhankhad et al. (2018), Yao et al. (2018), Trivedi et al. (2020), Mehbodniya et al. (2021), Severino & Yahao, (2021), Lokanan et al. (2022), Xu et al.

		(2022), Afriyie et al. (2023), Khosravi et al. (2023), Mohamed & Subramanian (2023).
Support Vector Machine	7	Thennakoon et al. (2019), Adepoju et al. (2019), Shen et al. (2021), Kaur et al. (2021), Xu et al. (2022), Aslam et al. (2022), Nalluri et al. (2023).
Logistic Regression	6	Itoo et al. (2021), Pani et al. (2021), Misra & Pandey (2021), Islam et al. (2022), Verma & Tyagi (2022), Debener et al. (2023).
Neural Network	3	Mittal & Tyagi (2021), Rai & Dwivedi (2020), Vuppula (2021)
LSTM	3	Yara, et al (2020), Sehrawat et al. (2023), Benchaji et al. (2021).
Wavelet	1	Ivanyuk (2023).
SMOTE	1	Varmedja wt al. (2019).
Gradient Boosting	1	Valavan & Rita (2023).
Deep Convolutional Neural Network	1	Chen & Lai (2021).
Catboost Classifier	1	Chen & Han (2021).
K-Nearest Neighbor	1	Mehbodniya, Abolfazl et al. (2021)

In the supervised learning, Random Forest is the most popular algorithm for fraud detection, because this algorithm is relatively easy to implement and configure Dhankhad et al. (2018). Additionally, it does not require much parameter tuning to provide good results Mehbodniya et al. (2021). Random Forest is often used when faced with data that is complex, unbalanced, and has patterns that are difficult to detect directly (Lokanan et al., 2022; Xu et al., 2022; Afriyie et al., 2023).

Unsupervised Learning

Unsupervised learning is a machine learning algorithm for analyzing and clustering unlabeled datasets (Goud & Premchand, 2019). This algorithm uncovers hidden patterns and groups data based on similarities among their characteristics (Agbakwuru & Elei, 2021). Some commonly used unsupervised learning methods for financial fraud detection include Hidden Markov Model, Isolation Forest, K-Means, Restricted Boltzmann, Self-Organizing Maps, and Spectral Clustering.

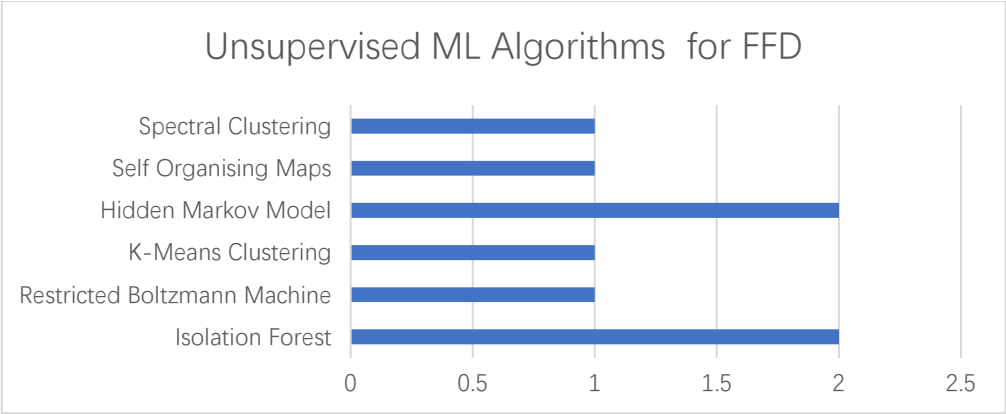

Fig.9: Unsupervised Algorithms for Financial Fraud Detection

The summary of article counts with unsupervised learning methods and relevant references can be seen in Table 3 below.

Tabel 3. Unsupervised learning methods and relevant references

Machine Learning Algorithm	Article Count	Reference
Hidden Markov Model	2	Goud & Premchand (2019), Agbakwuru & Elei (2021).
Isolation Forest	2	Bodepudi (2021), Shukur et al. (2019)
K-Means	1	Kanjanawattana (2019).
Restricted Boltzmann	1	Mubalaike and Adali (2020).
Self-Organizing Maps	1	Mongwe and Malan (2020).
Spectral Clustering	1	Roux et al. (2018)

In unsupervised learning, Isolation Forest and Hidden Markov Model are the most popular algorithms used. Isolation Forest was chosen because it has the advantage of being able to identify anomalies with few parameters (Bodepudi, 2021; Shukur et al., 2019). Isolation Forest has high computing speed because this algorithm focuses on finding solutions randomly and uses a simple tree structure. Meanwhile, the Hidden Markov Model is a probabilistic model that can overcome uncertainty in data, being effective in handling sequential data, such as time series or other sequential data (Goud et al., 2019, Agbakwuru et al., 2021).

Reinforcement Learning

Reinforcement learning is a machine learning technique that enables an agent (the entity taking actions) to learn in an interactive environment through a trial-and-error system of feedback based on their actions and experiences (Vimal et al., 2021). Reinforcement learning uses rewards and punishments as signals for making decisions or taking subsequent actions (Degirmenci, & Jones, 2022). Some reinforcement learning methods used in financial fraud detection cases include Deep Q-learning (Vimal et al., 2021) and Offline Reinforcement (Degirmenci, & Jones, 2022).

3.2.3. Strategies for Improving the Accuracy of Financial Fraud Detection

Enhancing accuracy in financial fraud detection is crucial, as detection errors can lead to inaccuracies in strategic decision-making. This research identified several strategies that can be employed to help improve the accuracy of financial fraud detection.

The first strategy is the collection of high-quality datasets (Al-Hashedi et al., 2021). The dataset used for model training must be complete, accurate, and up-to-date. It should be cleaned of noise and duplicates, and missing values should be handled appropriately (Dhankhad et al., 2018; Yao et al., 2018; Trivedi et al., 2020). The next strategy involves addressing imbalanced data (Abolfazl et al., 2021; Xu et al, 2022; Afriyie et al., 2023). In this regard, it's essential to ensure a balance between the fraud and non-fraud classes in the dataset. If the data is imbalanced, oversampling techniques (Shukur et al., 2019; Xu et al., 2022; Vuppula et al, 2021) or undersampling methods (Dhankhad et al., 2018; Mehbodniya et al., 2021) may need to be considered to address the imbalance.

Another important strategy is the identification of relevant features to enhance the model's ability to identify fraud patterns (Yao et al., 2018; Adepoju et al., 2019; Vuppula et al., 2021; Valavan & Rita, 2023). Various techniques to consider include dimension reduction, such as PCA (Roux et al., 2018; Shukur et al,. 2019; Vuppula, 2021), to reduce data complexity. Handling missing values can be done in various ways, such as imputing missing data using appropriate estimates or removing transactions if there is insufficient information (Afriyie et al., 2023; Khosravi et al., 2023).

Th next is the strategy for selecting a machine learning model that suits the data type and the fraud detection problem at hand. Hyperparameter optimization is also essential (Lokanan et al., 2022; Xu et al., 2023), such as finding the best combination of hyperparameters, like learning rate, the number of trees in random forest, or the number of layers in a neural network. Cross-validation is necessary to avoid overfitting and ensure that the model can generalize well to unseen data (Matheus & Peng, 2021; Shen, 2021). Another strategy is ensemble learning (Ivanyuk, 2023; Valavan, 2023), which involves

combining multiple models (bagging or boosting) to enhance model performance. Random Forest (Afriyie et al., 2023; Khosravi et al, 2023; Mohamed & Subramanian, 2023) and Gradient Boosting (Valavan, 2023) are popular examples.

Furthermore, it is essential to maintain a balance between accuracy and the number of false positives. Machine learning models that produce too many false positives can erode user trust in the model. Therefore, carefully considering the threshold for classifying transactions as fraud or non-fraud is crucial (Adepoju, 2019; Rai & Dwivedi, 2020). Adjusting the threshold up or down can control the model's sensitivity to false positives (Rai & Dwivedi, 2020).

3.2.4. Research Gaps, Trends, and Future Directions

Financial fraud detection using machine learning has become an essential and rapidly evolving topic. However, several research gaps warrant further investigation. One of the research gaps relates to class imbalance (Xu, 2022; Valavan, 2023; Yi, 2023). Many financial fraud detection datasets still exhibit class imbalance, where the number of fraudulent transactions is significantly lower than legitimate transactions. Further research is required to address the challenges associated with this class imbalance. Exploration of oversampling methods (Shukur et al., 2019; Vuppula, 2021), undersampling (Mehbodniya, 2021), or specialized techniques like SMOTE (Synthetic Minority Over-sampling Technique) is needed (Varmedja, 2019; Zamachsari, 2021).

Moreover, in more complex machine learning models, especially deep learning, results are often challenging to interpret (Matheus & Yaohao, 2021; Lokanan, 2022). In the context of financial fraud detection, it is essential for analysts to understand why a transaction is considered fraudulent by the model. Therefore, further research into interpretability and explainability concerning fraud detection in the financial sector is required (Lokanan, 2022; Xu, 2022).

From a methodological perspective, unsupervised learning, particularly clustering, groups transactions based on their similarities in various features (Shukur, 2018; Kanjanawattana, 2018). Transactions falling into unusual groups may be considered suspicious. Common challenges include distinguishing between suspicious transactions and anomalies that are not fraudulent. Therefore, the results of unsupervised learning often require verification and further analysis before recommending a course of action. Additionally, exploring the possibility of combining unsupervised and supervised learning methods to enhance fraud detection accuracy and effectiveness is worth considering.

On the other hand, the use of Reinforcement Learning is intriguing because most fraud detection models developed using supervised/unsupervised learning have limitations when dealing with increasingly sophisticated fraud patterns and evolving fraudulent tactics (Vimal, 2021; Degirmenci, 2022). Reinforcement learning enables systems to learn and adapt automatically to changes in fraud patterns. This can address the issue where fraudsters continually innovate to evade detection.

A final recommendation relates to the evaluation standards for fraud detection models. Various evaluation standards such as precision, recall, F1-score, ROC-AUC should be extended with more comprehensive metrics that consider potential losses resulting from false positives and false negatives. Furthermore, irrespective of model performance, it's essential to consider data transparency and security. Further research is needed to determine the extent to which privacy-sensitive financial data can be used for analysis without compromising data privacy.

4. Conclusion and Recommendations

This systematic review offers a comprehensive investigation into machine learning based financial fraud detection spanning techniques, optimizations and open challenges. Supervised methods dominate current literature, especially random forests, but unsupervised and reinforcement learning approaches exhibit untapped potential. Strategies like data balancing, feature engineering, hyperparameter tuning prove consistent accuracy improvements. At the same time, shortcomings exist regarding model transparency, generalizability and evaluative standards. As digital transactions scale in complexity and

volume, purposeful research and design of interpretable, ethical and robust intelligent fraud models remains imperative. This survey consolidates an informed outlook to guide this interdisciplinary mission towards secure and trustworthy financial ecosystems.

Furthermore, the research proposes some recommendations related to research gaps, trends, and future directions. Some of the research recommendations include addressing class imbalance in datasets, investigating interpretability and explainability, particularly in the context of fraud detection in the financial sector, and exploring the possibility of combining unsupervised and supervised learning methods to improve accuracy and effectiveness in fraud detection.

Acknowledgment

We thank the Satya Wacana Christian University for funding this research through the Fundamental Research Program under grant number 037/SPK-PF/RIK/7/2023.

References

ACFE. (2021). *Fraud in the Wake of COVID-19: Benchmarking Report*, Association of Certified Fraud Examiners.

ACFE (2022), *Occupational Fraud 2022: A Report to the nations*, Association of Certified Fraud Examiners.

Adepoju, O., Wosowei, J., Lawte, S. and Jaiman, H. (2019). Comparative Evaluation of Credit Card Fraud Detection Using Machine Learning Techniques, *Global Conference for Advancement in Technology (GCAT),* pp. 1-6. https://doi.org/10.1109/GCAT47503.2019.8978372.

Afriyie, J.K., Tawiah, K., et al. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions, *Decision Analytics Journal*, 6(1). https://doi.org/10.1016/j.dajour.2023.100163.

Agbakwuru, A. O., & Elei, F. O. (2021). Hidden Markov model application for credit card fraud detection systems. *International Journal of Innovative Science and Research*, 5(1).

Alghofaili, Y., Albattah, A., & Rassam, M. (2020). A Financial Fraud Detection Model Based on LSTM Deep Learning Technique A Financial Fraud Detection Model Based on LSTM Deep Learning Technique. *Journal of Applied Security Research*. 15(4). https://doi.org/10.1080/19361610.2020.1815491.

Ali A, Abd Razak S, Othman S.H, Eisa TAE, Al-Dhaqm A, Nasser M, Elhassan T, Elshafie H, Saif A. (2022). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied Sciences.* 12(19). https://doi.org/10.3390/app12199637.

Aslam, F., Hunjra, A.I., et al. (2022). Insurance fraud detection: Evidence from artificial intelligence and machine learning, *Research in International Business and Finance*, 62(1), https://doi.org/10.1016/j.ribaf.2022.101744.

Benchaji, I., Douzi, S., El Ouahidi, B. et al. (2021). Enhanced credit card fraud detection based on attention mechanism and LSTM deep model. *Journal of Big Data,* 8(151). https://doi.org/10.1186/s40537-021-00541-8.

Bodepudi, H. (2021). Credit Card Fraud Detection Using Unsupervised Machine Learning Algorithms, *International Journal of Computer Trends and Technology*, 69(8), pp. 1-3. https://doi.org/10.14445/22312803/IJCTT-V69I8P101.

Chen, Joy & Lai, Kong-Long. (2021). Deep Convolution Neural Network Model for Credit-Card Fraud Detection and Alert. *Journal of Artificial Intelligence and Capsule Networks*. 3(1). pp. 101-112. https://doi.org/10.36548/jaicn.2021.2.003.

D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic and A. Anderla. (2018). Credit Card Fraud Detection - Machine Learning methods, *18th International Symposium Infoteh-Jahorina (INFOTEH)*, pp. 1-5, https://doi.org/10.1109/INFOTEH.2019.8717766.

De Roux, D., Perez, B., et al. (2018). Tax Fraud Detection for Under-Reporting Declarations Using an Unsupervised Machine Learning Approach. *24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. pp. 215–222. https://doi.org/10.1145/3219819.3219878

Debener, J., Heinke, V., & Kriebel, J. (2023). Detecting insurance fraud using supervised and unsupervised machine learning. *Journal of Risk and Insurance Review*, 1– 26. https://doi.org/10.1111/jori.12427

Degirmenci, Soysal, and Chris Jones. (2022). Benchmarking offline reinforcement learning algorithms for e-commerce order fraud evaluation. *arXiv preprint*, arXiv:2212.02620.

Dhankhad, S., Mohammed, E., and Far, B. (2018). Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study. *IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 122-125, doi: 10.1109/IRI.2018.00025.

Goud, V. S., and Premchand, P. (2019). Enhanced Hidden Markov Model For Credit Card Fraud Detection. *Complexity International*, 23(2).

Shukur, H.A., and Kurnaz, S. (2019). Credit Card Fraud Detection using Machine Learning Methodology. *International Journal of Computer Science and Mobile Computing*, 8(3), pp. 257-260.

Xu, H., Fan, G., and Song, Y. (2022). Application Analysis of the Machine Learning Fusion Model in Building a Financial Fraud Prediction Model, *Security and Communication Networks*. https://doi.org/10.1155/2022/8402329

Hussain, S.K.S., Reddy, E S.C., Akshay, K.G., Akanksha, T. (2021). Fraud Detection in Credit Card Transactions Using SVM and Random Forest Algorithms, *Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, pp. 1013-1017, doi: 10.1109/I-SMAC52330.2021.9640631.

Islam, M.M., Ripan, R.C., et al. (2021). Intelligent Computing & Optimization. *ICO 2021 Lecture Notes in Networks and Systems*, *Springer,* vol 371. pp 217–226. https://doi.org/10.1007/978-3-030-93247-3_22

Itoo, F., Meenakshi and Singh, S.(2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *Int. journal. inf. tecnol*. 13(1), 1503–1511. https://doi.org/10.1007/s41870-020-00430-y

Ivanyuk, V. (2023). Forecasting of digital financial crimes in Russia based on machine learning methods. *Journal of Computer Virology and Hacking Techniques*, 19(2). https://doi.org/10.1007/s11416-023-00480-3

Joshi, Abhishek et al. (2021). An Experimental Study using Unsupervised Machine Learning Techniques for Credit Card Fraud Detection, *GIS Science Journal*, 8(5), pp. 1187-1206.

Kanjanawattana, S. (2019). A novel outlier detection applied to an adaptive k-means. *International Journal of Machine Learning and Computing*, 9(5), pp. 569-574.

Kaur, P., Sharma, A., Chahal, J. K., Sharma, T., and Sharma, V. K. (2021), Analysis on Credit Card Fraud Detection and Prevention using Data Mining and Machine Learning Techniques, *International*

Conference on Computational Intelligence and Computing Applications (ICCICA), pp. 1-4, doi: 10.1109/ICCICA52458.2021.9697172.

Khaled Gubran Al-Hashedi, Pritheega Magalingam, (2021). Financial fraud detection applying data mining techniques: A comprehensive review, *Computer Science Review*, vol 40. https://doi.org/10.1016/j.cosrev.2021.100402

Khosravi, S., Kargari, M., Teimourpour, B., Eshghi, A. and Aliabdi, A. (2023). Using Supervised Machine Learning Approaches To Detect Fraud In The Banking Transaction Network. *9th International Conference on Web Research (ICWR)*, pp. 115-119, doi: 10.1109/ICWR57742.2023.10139083.

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*, Department of Computer Science, University of Durham.

Kominfo. (2023). Tantangan Era Digital: Santernya Kasus Fraud di Industri Jasa Keuangan Indonesia. *Pandu Kominfo*, [Online] https://pandu.kominfo.go.id/blog/479 [Acessed: 10 May 2023].

Lokanan, M.E., Sharma, K. (2022). Fraud prediction using machine learning: The case of investment advisors in Canada*, Machine Learning with Applications*, vol 8. https://doi.org/10.1016/j.mlwa.2022.100269

Mehbodniya, A., et al. (2021). Financial Fraud Detection in Healthcare Using Machine Learning and Deep Learning Techniques. *Security and Communication Networks Journal*, vol 2021, https://doi.org/10.1155/2021/9293877.

Mishra, K.N., Pandey, S.C. (2021). Fraud Prediction in Smart Societies Using Logistic Regression and k-fold Machine Learning Techniques. *Wireless Pers Commun,* vol 119, pp. 1341–1367. https://doi.org/10.1007/s11277-021-08283-9

Mittal, S. and Tyagi, S. (2019) Performance Evaluation of Machine Learning Algorithms for Credit Card Fraud Detection, *9th International Conference on Cloud Computing, Data Science & Engineering (Confluence),* pp. 320-324, doi: 10.1109/CONFLUENCE.2019.8776925.

Mohamed, A. H. A. and Subramanian, S., (2023). Fraud Classification In Financial Statements Using Machine Learning Techniques. *International Conference on IT Innovation and Knowledge Discovery (ITIKD)*. pp. 1-4, doi: 10.1109/ITIKD56332.2023.10100257.

Mohmad, Y.A., (2022). Credit Card Fraud Detection Using LSTM Algorithm. *Wasit Journal of Computer and Mathematic Science*, 1(3), pp. 26–35. https://doi.org/10.31185/wjcm.60

Mongwe W. T., and Malan, K. M. (2020). The Efficacy of Financial Ratios for Fraud Detection Using Self Organising Maps, *IEEE Symposium Series on Computational Intelligence (SSCI),* pp. 1100-1106, doi: 10.1109/SSCI47803.2020.9308602.

Mubalaike, A. M. and Adali (2018). Deep Learning Approach for Intelligent Financial Fraud Detection System*, 3rd International Conference on Computer Science and Engineering (UBMK)*, pp. 598-603, doi: 10.1109/UBMK.2018.8566574.

Mustika, N. I., Nenda, B., Ramadhan, D. (2021). Machine Learning Algorithms in Fraud Detection: Case Study on Retail Consumer Financing Company, *Asia Pacific Fraud Journal*. 6(2), pp. 213-221. http://dx.doi.org/10.21532/apfjournal.v6i2.221

Nalluri, V., Chang, JR., Chen, LS. et al. (2023). Building prediction models and discovering important factors of health insurance fraud using machine learning methods. *Journal Ambient Intell Human Comput,* 14(1), pp. 9607–9619. https://doi.org/10.1007/s12652-023-04633-6

Pani, A. K., Kumar, P. (2021). An Approach for Detecting Frauds in E-Commerce Transactions using Machine Learning Techniques, *International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 826-831, doi: 10.1109/ICOSEC51865. 2021.9591720.

Rai A. K. and Dwivedi R. K. (2020). Fraud Detection in Credit Card Data using Unsupervised Machine Learning Based Scheme, *International Conference on Electronics and Sustainable Communication Systems (ICESC),* pp. 421-426, doi: 10.1109/ICESC48915.2020.9155615.

Sehrawat, D., Singh, Y. (2023) Auto-Encoder and LSTM-Based Credit Card Fraud Detection. *SN Computer Science*. 4(557). https://doi.org/10.1007/s42979-023-01977-w.

Severino, Matheus & Peng, Yaohao. (2021). Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata. *Machine Learning with Applications*. 5 (1), https://doi.org/10.1016/j.mlwa.2021.100074.

Shen, Y., Guo, C., et al. (2021). Financial Feature Embedding with Knowledge Representation Learning for Financial Statement Fraud Detection, *Procedia Computer Science*, vol. 187, pp. 420-425, https://doi.org/10.1016/j.procs.2021.04.110.

Siddharth, V., Kayathwal and Wadhwa, H and Dhama, (2021). Application of Deep Reinforcement Learning to Payment Fraud, *arXiv.* vol 2112. https://doi.org/10.48550/arXiv.2112.04236

Fulmer, L. (2015). Card Fraud Losses Reach $21.84 Billion in 2015. *The Nilson Report*. [Online]. Available: https://www.prweb.com/releases/creditcardfraud/2015/prweb13791784.htm. [Accessed: 31-May-2023].

Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S. and Kuruwitaarachchi, N. (2019). Real-time Credit Card Fraud Detection Using Machine Learning*. 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 488-493, doi: 10.1109/CONFLUENCE.2019.8776942.

Trivedi, Naresh & Simaiya, Sarita & Kumar Lilhore, Dr & Sharma, Sanjeev. (2020). An Efficient Credit Card Fraud Detection Model Based on Machine Learning Methods. *International Journal of Advanced Science and Technology,* 29(5), pp. 3414 - 3424.

Valavan, M. & Rita, S.. (2023). Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers. *Computer Systems Science and Engineering*. 45 (1), pp. 231-245. doi: 10.32604/csse.2023.026508.

Verma, P. and Tyagi, P., (2022), Analysis of Supervised Machine Learning Algorithms in the Context of Fraud Detection*, ECS Transactions*, 107(1), pp. 7189. doi: 10.1149/10701.7189ecst

Vuppula, K. (2021). An advanced machine learning algorithm for fraud financial transaction detection, *Journal For Innovative Development in Pharmaceutical and Technical Science (JIDPTS)*, 4(9), pp. 73-78.

Wahono, R.A. (2015). A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks*, Journal of Software Engineering*, 1(1), pp. 1-16.

Wahyudi, I., Boedi, S., Kadir, A. (2023). Kecurangan Laporan Keuangan (Fraudulent) Sektor Tambang di Indonesia, *Jurnal KRISNA: Kumpulan Riset Akuntansi*, 13(2), pp. 180-190.

Xu, H., Fan, G., Song, Y. (2022). Novel Key Indicators Selection Method of Financial Fraud Prediction Model Based on Machine Learning Hybrid Mode, *Mobile Information Systems*, vol. 2022, https://doi.org/10.1155/2022/6542652

Y. Chen and X. Han, (2021). CatBoost for Fraud Detection in Financial Transactions*, IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 176-179, doi: 10.1109/ICCECE51280.2021.9342475.

Yao, J., Zhang, J., and Wang, L. (2018). A financial statement fraud detection model based on hybrid data mining methods, *International Conference on Artificial Intelligence and Big Data (ICAIBD),* vol. 2018, pp. 57-61, doi: 10.1109/ICAIBD.2018.8396167.

Zamachsari, F., Puspitasari. (2021). Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik, *JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(2). pp. 203 – 212. https://doi.org/10.29207/resti.v5i2.2952

Zhu, Z., Ao, X., et al. (2021). Intelligent Financial Fraud Detection Practices in Post-Pandemic Era: A Survey. *The Innovation*. 2(4). https://doi.org/10.1016/j.xinn.2021.100176

Ziwei Yi, Xinwei Cao, et al. (2023). Fraud detection in capital markets: A novel machine learning approach, *Expert Systems with Applications*, 231(1). https://doi.org/10.1016/j.eswa.2023.120760