World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

Analyzing the use of machine learning techniques in detecting fraudulent activities

Joy Nnenna Okolo [1,*], Samuel A. Adeniji [2], Osondu Onwuegbuchi [2] and Samira Sanni [3]

[1] *McComish Department of Electrical Engineering and Computer Science, South Dakota State University, Brookings, South Dakota, United States.*

[2] *Department of Computer and Information Science, Western Illinois University, Macomb, Illinois, United States.*

[3] *Department of Technology and Industrial Management, University of Central Missouri, Warrensburg, Missouri.*

World Journal of Advanced Research and Reviews, 2025, 26(01), 1198-1209

Publication history: Received on 26 February 2025; revised on 07 April 2025; accepted on 09 April 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.26.1.1097

Abstract

Fraudulent activities have become a growing concern across industries such as finance, e-commerce, healthcare, and cybersecurity, necessitating the adoption of advanced detection mechanisms. Traditional rule-based fraud detection methods are increasingly ineffective in countering the evolving strategies of fraudsters. This paper explores the role of machine learning (ML) techniques in enhancing fraud detection capabilities, leveraging data-driven insights for more accurate and adaptive fraud prevention. The study categorizes ML approaches into supervised, unsupervised, and reinforcement learning methods, each offering distinct advantages in identifying fraudulent patterns. While supervised models rely on labeled datasets for classification, unsupervised techniques excel in detecting anomalies in unlabeled data, and reinforcement learning dynamically refines detection strategies based on real-time feedback. The paper also examines emerging hybrid frameworks that integrate ML with rule-based systems to improve accuracy, interpretability, and scalability. Despite the promise of ML-driven fraud detection, challenges such as data imbalance, model explainability, and regulatory compliance persist. Additionally, advancements in AI, federated learning, and blockchain technology present new opportunities for enhancing fraud detection while ensuring data privacy and security. This conceptual study provides a comprehensive analysis of ML applications in fraud detection, offering insights into current trends, challenges, and future directions for AI-driven fraud prevention strategies.

Keywords: Machine Learning; Artificial Intelligence; Blockchain; Cybersecurity; Federated Learning; Financial Fraud Detection

1. Introduction

Fraudulent activities have become a big threat in several areas, particularly in financial services, e-commerce, healthcare, and cybersecurity [23]. The increased complexity and volume of transactions in digital platforms have made traditional fraud detection approaches less effective. Conventional rule-based techniques rely on established patterns and manual interventions, which are often incapable of adjusting to new and developing fraud tactics [77]. Organizations are looking for more sophisticated solutions to effectively detect and mitigate fraudulent activities as fraudsters continue to use sophisticated techniques. As a result, there is increasing interest in machine learning (ML) as a potent fraud detection technique that uses data-driven insights to more quickly and accurately identify suspicious actions [15:85].

Without the need for explicit programming, machine learning approaches provide the capacity to analyze enormous volumes of transactional data, identify anomalies, and identify fraudulent trends [27]. In contrast to conventional methods, machine learning models are able to generalize to identify new fraud situations by learning from past fraud cases. Because they use labeled datasets to distinguish between fraudulent and genuine transactions, supervised

[*] Corresponding author: Joy Nnenna Okolo.

learning techniques like decision trees, support vector machines (SVM), and neural networks are frequently used in fraud detection [16]. However, when labeled fraud data is scarce, unsupervised learning techniques—such as clustering algorithms and autoencoders—are employed to identify anomalies and odd behavior. By dynamically modifying detection algorithms in response to real-time input, reinforcement learning is also becoming a viable technique for fraud detection [9].

The flexibility of ML-based fraud detection to adjust to novel fraudulent schemes is one of its key benefits [37]. Static rule-based systems eventually lose their effectiveness since fraudsters constantly adapt their methods to avoid detection. Through continuous learning, machine learning models—especially deep learning architectures—are able to modify their predictions and identify intricate patterns in unprocessed data [15]. Furthermore, by utilizing a variety of algorithms to lower false positives and false negatives, ensemble learning techniques—which integrate several machine learning models—improve detection accuracy. These developments assist e-commerce platforms, payment processors, and financial institutions in proactively detecting fraudulent activity and reducing risks before significant losses happen [47].

Despite its potential, there are a number of obstacles to overcome when using machine learning for fraud detection. Since ML models need a lot of high-quality labeled data to function well, data availability and quality continue to be major challenges [62]. Furthermore, fraudulent transactions make up a very small portion of the entire data in fraud datasets, which are frequently extremely unbalanced. This class imbalance can lead to biased models that fail to detect rare fraud cases accurately [17]. Furthermore, explainability and interpretability of ML models represent another hurdle, particularly in regulatory situations where decision-making transparency is critical. Deep learning and other black-box models may be very accurate, but they don't give a clear explanation for their predictions, which raises questions about compliance and accountability [34].

Hybrid fraud detection models that combine rule-based systems and machine learning techniques have been developed as a result of recent developments in AI and ML. By fusing the predictive capabilities of machine learning with the interpretability of conventional systems, these hybrid approaches increase accuracy and transparency [40]. Additionally, companies can use distributed datasets while preserving data security and compliance thanks to the growth of federated learning and privacy-preserving machine learning techniques [4]. Real-time fraud detection systems driven by artificial intelligence will be integrated more frequently across industries as machine learning (ML) develops further, providing more reliable and scalable ways to fight fraud [79].

This study intends to investigate the efficacy of machine learning techniques in fraud detection, given the increasing dependence on digital transactions and the complexity of fraudulent activities [61]. This study aims to offer important insights into the function of AI-driven fraud detection systems by examining different machine learning algorithms, their applications, difficulties, and new developments [83]. The study's conclusions will advance knowledge of how machine learning can improve efforts to prevent fraud, allowing companies and organizations to fortify their defenses against cyber threats and financial crimes [62].

1.1. Research Objectives

The study, being a conceptual paper relying on existing literature and theoretical models, has the following objectives:

- To explore the role of machine learning in fraud detection
- To identify key machine learning techniques and assess their effectiveness in fraud detection
- To discuss conceptual frameworks that can enhance fraud detection using machine learning

2. Theoretical Foundations of Fraud Detection

2.1. Concepts and forms of Fraud

Often at the expense of individuals, businesses, or governments, fraud is a dishonest behavior done to obtain an unfair or illegal benefit. Usually, it entails dishonesty, hiding, or trust abuse to get victims to provide either personal or financial information [1]. From banking to cybersecurity to healthcare to business, fraud can strike any field and changes with technology. The digital era has greatly enlarged the scope and complexity of fraudulent activity, therefore reducing the efficacy of conventional detection techniques [24]. Advanced fraud detection systems like machine learning and artificial intelligence become more important as fraudsters create more complex plans [71].

Financial fraud, in which dishonest behavior meant for financial benefit, is one of the most common forms of fraud. Credit card fraud, insurance fraud, mortgage fraud, and Ponzi schemes—Kasim et al., [44]—are among these include. Often using weaknesses in banking and payment systems, financial fraud can be committed by individuals or organized crime groups. While insurance fraud happens when false claims are filed to get undue compensation, credit card fraud—for example, entails illegal transactions performed using stolen card information—Levi &Soudijn, [50]. Real-time fraud detection systems that examine transaction patterns for anomalies have been adopted since the growing use of digital banking and online transactions raises the financial fraud risks [29].

Identity theft is another significant type of fraud whereby con artists utilize personal data to pass for victims. Unauthorized access to bank accounts, false loan applications, and social security fraud—Ilzan et al., [39]—can follow from this. Phishing attacks, data leaks, or malware infections compromising private personal data including social security numbers, addresses, and banking credentials usually start identity theft [58]. Identity theft is harder to find as deepfake technology and AI-driven social engineering techniques proliferate. Because of illegal acts carried out in their names, victims of identity fraud often suffer significant financial and legal repercussions [73].

Cyber fraud, which includes hacking, phishing, ransomware attacks, and online scams [48], spans a wide spectrum of dishonest behavior carried out over digital networks. Skilled criminals use clever methods to take advantage of system security flaws and fool people into revealing private data. While ransomware strikes lock victims out of their own systems until a ransom is paid, phishing scams—for example, involve false emails or messages meant to steal login credentials— [58]. Cyber fraud is now a major threat to people and companies both since financial transactions and personal contacts depend on digital platforms more and more. Although governments and companies are always improving cybersecurity policies, fraudsters always change their methods, thus artificial intelligence-driven fraud detection and prevention techniques are becoming more and more important [41].

3. Material and Methods

This study adopts a conceptual research approach, relying on extensive review and analysis of existing literature to explore the role of machine learning (ML) techniques in detecting fraudulent activities. The objective was to synthesize key developments, identify effective ML approaches, and highlight conceptual frameworks used in fraud detection.

3.1. Traditional Approaches vs. Machine Learning in Fraud Detection

Traditionally, rule-based systems—which run by specifying a set of predefined criteria or rules to flag dubious transactions—have been the foundation of fraud detection. Applying rigorous if-then conditions to identify abnormalities, these systems are built on expert knowledge and historical fraud trends [37]. In banking, for instance, a rule-based system might reject a transaction if it comes from an odd source or amounts more than a certain limit. These systems struggle with changing fraud techniques even if they offer a methodical and understandable way to identify fraud. In contrast, machine learning (ML)-based adaptive models can analyze vast amounts of data, learn from past fraud cases, and adjust their detection strategies dynamically. ML algorithms, unlike rule-based systems, focus on statistical patterns and predictive analytics instead of fixed conditions to more precisely detect fraudulent activity [21].

One of the main drawbacks of rule-based fraud detection is its incapacity to change with new, unanticipated fraud tactics. Because fraudsters always change their strategies, rule-based systems find it challenging to keep up [11]. These systems need continuous updates and human intervention to change policies as fresh fraud trends show themselves. Furthermore, because of their strict criteria, rule-based methods often result in large false-positive rates, therefore identifying legal transactions as fraudulent [64]. Financial institutions and companies have to manually analyze flagged transactions, therefore raising costs and delays in transaction processing, which influences not only customer experience but also operational inefficiencies [64].

Conversely, by using adaptive learning approaches, machine learning models get above these restrictions. Supervised learning algorithms, including as decision trees, support vector machines (SVM), and neural networks, can be trained on huge datasets to distinguish between fraudulent and genuine transactions [33]. Unsupervised learning methods, such clustering algorithms and autoencoders, identify abnormalities without relying on labeled fraud data, making them effective for detecting new and unknown fraud schemes [12]. Additionally, reinforcement learning allows models to change their fraud detection tactics in real time based on developing risks. By continuously learning from new data, ML-based fraud detection systems offer improved accuracy, minimizing false positives while improving fraud detection rates (Faisa et al., 2024).

Despite its advantages, machine learning-based fraud detection still offers problems, such as data reliance, model interpretability, and computing costs [74]. ML models require big, high-quality datasets to train well, yet fraud detection data is generally imbalanced, with fraudulent transactions representing only a small percentage of total transactions [16]. Furthermore, although very successful, deep learning models sometimes function as "black boxes," which makes it challenging to explain their decision-making process. In regulatory settings where explainability is essential, this lack of openness can raise questions. Nevertheless, hybrid fraud detection strategies combining machine learning models with rule-based techniques are becoming increasingly interesting since they provide interpretability and adaptability to fight fraud more effectively [78].

3.2. Machine Learning Techniques in Fraud Detection

3.2.1. Overview of ML Techniques:

Modern data-driven applications are essentially based on machine learning (ML) methods; one of the most often used paradigms is supervised learning. A pillar of binary classification, logistic regression is computationally efficient and interpretable but implies linearity in correlations between characteristics and outcomes, therefore restricting its efficacy for complicated datasets [67]. Although they offer a clear method for representing hierarchical decisions, decision trees are prone to overfitting—especially in noisy or high-dimensional data. By optimizing margins between classes, support vector machines (SVMs) provide stable performance; nonetheless, their computational complexity rises dramatically with increasing datasets [76]. Although they require large volumes of labeled data and significant computer resources, neural networks—especially deep learning models—have shown remarkable skill in handling complex patterns. For smaller-scale tasks, they are less relevant. Therefore, even if supervised learning techniques shine in organized tasks, their relevance mostly relies on the quality and volume of the accessible data [26].

Though they come with their own set of difficulties, unsupervised learning methods handle situations when labeled data is either absent or impractical to acquire [81]. Crucially in sectors like cybersecurity and manufacturing, anomaly detection techniques seek to find rare events or outliers. These approaches, meanwhile, frequently rely on presumptions about usual behavior, which might not necessarily be accurate in dynamic settings. Powerful tools for grouping comparable data items, clustering techniques as K-means and DBSCAN also have drawbacks [20]. K-means requires a fixed number of clusters and assumes spherical clusters, hence violating these assumptions could produce less than ideal results. DBSCAN is susceptible to parameter tuning and suffers with different densities in datasets even if it is more versatile in spotting randomly formed clusters. These negative aspects emphasize the need of choosing suitable methods depending on the particular features of the dataset and the current challenge [82].

With regard to real-time decision-making and fraud pattern adaptation, reinforcement learning (RL) presents great promise since it allows agents to learn optimal behaviors by means of interactions with an environment. By always improving techniques depending on input from prior events, RL systems can dynamically learn changing fraud trends in fraud detection. This dynamic learning capability makes RL particularly suited for combating sophisticated fraud schemes that traditional rule-based systems might fail to detect [42, 43]. However, RL faces several challenges, including the need for extensive training periods and the risk of instability during learning. Additionally, in high-stakes applications like fraud detection, the exploration phase of RL—where the agent tries out various actions—can introduce risks if it leads to incorrect classifications or false positives. Balancing exploration and exploitation remains a critical issue, as overly aggressive exploration could compromise system reliability [10]. Despite its promise, reinforcement learning's application in real-time decision-making contexts is not without limitations. Effective learning may be hampered by the difficulty of creating reward systems that fairly represent desired results since badly crafted rewards could cause unexpected actions [28]. Moreover, to converge to optimal policies, which might not be possible in every situation, RL models can demand big amounts of interaction data and substantial processing resources. For autonomous systems needing instantaneous responses, for example, delays resulting from extended learning phases could prove negative. Therefore, even although reinforcement learning has great potential to increase artificial intelligence capacities, its implementation has to be properly thought out considering the trade-offs between safety, performance, and resource constraints. Future developments in algorithmic efficiency and interpretability will probably improve its applicability over several fields [19, 86].

4. Results and Discussion

4.1. Evaluation Metrics for Fraud Detection:

Analyzing fraud detection models calls for careful evaluation of suitable metrics since the class imbalance in such datasets makes conventional measures like accuracy inadequate [63, 87]. While accuracy gauges the percentage of properly categorized events, it can be deceptive in imbalanced situations whereby the majority class rules. For example, a model that forecasts all transactions as non-fraudulent may have great accuracy but neglect to find any fraudulent events. Focusing on several facets of model performance, precision and recall offer more complex insights [46, 88]. Maintaining user confidence and limiting false alarms depend on precision, which measures the percentage of genuine positives among expected positives thereby guaranteeing that flagged transactions are indeed fraudulent. Conversely, recall emphasizes the need of spotting as many fraudulent transactions as possible to stop financial losses since it gauges the capacity of the model to recognize all true fraud incidents. Particularly helpful when both features need to be enhanced concurrently is the F1-score, which combines precision and recall into a single value using harmonic mean, therefore providing a balanced statistic.

Another crucial statistic for assessing fraud detection systems is ROC-AUC (Receptor Operating Characteristic - Area Under Curve), which offers a whole picture of the trade-off between false positive rate across many classification thresholds and true positive rate (sensitivity). This makes ROC-AUC especially useful in situations when the cost of false positives (incorrectly reported normal transactions) and false negatives (missed fraud) may vary greatly. Though ROC-AUC is useful for comparing models, it does not directly address the particular operational limitations of fraud detection systems, including the tolerance for false positives or the threshold at which decisions are taken [5]. Consequently, choosing evaluation criteria has to strike a compromise between theoretical performance and pragmatic concerns so that the selected metrics fit the company risk tolerance and business goals. In the end, no one statistic can adequately reflect the complexity of fraud detection; so, a multi-metric approach catered to the particular needs of the application is necessary.

4.2. Conceptual Framework for Fraud Detection Using Machine Learning

4.2.1. Proposed Framework Components:

Data collection and preprocessing are fundamental to the suggested fraud detection strategy since they guarantee the efficacy of later phases. Since fraudulent transactions usually make up a small portion of all transactions, handling imbalanced datasets is especially important [84]. Conventional machine learning algorithms do poorly in identifying minority-class instances (fraudulent cases) because they prefer the majority class. By balancing the dataset, techniques like oversampling (like SMOTE) or undersampling can assist reduce this problem, but they also carry some dangers, like overfitting or information loss [70]. Another crucial component is featuring engineering, since the performance of the model is greatly impacted by the selection of pertinent features and the conversion of unstructured input into meaningful representations. The development of engineering features, including transaction frequency, geographic anomalies, or behavioral patterns, which improve the model's capacity to differentiate between authentic and fraudulent activity, is frequently guided by domain knowledge [18].

The framework's core components are model selection and training, where the type of fraud patterns and dataset properties determine which method is best. For example, ensemble techniques such as Random Forests and decision trees work well for capturing intricate correlations in structured data, but they might not work as well with high-dimensional or unstructured datasets [38]. While neural networks—particularly deep learning architectures—are excellent at processing complex patterns, they also demand a significant amount of labeled data and processing power. While anomaly detection methods like Isolation Forests are well-suited for detecting uncommon occurrences, support vector machines (SVMs) provide reliable performance for smaller datasets with distinct class boundaries [69]. Taking into account the particular needs of the fraud detection system, the algorithm selection must strike a balance between computational efficiency, interpretability, and accuracy. Moreover, cross-validation and hyperparameter tuning are crucial procedures to maximize model performance and avoid overfitting, guaranteeing that the chosen model performs well when applied to new data [56].

When the trained model is integrated into operational fraud monitoring systems in real-time, it presents special difficulties that go beyond simple forecast accuracy. Fraud detection systems must process transactions quickly and with minimal latency in real-world situations since delays could lead to large financial losses [54]. This calls for effective deployment techniques, including using edge computing for quick inference or lightweight models. According to Bello et al. [16], the system must also have features that notify stakeholders, when possible, fraud is identified and allow for human intervention to validate cases that have been highlighted and lower false positives. Another important consideration is scalability, since the system must be able to manage growing transaction volumes without experiencing any degradation in performance. The integration process is made more difficult by the need to ensure compatibility with current infrastructure and comply with regulations, which calls for meticulous planning and cooperation between the technical and business domains [55].

The framework's essential elements of ongoing learning and adaptation address the dynamic nature of fraud patterns.

Static models become outdated over time since fraudsters constantly modify their strategies to evade detection systems [68]. Models are kept effective against new threats by routinely updating them with fresh data. Faster adaptation to changing surroundings is made possible by techniques like online learning, which enable models to gradually absorb new knowledge without retraining from scratch. Nevertheless, idea drift—a phenomenon in which the statistical characteristics of the data alter over time and may impair model performance—is a problem associated with continuous learning [72]. To identify and proactively fix such problems, model performance must be regularly monitored and assessed. Furthermore, adding feedback loops improves the system's capacity to learn and get better over time by having human analysts verify predictions and offer more insights. The system guarantees ongoing efficacy in thwarting complex and ever-evolving fraud schemes by incorporating mechanisms for continual development [72].

4.2.2. Ethical and Regulatory Considerations

The ethical considerations in fraud detection models center on questions of fairness and prejudice, which can have a big impact on people and organizations. Unfair treatment of particular groups based on traits like race, gender, or socioeconomic position can result from bias that occurs during data collection, preprocessing, or model training [80]. The algorithm may produce unfair results, for example, if previous transaction data has systematic biases, such as disproportionately flagging transactions from particular demographic groups. Data representation and algorithmic design must be carefully considered in order to ensure fairness in fraud detection [31]. By specifically taking sensitive traits into consideration during training, methods such as adversarial debiasing or fairness-aware learning can help alleviate these problems. Achieving fairness is not simple, though, as various notions of fairness (such as equalized odds and demographic parity) may clash, requiring a balancing act between conflicting ethical values [22].

Because businesses must abide by legislative frameworks controlling data privacy, consumer protection, and financial integrity, regulatory compliance adds another level of complexity to fraud detection systems [13]. Strict guidelines for data handling are enforced by laws like the CCPA and GDPR, which mandate openness in the gathering, processing, and use of personal data. Explainability issues arise in fraud detection because many sophisticated machine learning models, especially deep neural networks, function as "black boxes," making it challenging to defend choices to regulators or impacted parties. In order to preserve model efficacy and safeguard sensitive data, compliance also necessitates strong data protection and anonymization procedures [65]. Additionally, automated systems must not unfairly disfavor any group in accordance with anti-discrimination rules, which increases pressure to make sure fraud detection algorithms are both equitable and accurate. A multidisciplinary approach is necessary to navigate these ethical and regulatory considerations, fusing technical know-how with ethical and legal insights to create reliable and compliant systems [6].

5. Challenges and Future Directions

5.1.1. Challenges

In fraud detection systems, false positives and false negatives pose serious problems because they can both result in high expenses and inefficient operations. False positives happen when valid transactions are mistakenly reported as fraudulent, which can be inconvenient for clients, erode their trust, and possibly result in lost revenue (Hila et al., 2022). Reducing false positives is essential in sectors where the client experience is crucial, like banking and e-commerce. Conversely, false negatives, in which real fraudulent activity is missed, can lead to monetary losses, harm to one's reputation, and legal repercussions. It is difficult to strike a balance between these two kinds of errors because decreasing one frequently results in an increase in the other [47]. Because of this trade-off, sophisticated evaluation criteria like precision-recall curves must be used to maximize model performance in accordance with the organization's unique risk tolerance. Furthermore, decision thresholds can be improved by integrating contextual data and domain expertise, which raises the overall efficacy of the system [57].

Another significant obstacle for fraud detection systems is idea drift or changing fraud trends. Over time, static models lose their effectiveness since fraudsters constantly modify their strategies to take advantage of weaknesses and get beyond current safeguards [2, 89]. Changes in transaction behavior, new fraud schemes, or changes in demographic patterns are just a few examples of how concept drift might appear. Performance may suffer if traditional machine learning models that were trained on historical data find it difficult to generalize to new patterns. Systems need to include methods for ongoing learning and adaptation in order to solve this problem [8]. While ensemble approaches, which integrate predictions from several models, might increase resilience against changing threats, techniques such as online learning enable models to update gradually as new data becomes available. To further guarantee that models stay in line with contemporary fraud dynamics, frequent monitoring and retraining with new datasets are crucial. But putting these ideas into practice adds more operational and computational complexity, necessitating careful resource allocation and planning [16].

One of the biggest obstacles to fraud detection in regulated businesses, where accountability and transparency are crucial, is the interpretability of AI-based conclusions. As "black boxes," many sophisticated machine learning models, such deep neural networks, make it challenging to explain how particular predictions were achieved [83]. Stakeholders, such as regulators, auditors, and end users, are concerned about this lack of interpretability and want explicit explanations for actions that affect people or organizations. By shedding light on important variables affecting predictions and offering insights into model behavior, explainable AI (XAI) approaches like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) present viable remedies. However, these approaches frequently have drawbacks, like higher complexity or decreased accuracy [66]. A careful integration of XAI tools into the fraud detection pipeline and open lines of communication to explain the reasoning behind automated judgments are necessary to ensure interpretability without sacrificing efficiency. Addressing interpretability will continue to be a top priority as fraud detection systems advance in order to foster trust and adhere to legal standards [3].

5.1.2. Future Research Opportunities

Future research on explainable AI (XAI) in fraud detection looks potential as the demand for accountability and transparency in automated decision-making increases. By offering comprehensible justifications for their predictions, recent developments in XAI seek to demystify intricate machine learning algorithms, including deep neural networks [14]. Knowing why a transaction was reported as suspicious is essential in fraud detection, because choices can have serious financial and reputational repercussions. In order to ensure that explanations are correct and useful for stakeholders, future research could concentrate on creating more reliable XAI techniques that are specifically suited to fraud detection scenarios [49, 87]. Furthermore, by enabling human analysts to validate and improve model outputs, XAI integration with real-time systems may increase confidence in automated fraud detection procedures. Investigating hybrid strategies that blend explainable machine learning models with rule-based systems may enhance interpretability and performance even further, opening the door for more transparent and dependable fraud prevention solutions [45].

Blockchain technology and machine learning (ML) combine to provide novel ways to improve fraud prevention systems. The irreversible and decentralized ledger of blockchain technology offers a safe basis for transaction recording, making it more difficult for fraudsters to change or manipulate data [35]. Businesses may use the advantages of both blockchain and machine learning to build robust fraud detection systems. For example, while the blockchain maintains the integrity and traceability of the underlying data, machine learning algorithms can examine blockchain data to find patterns suggestive of fraudulent activity [75]. Future studies might look into creating smart contracts that, using insights from machine learning, automatically sound an alarm or take preventative action. Furthermore, worries about the exposure of sensitive information may be allayed by combining privacy-preserving strategies, like zero-knowledge proofs, with machine learning models that use blockchain data. By developing systems that are not only safe but also extremely responsive to changing risks, this combination of technologies has the potential to completely transform the prevention of fraud [32].

Federated learning, which overcomes the drawbacks of centralized data gathering and processing, is another fascinating development in privacy-preserving fraud detection. Particularly in sectors like finance and healthcare, traditional machine learning techniques frequently call for combining sizable datasets from many sources, which raises privacy and security issues [60]. Federated learning eliminates the need for parties to exchange raw data by enabling models to be trained across decentralized devices or servers that store local data. This method permits collaborative learning while guaranteeing that private data stays in its original setting. Federated learning has the potential to improve overall detection accuracy in fraud detection by enabling the creation of global models that capture a variety of fraud trends across various institutions or areas [59]. Future studies should concentrate on refining federated learning algorithms for fraud-specific applications, tackling problems including model convergence, data heterogeneity, and communication overhead. In an increasingly linked world, investigating hybrid architectures that integrate federated learning with additional privacy-enhancing technologies, including homomorphic encryption, may also improve the security and usefulness of fraud detection systems [52].

6. Conclusion

With its sophisticated skills to detect and stop fraudulent activity in a variety of industries, machine learning has become a revolutionary force in fraud detection. Important realizations show that machine learning (ML) approaches, from supervised learning algorithms like logistic regression and neural networks to unsupervised techniques like clustering and anomaly detection, offer strong instruments for identifying trends and abnormalities in big datasets. These models improve the speed and precision of fraud detection systems by facilitating real-time decision-making. The success of machine learning in this field, however, depends on resolving issues including unbalanced datasets, changing fraud trends, and the requirement for ongoing learning. Organizations might greatly lower financial losses and increase client trust by utilizing complex algorithms and incorporating them into operational frameworks.

The long-term viability of fraud detection systems depends on the creation of strong, moral, and flexible machine learning models. While ethical issues demand fairness, openness, and adherence to regulatory requirements, robustness necessitates that model be able to manage noisy data, class imbalance, and idea drift successfully. In order to provide fair treatment for all users and avoid prejudice against particular groups, bias in data and algorithms must be addressed. Furthermore, as stakeholders call for more justifications for automated judgments, interpretability is essential to fostering trust. Maintaining system efficacy over time requires adaptive models that can learn from fresh data and adapt to new fraud strategies. Developing ML-based solutions that support business goals and societal values requires striking a balance between four factors: robustness, ethics, and adaptability.

Future research fields have a lot of potential to advance machine learning-based fraud detection. Explainable AI (XAI), which promotes greater openness and allows for clearer insights into model predictions, will continue to gain traction. While federated learning offers a privacy-preserving method of collaborative model training without jeopardizing sensitive data, blockchain technology combined with machine learning offers novel possibilities for safe and impenetrable transaction monitoring. These developments have broad ramifications for sectors including finance, insurance, and e-commerce that depend on preventing fraud. ML's influence on fraud detection is expected to grow as it develops further, leading to more reliable, secure, and effective systems. For organizations to keep ahead of increasingly complex fraud schemes, they must continue to be proactive in implementing cutting-edge technologies and processes.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abu Amuna, Y. M., & Abu Mouamer, F. (2020). Impact of applying fraud detection and prevention instruments in reducing occupational fraud: case study: ministry of health in Gaza Strip. International Journal of Advanced Studies of Scientific Research, 4(6).

[2] Adebayo, O. S., Favour-Bethy, T. A., Otasowie, O., & Okunola, O. A. (2023). Comparative Review of Credit Card Fraud Detection using Machine Learning and Concept Drift Techniques. International Journal of Computer Science, 12(2), 24-48.

[3] Adelakun, B. O., Onwubuariri, E. R., Adeniran, G. A., &Ntiakoh, A. (2024). Enhancing fraud detection in accounting through AI: Techniques and case studies. Finance & Accounting Research Journal, 6(6), 978-999.

[4] Aggarwal, M., Khullar, V., & Goyal, N. (2024). A comprehensive review of federated learning: Methods, applications, and challenges in privacy-preserving collaborative model training. Applied Data Science and Smart Systems, 570-575.

[5] Airlangga, G. (2024). Comparative Analysis of Machine Learning Models for Credit Card Fraud Detection in Imbalanced Datasets. Journal of Computer Networks, Architecture and High Performance Computing, 6(2), 858-866.

[6] Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. GSC Advanced Research and Reviews, 18(3), 050-058.

[7] Al-dahasi, E. M., Alsheikh, R. K., Khan, F. A., & Jeon, G. (2024). Optimizing fraud detection in financial transactions with machine learning and imbalance mitigation. Expert Systems, e13682.

[8] Alexandre, D. S. (2024). Fraud Detection Systems Empowered by Context-Awareness: Leveraging Dynamic Machine Learning Techniques (Master's thesis, Universidade NOVA de Lisboa (Portugal)).

[9] Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., ... & Saif, A. (2022). Financial fraud detection based on machine learning: a systematic literature review. Applied Sciences, 12(19), 9637.

[10] Arshad, K., Ali, R. F., Muneer, A., Aziz, I. A., Naseer, S., Khan, N. S., & Taib, S. M. (2022). Deep reinforcement learning for anomaly detection: A systematic review. IEEE Access, 10, 124017-124035.

[11] Ayodeji, I. A. (2024). Fraud Detection and Prevention in the Nigerian Financial Industry (Doctoral dissertation, Walden University).

[12] Bakumenko, A., & Elragal, A. (2022). Detecting anomalies in financial data using machine learning algorithms. Systems, 10(5), 130.

[13] Balakrishnan, A. (2024). International Journal of Computer Trends and Technology.

[14] Belghachi, M. (2023). A Review on Explainable Artificial Intelligence Methods, Applications, and Challenges. Indonesian Journal of Electrical Engineering and Informatics, 11(4), 1007-1024.

[15] Bello, H. O., Ige, A. B., & Ameyaw, M. N. (2024). Adaptive machine learning models: concepts for real-time financial fraud prevention in dynamic environments. World Journal of Advanced Engineering Technology and Sciences, 12(02), 021-034.

[16] Bello, O. A., Folorunso, A., Ejiofor, O. E., Budale, F. Z., Adebayo, K., & Babatunde, O. A. (2023). Machine Learning Approaches for Enhancing Fraud Prevention in Financial Transactions. International Journal of Management Technology, 10(1), 85-108.

[17] Bin Sulaiman, R., Schetinin, V., & Sant, P. (2022). Review of machine learning approach on credit card fraud detection. Human-Centric Intelligent Systems, 2(1), 55-68.

[18] Breskuvienė, D., &Dzemyda, G. (2022). Autoencoder for fraudulent transactions data feature engineering. In DAMSS: 13th conference on data analysis methods for software systems, Druskininkai, Lithuania, December 1–3, 2022. Vilniaus universitetoleidykla.

[19] Chapman, M., Xu, L., Lapeyrolerie, M., & Boettiger, C. (2023). Bridging adaptive management and reinforcement learning for more robust decisions. Philosophical Transactions of the Royal Society B, 378(1881), 20220195.

[20] Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. Symmetry, 15(9), 1679.

[21] Chy, M. K. H. (2024). Proactive Fraud Defense: Machine Learning's Evolving Role in Protecting Against Online Fraud. arXiv preprint arXiv:2410.20281.

[22] Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R., & Goel, S. (2023). The measure and mismeasure of fairness. The Journal of Machine Learning Research, 24(1), 14730-14846.

[23] Dahal, S. B. (2023). Enhancing E-commerce Security: The Effectiveness of Blockchain Technology in Protecting Against Fraudulent Transactions. International Journal of Information and Cybersecurity, 7(1), 1-12.

[24] Daraojimba, R. E., Farayola, O. A., Olatoye, F. O., Mhlongo, N., &Oke, T. T. (2023). Forensic accounting in the digital age: a US perspective: scrutinizing methods and challenges in digital financial fraud prevention. Finance & Accounting Research Journal, 5(11), 342-360.

[25] David, E. (2023). Evaluating the Impact of Metric-based Security Tools on Company Performance and Decision-making (Doctoral dissertation, University of the Cumberlands).

[26] Degen, D., Caviedes Voullième, D., Buiter, S., Hendricks Franssen, H. J., Vereecken, H., González-Nicolás, A., & Wellmann, F. (2023). Perspectives of physics-based machine learning strategies for geoscientific applications governed by partial differential equations. Geoscientific Model Development, 16(24), 7375-7409.

[27] Dhanawat, V. (2022). Anomaly Detection in Financial Transactions using Machine Learning and Blockchain Technology. International Journal of Business Management and Visuals, ISSN: 3006-2705, 5(1), 34-41.

[28] Ekundayo, F. (2024). Reinforcement learning in treatment pathway optimization: A case study in oncology. International Journal of Science and Research Archive, 13(02), 2187- 2205.

[29] Faisal, N. A., Nahar, J., Sultana, N., &Mintoo, A. A. (2024). Fraud Detection In Banking Leveraging Ai To Identify And Prevent Fraudulent Activities In Real-Time. Journal of Machine Learning, Data Engineering and Data Science, 1(01), 181-197.

[30] Fatih, C. (2023). Comparing machine learning algorithms on credit card fraud problem (Doctoral dissertation, Dublin Business School).

[31] Gautam, A. (2023). The evaluating the impact of artificial intelligence on risk management and fraud detection in the banking sector. AI, IoT and the Fourth Industrial Revolution Review, 13(11), 9-18.

[32] George, A. S. (2023). Securing the future of finance: how AI, Blockchain, and machine learning safeguard emerging Neobank technology against evolving cyber threats. Partners Universal Innovative Research Publication, 1(1), 54-66.

[33] Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: a literature review. Education Sciences, 13(12), 1216.

[34] Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. Cognitive Computation, 16(1), 45-74.

[35] Hayadi, B. H., & El Emary, I. M. (2024). Enhancing Security and Efficiency in Decentralized Smart Applications through Blockchain Machine Learning Integration. Journal of Current Research in Blockchain, 1(2), 139-154.

[36] Hilal, W., Gadsden, S. A., &Yawney, J. (2022). Financial fraud: a review of anomaly detection techniques and recent advances. Expert systems With applications, 193, 116429.

[37] Ikemefuna, C. D., Okusi, O., Iwuh, A. C., & Yusuf, S. (2024). Adaptive Fraud Detection Systems: Using MlTo Identify And Respond To Evolving Financial Threats. International Research Journal of Modernization in Engineering Technology and Science, 6(9), 1727-1735.

[38] Ileberi, E., Sun, Y., & Wang, Z. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. Journal of Big Data, 9(1), 24.

[39] Ilzan, A. R., Oktaviani, R. F. B., Yusuf, F. M., Wegman, D. J., & Imtiyaz, N. Y. (2023). Understanding the phenomenon and risks of identity theft and fraud on social media. Asia Pacific Journal of Information System and Digital Transformation, 1(01), 23-32.

[40] Islam, S., Haque, M. M., & Karim, A. N. M. R. (2024). A rule-based machine learning model for financial fraud detection. International Journal of Electrical & Computer Engineering (2088-8708), 14(1).

[41] Jimmy, F. (2021). Emerging threats: The latest cybersecurity risks and the role of artificial intelligence in enhancing cybersecurity defenses. Valley International Journal Digital Library, 564-574.

[42] Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V. (2020). Enhancing Process Automation Using Reinforcement Learning and Deep Neural Networks. International Journal of AI and ML, 1(3).

[43] Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V. (2020). Optimizing Decision- Making with AI-Enhanced Support Systems: Leveraging Reinforcement Learning and Bayesian Networks. International Journal of AI and ML, 1(2).

[44] Kasim, E. S., Md Zina, N., Mohd Padil, H., & Omar, N. (2020). Ponzi schemes and its prevention: Insights from Malaysia. Management & Accounting Review (MAR), 19(3), 89-118.

[45] Kaushik, K., Pavithra, L. K., & Subbulakshmi, P. (2024). 3 Applications of XAI in. Explainable, Interpretable, and Transparent AI Systems, 31.

[46] Khalid, A. R., Owoh, N., Uthmani, O., Ashawa, M., Osamor, J., &Adejoh, J. (2024). Enhancing credit card fraud detection: an ensemble machine learning approach. Big Data and Cognitive Computing, 8(1), 6.

[47] Khurana, R. (2020). Fraud detection in ecommerce payment systems: The role of predictive ai in real-time transaction security and risk management. International Journal of Applied Machine Learning and Computational Intelligence, 10(6), 1-32.

[48] Kumar, S. (2023). Cyber Crime: A Review. International Journal of Advanced Scientific Innovation, 5(12).

[49] Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a critical review. IEEE access, 9, 82300-82317.

[50] Levi, M., &Soudijn, M. (2020). Understanding the laundering of organized crime money. Crime and Justice, 49(1), 579-631.

[51] Louati, F., Ktata, F. B., & Amous, I. (2024). Enhancing Intrusion Detection Systems with Reinforcement Learning: A Comprehensive Survey of RL-based Approaches and Techniques. SN Computer Science, 5(6), 665.

[52] Manzoor, H. U., Shabbir, A., Chen, A., Flynn, D., & Zoha, A. (2024). A survey of security strategies in federated learning: Defending models, data, and privacy. Future Internet, 16(10), 374.

[53] Mill, E. R., Garn, W., Ryman-Tubb, N. F., & Turner, C. (2023). Opportunities in real time fraud detection: An explainable artificial intelligence (XAI) research agenda. International Journal of Advanced Computer Science and Applications, 14(5), 1172-1186.

[54] Mohammad, N., Prabha, M., Sharmin, S., Khatoon, R., & Imran, M. A. U. (2024). Combating banking fraud with it: integrating machine learning and data analytics. The American Journal of Management and Economics Innovations, 6(07), 39-56.

[55] Mohammed Abdul, S. S. (2024). Navigating blockchain's twin challenges: Scalability and regulatory compliance. Blockchains, 2(3), 265-298.

[56] Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Overfitting, model tuning, and evaluation of prediction performance. In Multivariate statistical machine learning methods for genomic prediction (pp. 109-139). Cham: Springer International Publishing.

[57] Mutemi, A., &Bacao, F. (2024). E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review. Big Data Mining and Analytics, 7(2), 419-444.

[58] Nadeem, M., Zahra, S., Abbasi, M., Arshad, A., Riaz, S., & Ahmed, W. (2023). Phishing Attack, Its Detections and Prevention Techniques. International Journal of Wireless Security and Networks, 1(2), 13-25p.

[59] Nguyen, D. C., Ding, M., Pham, Q. V., Pathirana, P. N., Le, L. B., Seneviratne, A., ... & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. Internet of Things Journal, 8(16), 12806-12825.

[60] Nicolazzo, S., Arazzi, M., Nocera, A., & Conti, M. (2024). Privacy-Preserving in Blockchain-based Federated Learning Systems. arXiv preprint arXiv:2401.03552.

[61] Njoku, D. O., Iwuchukwu, V. C., Jibiri, J. E., Ikwuazom, C. T., Ofoegbu, C. I., & Nwokoma, F. O. (2024). Machine learning approach for fraud detection system in financial institution: a web base application. Machine Learning, 20(4), 01-12.

[62] Obeng, S., Iyelolu, T. V., Akinsulire, A. A., & Idemudia, C. (2024). Utilizing machine learning algorithms to prevent financial fraud and ensure transaction security. World Journal of Advanced Research and Reviews, 23(1), 1972-1980.

[63] Olushola, A., & Mart, J. (2024). Fraud Detection using Machine Learning. ScienceOpen Preprints.

[64] Owoade, S. J., Uzoka, A., Akerele, J. I., & Ojukwu, P. U. (2024). Automating fraud prevention in credit and debit transactions through intelligent queue systems and regression testing. International Journal of Frontline Research in Science and Technology, 4(1), 45- 62.

[65] Patel, S., & Rahman, A. (2024). Data Privacy in the Digital Age: Navigating Compliance and Ethical Challenges. Baltic Multidisciplinary Research Letters Journal, 1(3), 13-24.

[66] Prakash, V., &Deokar, R. (2025). Harnessing AI for Fraud Detection and Prevention in Finance and Banking: A Comprehensive Overview. Real-World Applications of AI Innovation, 389-406.

[67] Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Machine learning and deep learning for big data analytics: A review of methods and applications. Partners Universal International Innovation Journal, 2(3), 172-197.

[68] Rohilla, A. (2024). Strengthening Financial Resilience: A Holistic Approach to Combatting Fraud. Indian Journal of Economics and Finance (IJEF), 4(1), 20-31.

[69] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2(3), 160.

[70] Sharief, F., Ijaz, H., Shojafar, M., & Naeem, M. A. (2024). Multi-Class Imbalanced Data Handling with Concept Drift in Fog Computing: A Taxonomy, Review, and Future Directions. ACM Computing Surveys, 57(1), 1-48.

[71] Sharma, R., Mehta, K., & Sharma, P. (2024). Role of Artificial Intelligence and Machine Learning in Fraud Detection and Prevention. In Risks and Challenges of AI-Driven Finance: Bias, Ethics, and Security (pp. 90-120). IGI Global.

[72] Shoetan, P. O., Oyewole, A. T., Okoye, C. C., &Ofodile, O. C. (2024). Reviewing the role of big data analytics in financial fraud detection. Finance & Accounting Research Journal, 6(3), 384-394.

[73] Sonowal, G., & Sonowal, G. (2022). Introduction to phishing. Phishing and Communication Channels: A Guide to Identifying and Mitigating Phishing Attacks, 1-24.

[74] Sontakke, D. (2023). Fraud Detection in Insurance: A Data-Driven Approach Using Machine Learning Techniques. Journal of Science & Technology, 4(1), 66-88.

[75] Tatineni, S. (2020). Enhancing Fraud Detection in Financial Transactions using Machine Learning and Blockchain. International Journal of Information Technology and Management Information Systems, 11(1), 8-15.

[76] Thomas, R., Sujithra, M., & Senthilkumar, B. (2025). The Role of AI and ML in Shaping Predictive Analytics for Modern Business Intelligence: Techniques, Challenges, and Applications for Data-Driven Decision-Making. AI-Powered Business Intelligence for Modern Organizations, 51-78.

[77] Van Duc, N., Chau, T. T. M., Long, P. H., Nhung, L. T. C., Huy, B. Q., Bin, Z., & Yusof, A. F. B. H. (2024). Modernizing Taxation, Fraud Detection, and Revenue Management in Public Institutions Using AI-Driven Approaches.

[78] Varga, G. (2024). Data-Driven Methods for Machine Learning-Based Fraud Detection and Cyber Risk Mitigation in National Banking Infrastructure. Nuvern Machine Learning Reviews, 1(1), 33-40.

[79] Vyas, B. (2023). Java in Action: AI for Fraud Detection and Prevention. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 58-69.

[80] Yalamati, S. (2023). Identify fraud detection in corporate tax using Artificial Intelligence advancements. International Journal of Machine Learning for Sustainable Development, 5(2), 1-15.

[81] Yan, J., & Wang, X. (2022). Unsupervised and semi-supervised learning: The next frontier in machine learning for plant systems biology. The Plant Journal, 111(6), 1527-1538.

[82] Zangana, H. M., & Abdulazeez, A. M. (2023). Developed Clustering Algorithms for Engineering Applications: A Review. International Journal of Informatics, Information System and Computer Engineering, 4(2), 147-169.

[83] Zanke, P. (2023). AI-Driven fraud detection systems: a comparative study across banking, insurance, and healthcare. Advances in Deep Learning Techniques, 3(2), 1-22.

[84] Zhu, M., Zhang, Y., Gong, Y., Xu, C., & Xiang, Y. (2024). Enhancing Credit Card Fraud Detection A Neural Network and SMOTE Integrated Approach. arXiv preprint arXiv:2405.00026.

[85] David, A. A., & Edoise, A. (2025). Cloud computing and Machine Learning for Scalable Predictive Analytics and Automation: A Framework for Solving Real-world Problem.

[86] Temitope, S. A., & Onah, l. k. (2024). Exploring machine learning algorithms for automating complex processes in building landscape architecture.

[87] Olowu, O., Adeleye, A. O., Omokanye, A. O., Ajayi, A. M., Adepoju, A. O., Omole, O. M., &Chianumba, E. C. (2024). AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity.

[88] Ariyibi, K. O., Bello, O. F., Ekundayo, T. F., &Ishola, O. (2024). Leveraging Artificial Intelligence for enhanced tax fraud detection in modern fiscal systems.

[89] Adeusi, O. C., Adebayo, Y. O., Ayodele, P. A., Onikoyi, T. T., Adebayo, K. B., &Adenekan, I. O. (2024). IT standardization in cloud computing: Security challenges, benefits, and future directions. World Journal of Advanced Research and Reviews, 22(3), 2050-2057.