ASEAN Journal of
Economic and Economic Education

Journal homepage: https://ejournal.bumipublikasinusantara.id/index.php/ajeee

Advancing Machine Learning for Financial Fraud Detection: A Comprehensive Review of Algorithms, Challenges, and Future Directions

*H. M. M. N. Herath**

Department of Accountancy, University of Kelaniya, Sri Lanka
*Correspondence: E-mail: malindah@kln.ac.lk

ABSTRACT

The increasing prevalence of credit card fraud in digital financial transactions necessitates more advanced detection techniques beyond traditional rule-based methods. This study explores the application of machine learning (ML) algorithms in fraud detection, highlighting the effectiveness of models such as Random Forest, LightGBM, and Artificial Neural Networks in identifying fraudulent activities with high accuracy and recall. Additionally, the research examines the role of feature engineering, data augmentation (SMOTE, K-CGAN), and ensemble learning in enhancing fraud detection capabilities. Key challenges, including data privacy concerns, regulatory compliance, and the evolving nature of fraud tactics, are discussed, along with proposed solutions such as federated learning and adaptive fraud detection models. The study also identifies future research directions, emphasizing the integration of deep learning, reinforcement learning, and cross-domain data sources for more effective fraud prevention. By leveraging advanced ML techniques, financial institutions can improve fraud detection accuracy, reduce false positives, and enhance the security of digital transactions while ensuring compliance with privacy regulations.

ARTICLE INFO

Article History:
Submitted/Received 27 Nov 2024
First Revised 30 Dec 2024
Accepted 27 Feb 2025
First Available online 28 Feb 2025
Publication Date 01 Mar 2025

Keyword:
Credit card fraud detection,
Data augmentation,
Deep learning,
Fraud prevention,
Machine learning.

1. INTRODUCTION

Credit card fraud has become a major concern in the financial industry, with the increasing adoption of digital transactions. Traditional fraud detection methods, such as rule-based systems, have limitations in handling complex fraud patterns. As fraudulent activities evolve, machine learning (ML) techniques have emerged as an effective approach for detecting anomalies in financial transactions. This study explores the effectiveness of ML algorithms in credit card fraud detection, highlighting key advancements in the field.

The integration of ML techniques has revolutionized fraud detection by enabling the analysis of large-scale transactional data in real time. Unlike rule-based methods that rely on predefined patterns, ML models learn from historical data to detect new fraudulent behaviors. Various supervised and unsupervised learning approaches, including Random Forest, Support Vector Machine (SVM), Logistic Regression, Decision Trees, LightGBM, and XGBoost, have been successfully applied to fraud detection (Mohsen *et al*., 2023). These algorithms have demonstrated their ability to enhance fraud detection accuracy by identifying intricate patterns within transaction datasets, which would be difficult to capture using traditional techniques (Maniraj *et al*., 2019).

Traditional fraud detection systems primarily rely on manually defined rules and statistical methods to identify suspicious activities. While these methods provide a foundational approach to fraud detection, they suffer from several limitations. High false positives and false negatives are common, as rule-based systems often incorrectly flag legitimate transactions while failing to detect sophisticated fraud patterns (Al-Hashedi & Magalingam, 2021). These systems also lack adaptability, making them ineffective against new fraud schemes that continuously evolve (Chaquet-Ulldemolins *et al*., 2022). Scalability issues arise due to the heavy reliance on human intervention, which makes real-time fraud detection impractical in large-scale financial transactions (Da'U & Salim, 2019). The problem of imbalanced data further complicates fraud detection, as fraudulent transactions make up only a small fraction of total transactions, leading to biased models that struggle to identify rare fraudulent activities (Choi & Lee, 2018).

Machine learning techniques address these challenges by learning patterns from large datasets and adapting to new fraud strategies dynamically. Through advanced methodologies such as feature engineering and anomaly detection, ML models significantly improve the accuracy and efficiency of fraud detection systems (Bello *et al*., 2023).

Machine learning models for fraud detection can be broadly classified into supervised learning models, unsupervised learning models, and ensemble learning approaches. Supervised learning models require labeled transaction data to train classification algorithms. Among the most commonly used models, Random Forest is recognized for its ability to handle imbalanced datasets and reduce overfitting, making it an effective fraud detection tool (Aghware *et al*., 2024). Support Vector Machine (SVM) is another widely used algorithm that separates fraudulent and legitimate transactions using optimized hyperplanes, ensuring accurate classification (Mohsen *et al*., 2023). Logistic Regression remains a popular choice for binary classification tasks, while Artificial Neural Networks (ANNs) excel at capturing complex transaction patterns, offering a deep-learning approach to fraud detection (Hossain *et al*., 2022).

Unsupervised learning models identify anomalies without requiring labeled fraud data. Autoencoders are neural networks trained to reconstruct normal transactions, with high reconstruction errors serving as an indicator of fraud (Pitsane, 2022). Clustering algorithms

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

such as k-means and DBSCAN have also been applied to group transactions based on similarity, helping detect suspicious activities without prior labeling (He, 2022).

Ensemble learning and hybrid approaches further enhance fraud detection performance by combining multiple ML models. XGBoost and LightGBM leverage the strengths of multiple algorithms, leading to improved classification accuracy (Patel & Panday, 2023). Studies have shown that integrating these models with anomaly detection techniques significantly boosts fraud detection rates, making them highly effective in real-world applications (Ahmed & Shamsuddin, 2021).

Feature engineering plays a crucial role in improving fraud detection accuracy by selecting and transforming relevant transaction attributes. Studies have highlighted the importance of incorporating behavioral features, such as spending patterns, transaction frequency, and merchant categories, to enhance ML model performance (Mohsen *et al*., 2023). Temporal features, including transaction timestamps and recurrence, provide additional insights into unusual patterns, improving fraud detection capabilities (Aghware *et al*., 2024). Geospatial features, based on location data, have also been successfully used to flag transactions that deviate from a user's normal spending habits (Maniraj *et al*., 2019).

To address the issue of class imbalance, data augmentation techniques such as the Synthetic Minority Oversampling Technique (SMOTE) and Generative Adversarial Networks (GANs) have been applied to generate synthetic fraudulent transactions. SMOTE creates new synthetic samples of fraudulent transactions to balance training datasets, significantly improving model generalization and reducing bias (Ahmed & Shamsuddin, 2021). The introduction of advanced augmentation techniques, such as K-CGAN, has further enhanced fraud detection accuracy by providing synthetic data that closely resembles real fraudulent transactions (Strelcenia & Prakoonwit, 2023). The combination of ML algorithms with data augmentation techniques has demonstrated remarkable success in reducing false negatives and improving fraud detection efficiency (Mienye & Sun, 2023).

As fraudsters continue to develop sophisticated methods, the future of ML-driven fraud detection lies in continuous innovation and adaptation. One major direction is the integration of real-time detection systems, which can analyze transactions within milliseconds to prevent fraud before completion (Patel & Panday, 2023). Explainable AI (XAI) is another critical area of research, aiming to improve model interpretability and enhance trust in AI-driven decision-making, ensuring financial institutions can understand and validate fraud detection decisions (He, 2022).

Graph-based fraud detection, which utilizes network analysis to identify fraud rings and complex transaction structures, has shown promise in uncovering organized fraud patterns (Aghware *et al*., 2024). The use of federated learning, allowing financial institutions to collaborate on fraud detection while preserving user privacy, is another emerging approach that enhances security without compromising sensitive data (Mienye & Sun, 2023).

By incorporating these advancements, financial institutions can stay ahead of emerging fraud threats and enhance security in digital transactions. The combination of ML algorithms, advanced feature selection, and real-time detection strategies ensures a robust defense against evolving fraudulent activities, paving the way for a more secure financial ecosystem (Alarfaj *et al*., 2022).

The research questions formulated in this study serve as a foundation for analyzing the role of ML in fraud detection and identifying gaps in existing literature. The study is guided by the following key questions:

(i) How effective are machine learning techniques in detecting credit card fraud compared to traditional rule-based methods?

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

(ii) What are the most common types of credit card fraud, and which machine learning algorithms demonstrate the highest accuracy in detecting them?
(iii) Can the performance of machine learning models for credit card fraud detection be improved by incorporating additional data sources or feature selection techniques?

By systematically addressing these research questions, this study aims to provide a comprehensive understanding of ML-based fraud detection methodologies, evaluate their effectiveness, and suggest potential improvements for future research and practical applications.

2. METHODS

This study employs a Systematic Literature Review (SLR) to analyze the application of machine learning (ML) techniques in detecting credit card fraud. The SLR methodology is a well-recognized approach for conducting structured reviews in specific research fields (Serenko, 2021; Chen *et al*., 2022; Compagnucci *et al*., 2023). This method ensures a comprehensive and methodical examination of the literature, facilitating an in-depth understanding of how ML techniques operate in identifying fraudulent activities (Kraus *et al*, 2022).

The systematic review process is structured into eight distinct phases, as outlined in **Figure 1** (Kraus *et al*, 2022). These phases guide the identification, selection, and synthesis of relevant studies to provide a robust analysis of the effectiveness of ML models in fraud detection.

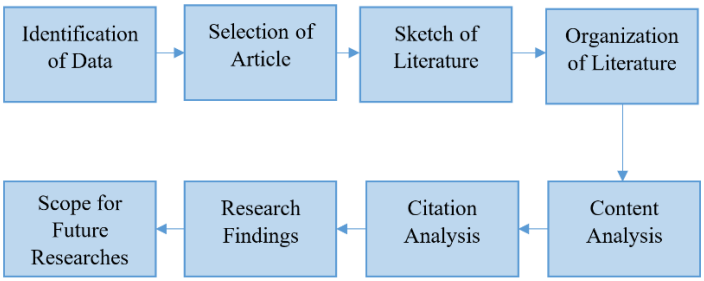

Figure 1. Systematic review process (Kraus *et al*., 2022).

To ensure a thorough examination of existing research, a comprehensive search was conducted across multiple electronic databases. The selected databases—Emerald Insight, IEEE, Google Scholar, and ResearchGate—were chosen due to their extensive repository of peer-reviewed literature on machine learning and fraud detection.

The search queries were constructed using the key term “Machine Learning Techniques for Detection of Credit Card Frauds” to retrieve relevant studies. To maintain the quality and relevance of the literature, only peer-reviewed journal articles published between 2018 and June 2023 were considered.

The selection of studies followed a five-step screening approach to ensure the inclusion of the most relevant and high-quality research:

(i) Initial Search: The first step involved retrieving studies using the key term "Machine Learning Techniques for Detection of Credit Card Frauds" across all databases.
(ii) Time-Based Filtering: Studies published between 2018 and June 2023 were selected to ensure that the research reflects the latest advancements in ML-driven fraud detection.
(iii) Field-Specific Screening: The search was refined using advanced search filters to focus on studies in finance, machine learning, cybersecurity, and fraud detection.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

(iv) Content Screening: Studies were evaluated based on titles, abstracts, and full texts, ensuring they explicitly discuss ML applications for credit card fraud detection.

(v) Final Selection: Only studies accessible to the researchers were included, and irrelevant studies were removed. After this rigorous selection process, a final sample of 24 research articles was chosen for review.

The number of studies retained at each stage of filtering across the databases is summarized in **Table 1**.

Table 1. Sample selection across databases.

Electronic Database	Step 1 (Initial Search)	Step 2 (Time-Based Filtering)	Step 3 (Field-Specific Screening)	Step 4 (Content Screening)	Final Sample
Emerald Insight	12	8	5	3	3
IEEE	8	5	4	2	2
Google Scholar	40	25	15	10	10
ResearchGate	30	20	12	9	9
Total	**90**	**58**	**36**	**24**	**24**

To enhance the transparency and rigor of the review process, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework was applied. PRISMA ensures a structured methodology for selecting, screening, and analyzing relevant studies, mitigating the risk of bias in the selection process.

The PRISMA Flow Diagram (**Figure 2**) illustrates the stepwise approach adopted for study selection, detailing the number of records screened, excluded, and included in the final review. This structured approach guarantees adherence to best practices in systematic reviews and meta-analyses.

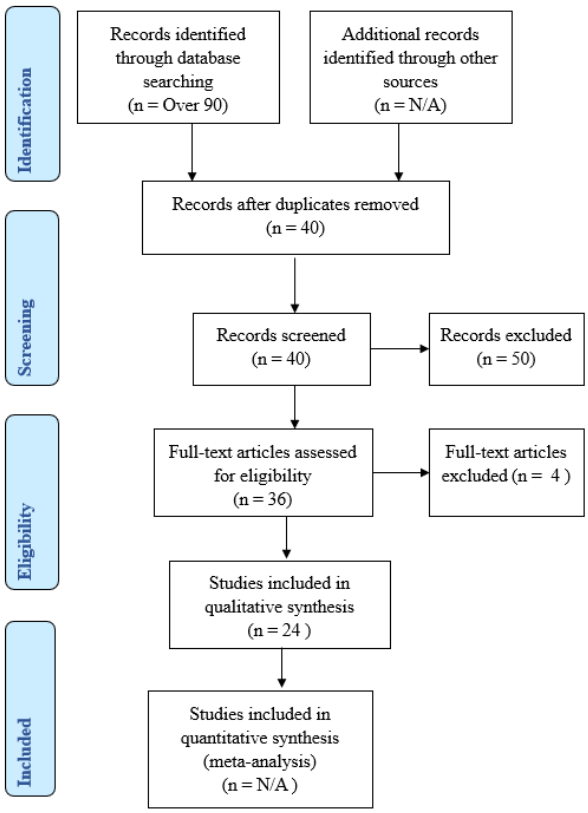

Figure 2. PRISMA flow diagram.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

3. RESULTS AND DISCUSSION

3.1. Effectiveness of Machine Learning Techniques in Fraud Detection

3.1.1. Superiority of machine learning in fraud detection

Machine learning (ML) techniques have revolutionized fraud detection, offering a significant improvement over traditional rule-based system. Traditional methods primarily rely on predefined thresholds and rigid rules, which often result in high false positives and an inability to adapt to evolving fraud patterns (Potla, 2023). As financial transactions become more complex, the limitations of these static systems become apparent. In contrast, ML models analyze vast amounts of data, recognize hidden patterns, and continuously improve their detection capabilities, making them a more dynamic and effective solution (Pan, 2024).

One of the primary advantages of machine learning in fraud detection is its ability to process large datasets efficiently and extract meaningful insights. ML models leverage historical transaction data to identify anomalies and suspicious behaviors in real time. This capability allows financial institutions to mitigate risks more effectively and respond to fraud attempts swiftly. Ilori *et al*. (2024) highlight that adaptive analytical models driven by data can significantly improve the efficiency of fraud detection systems, making them easier to maintain and more objective. In particular, supervised learning techniques, such as Random Forest and Support Vector Machines (SVM), enhance precision by learning from labeled transaction data and making accurate classifications of fraudulent and legitimate transactions (Cho & Lee, 2018).

Another critical benefit of ML-based fraud detection is its capacity to reduce false positives, which is a common challenge with rule-based systems. Traditional fraud detection methods often flag legitimate transactions as suspicious, leading to unnecessary investigations and customer dissatisfaction (Kenyon & Tilton, 2012). ML algorithms minimize these errors by learning intricate transaction behaviors, thereby distinguishing fraudulent activities with greater accuracy. Kumar *et al*. (2022) emphasize that as financial transactions grow in complexity, ML algorithms provide an advanced approach to detecting fraud while reducing operational inefficiencies.

Furthermore, the adaptability of machine learning ensures continuous improvement in fraud detection. Traditional systems struggle to keep up with evolving fraud tactics, whereas ML models can update their detection mechanisms through incremental learning and data-driven optimization (Al-Dahidi *et al*., 2024). By integrating ML with big data technologies, institutions can monitor transactions in real time, allowing for immediate identification and mitigation of fraudulent activities (Ilori *et al*., 2024). This adaptability is essential in today's financial landscape, where fraudsters continually develop new methods to bypass conventional security measures.

3.1.2. Performance of machine learning in fraud detection

Among the various ML techniques used for fraud detection, Random Forest, Support Vector Machines (SVM), Logistic Regression, and Decision Trees have demonstrated exceptional accuracy in identifying fraudulent transactions. These algorithms are particularly effective because they can handle large and complex datasets while maintaining high classification accuracy.

Random Forest is widely regarded as one of the most effective ML models for fraud detection due to its ability to handle high-dimensional data and balance between false positives and false negatives. It operates as an ensemble method, combining multiple decision trees to improve classification performance. Studies have shown that Random Forest is highly robust against overfitting and can effectively classify transactions by analyzing

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

numerous features simultaneously. The ensemble nature of the model allows it to aggregate predictions from different decision trees, reducing the likelihood of misclassification and improving accuracy (Parvin *et al*., 2015).

Support Vector Machines (SVM) have also proven highly effective in fraud detection, especially in scenarios where the data is not linearly separable. SVMs work by identifying an optimal hyperplane that maximizes the margin between different classes, thereby distinguishing fraudulent and legitimate transactions with high precision. This feature is crucial in financial fraud detection, where fraudulent transactions often exhibit subtle differences from legitimate ones. Additionally, SVMs can be enhanced using different kernel functions, allowing them to model complex decision boundaries and improve classification accuracy (Pan, 2024).

Decision Trees are another commonly used ML technique in fraud detection due to their transparency and interpretability. They work by splitting data into subsets based on key features, making them particularly useful for identifying patterns associated with fraudulent activities (Cho & Lee, 2018). The interpretability of Decision Trees is especially valuable in financial settings, where regulatory compliance and transparency in fraud detection models are essential (Mohammed, 2022).

The ensemble approach of Random Forest further enhances fraud detection capabilities by reducing variance and bias in classification. By averaging the outputs of multiple decision trees, Random Forest achieves a higher degree of accuracy and robustness, particularly when dealing with imbalanced datasets, where fraudulent transactions are significantly fewer than legitimate ones (Al-dahasi *et al.*, 2025). This is a crucial advantage over simpler models, which may struggle to correctly classify minority-class instances due to dataset skewness.

3.1.3. Enhancing fraud detection with feature engineering and adaptive learning

The effectiveness of ML algorithms in fraud detection is further amplified by feature engineering and ensemble learning techniques, which optimize model performance by refining the data used for training. Feature engineering involves selecting and transforming raw transaction data into meaningful representations that improve classification accuracy. Behavioral features, such as transaction frequency, spending patterns, and merchant types, provide valuable insights into fraudulent activities (Mohsen *et al*., 2023). Additionally, temporal and geospatial features help detect anomalies by analyzing transaction timestamps and geographic locations, further strengthening fraud detection models (Aghware *et al*., 2024).

Ensemble learning, which combines multiple ML models to improve classification performance, has proven to be a highly effective approach in fraud detection. Gradient Boosting algorithms such as XGBoost and LightGBM leverage the strengths of multiple weak learners, thereby enhancing fraud detection accuracy while minimizing false positives (Patel & Panday, 2023). Studies indicate that ensemble methods outperform individual models by reducing overfitting and increasing robustness against noise in the dataset (Ahmed & Shamsuddin, 2021).

Moreover, real-time processing and adaptive learning mechanisms play a crucial role in modern fraud detection systems. Traditional rule-based models often require manual updates, making them slow to adapt to new fraud patterns. In contrast, ML models equipped with incremental learning can continuously update themselves with new transactional data, ensuring they remain effective against emerging fraud tactics (Al-Dahidi *et al*., 2024). This dynamic learning capability is critical for financial institutions, as fraudsters continuously evolve their methods to bypass detection mechanisms (Pan, 2024).

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

However, despite their effectiveness, ML-based fraud detection systems still face challenges related to interpretability and bias. Many complex ML models, such as deep learning-based approaches, function as "black boxes," making it difficult to explain their decision-making processes (Cho & Lee, 2018). Ensuring regulatory compliance and maintaining customer trust require the development of more explainable AI (XAI) frameworks that provide insights into model decisions. Additionally, addressing bias in ML models is essential to prevent discriminatory outcomes in fraud detection. Biased training data can lead to unfair classification of transactions, impacting specific demographic groups disproportionately (Kenyon & Tilton, 2012).

In conclusion, ML-based fraud detection systems provide significant advantages over traditional rule-based methods through enhanced accuracy, reduced false positives, and real-time adaptability. The integration of feature engineering, ensemble learning, and adaptive learning techniques further strengthens their performance, ensuring they remain effective in the ever-changing landscape of financial fraud. However, continued research is required to enhance model interpretability and mitigate biases, ensuring ML-driven fraud detection remains both accurate and fair in real-world applications.

3.2. Comparison of Machine Learning Algorithms for Credit Card Fraud Detection

3.2.1. Effectiveness of machine learning algorithms in fraud detection

Machine learning algorithms have significantly advanced credit card fraud detection, with some models demonstrating superior accuracy and recall compared to traditional rule-based approaches. Among the most effective algorithms, Random Forest (RF) and LightGBM have consistently achieved high performance in detecting fraudulent transactions due to their ensemble learning capabilities. These models leverage multiple decision trees to enhance predictive accuracy and reduce overfitting, making them ideal for handling imbalanced datasets commonly encountered in fraud detection (Dube & Verster, 2023).

Random Forest has proven particularly robust, with studies reporting accuracy rates of 98.8% (Poojitha & Malathi, 2022) and 99.96% (Rodriguez-Galiano *et al*., 2015). Its high recall rate of 80.22% further highlights its effectiveness in distinguishing fraudulent from legitimate transactions. LightGBM, another ensemble learning technique, has also demonstrated superior performance, especially in large-scale datasets where computational efficiency is crucial (El-Hasani *et al*., 2024). In comparative studies, these models have consistently outperformed Logistic Regression and Decision Trees, which often struggle with complex data patterns and exhibit lower accuracy rates.

Artificial Neural Networks (ANNs) have also been widely utilized in fraud detection, offering 99.96% accuracy (Dube & Verster, 2023). ANN models excel in learning nonlinear relationships in financial transaction data, making them highly effective for detecting anomalies. However, their high computational cost and difficulty in interpretability limit their practical implementation in real-world fraud detection systems.

The effectiveness of Support Vector Machines (SVMs) has also been explored in multiple studies. While SVM models provide reliable classification performance, they do not always outperform ensemble learning techniques like Random Forest or hybrid models that combine multiple algorithms. Mohammed (2022) emphasized that SVM is a strong contender in fraud detection tasks but may require additional feature selection and hyperparameter tuning to reach optimal performance.

These findings indicate that Random Forest and LightGBM consistently outperform traditional models in terms of accuracy and recall. Meanwhile, hybrid models that combine multiple machine learning techniques often yield even greater detection results, highlighting

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

the need for adaptive, ensemble-based approaches in fraud detection (Ogundokun *et al*., 2023).

3.2.2. Addressing class imbalance with SMOTE and K-CGAN

A significant challenge in fraud detection is the class imbalance in datasets, where fraudulent transactions are significantly fewer than legitimate ones. This imbalance can cause machine learning models to favor the majority class, reducing their ability to correctly identify fraud. To mitigate this issue, data balancing techniques such as Synthetic Minority Over-sampling Technique (SMOTE) and K-Conditional Generative Adversarial Network (K-CGAN) have been employed to enhance model performance.

SMOTE is a widely used method that generates synthetic samples of fraudulent transactions, improving model learning and reducing bias. Research by Aghware *et al*. (2024) demonstrated that applying SMOTE increased Random Forest accuracy from 98.02% to 99.19%, highlighting its effectiveness in improving model performance. Similarly, Setiawan *et al*. (2023) found that combining SMOTE with clustering techniques such as HDBSCAN led to better fraud detection results, improving recall and precision scores.

In contrast, K-CGAN represents a more advanced data augmentation approach. Unlike traditional oversampling techniques, K-CGAN generates synthetic fraudulent transactions while preserving the underlying relationships in transaction data, allowing for more realistic data augmentation (Strelcenia & Prakoonwit, 2023). Studies have shown that K-CGAN significantly enhances fraud detection accuracy when integrated with XGBoost and Random Forest models, as it allows classifiers to generalize better to unseen fraudulent patterns.

Cheah *et al*. (2023) reported that combining K-CGAN with ensemble methods reduced false positives and increased the true positive rate, making fraud detection systems more reliable. Similarly, Nayyer *et al*. (2024) and Cheah *et al*. (2023) demonstrated that employing ensemble models alongside SMOTE improved fraud detection rates compared to using SMOTE alone. Their findings suggest that a hybrid approach integrating ensemble learning techniques with data balancing strategies yields the best fraud detection results.

Overall, both SMOTE and K-CGAN significantly enhance fraud detection models by addressing class imbalance. While SMOTE effectively increases model recall and precision, K-CGAN generates more realistic synthetic data, further improving fraud detection systems. The integration of these techniques with ensemble learning models such as Random Forest and XGBoost ensures more accurate fraud detection across large financial datasets.

3.2.3. Feature selection, hyperparameter tuning, and computational efficiency

The performance of machine learning models in fraud detection is not only dependent on the choice of algorithms but also on feature selection and hyperparameter tuning. Selecting the most relevant features significantly improves model accuracy by reducing noise and focusing on the most critical fraud indicators. Studies suggest that feature engineering techniques such as Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) play a crucial role in optimizing fraud detection models (Kumar *et al*., 2022).

Hyperparameter tuning further refines model performance by adjusting learning parameters to optimize accuracy, precision, and recall. For example, tuning the number of estimators in Random Forest or learning rate in LightGBM can drastically impact a model's predictive power. El-Hasani *et al.* (2024) found that hyperparameter-optimized Random Forest models consistently outperformed default models, demonstrating the importance of fine-tuning machine learning algorithms for fraud detection.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Despite their high accuracy, different models exhibit varying computational efficiencies, impacting their real-world deployment. While LightGBM is highly efficient and performs well on large datasets with minimal computational cost, Random Forest and ANN models require significant processing power (Rodriguez-Galiano *et al*., 2015). Studies have shown that ensemble learning models, although highly accurate, require substantial memory and processing resources, making them less suitable for real-time fraud detection in high-volume financial systems (Mohammed, 2022).

Hybrid models combining multiple algorithms often yield superior results but at a higher computational cost. Ogundokun *et al*. (2022) demonstrated that hybrid architectures combining ANN, Random Forest, and XGBoost achieved nearly 100% fraud detection accuracy but required extensive computational resources, limiting their real-world applicability.

Thus, selecting the right balance between accuracy and efficiency is critical for deploying ML-based fraud detection models in financial institutions. While Random Forest and LightGBM provide the best trade-off between accuracy and computational efficiency, ANN models, despite their high accuracy, require more processing power, making them suitable for offline fraud detection rather than real-time applications.

The comparison of machine learning algorithms for credit card fraud detection reveals that Random Forest and LightGBM outperform traditional models in terms of accuracy and recall, while hybrid models incorporating multiple algorithms yield superior detection results. SMOTE and K-CGAN play crucial roles in mitigating class imbalance, enhancing model performance across fraud detection datasets. However, computational efficiency remains a challenge, with models such as ANN requiring high processing power, limiting their real-time applicability.

Future advancements in fraud detection should focus on developing efficient hybrid models that balance accuracy, interpretability, and computational cost. By integrating data balancing techniques, hyperparameter tuning, and feature selection, financial institutions can enhance their fraud detection systems, ensuring more reliable, real-time fraud prevention.

3.3. Role of Additional Data Sources and Feature Engineering

3.3.1. Enhancing fraud detection with additional transaction-related features

The integration of additional transaction-related data into machine learning (ML) models significantly improves the accuracy and robustness of fraud detection systems. Traditional fraud detection models often rely solely on transaction amount, location, and time, which limits their ability to differentiate between fraudulent and legitimate activities. By incorporating behavioral and contextual data, ML algorithms gain a more comprehensive understanding of transaction patterns, allowing for more precise fraud identification.

Behavioral data, including spending patterns, transaction frequency, and user behavior, enables fraud detection models to distinguish normal transactions from fraudulent ones. Studies have shown that historical patterns of fraudulent behavior serve as strong indicators for future fraud detection. ML models trained on extensive historical transaction data achieve significantly higher predictive accuracy (Potla, 2023). Additionally, Hidden Markov Model (HMM)-based behavioral features enhance fraud detection in e-commerce and face-to-face transactions by identifying deviations from normal spending behavior (Hilal *et al*., 2022).

Beyond behavioral data, contextual features, such as the device type, IP address, geolocation, and transaction time, provide additional layers of fraud detection. Fraudsters often attempt to mimic legitimate transactions, but anomalies in contextual features can reveal fraudulent intent. Hossain *al*. (2024) discuss how integrating data analytics with ML techniques enables real-time fraud monitoring, improving detection rates by identifying

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

inconsistencies in transactional context. This adaptability is particularly crucial in rapidly evolving fraud schemes, as ML models must continuously update their learning patterns to detect new fraudulent behaviors.

Furthermore, class imbalance remains a major challenge in fraud detection, as fraudulent transactions constitute only a small fraction of total transactions. Models trained on imbalanced datasets may fail to recognize fraudulent cases effectively, leading to high false-negative rates. To address this, Pitsane *et al*. (2022) suggest that enriching datasets with additional behavioral and contextual features significantly improves fraud detection rates, particularly when combined with ensemble learning techniques. Nayyer *et al*. (2023) further argue that combining multiple ML models in ensemble frameworks mitigates data imbalance and enhances overall fraud detection performance.

These findings underscore the importance of leveraging diverse and comprehensive transaction-related features to improve the precision and recall of fraud detection models. By continuously refining these features, ML-based fraud detection systems become more effective in combating increasingly sophisticated fraud tactics in the financial sector.

3.3.2. Addressing data imbalance and improving model robustness with data augmentation

Handling imbalanced datasets is a critical challenge in fraud detection, as the low occurrence of fraudulent transactions skews ML models toward favouring legitimate transactions. Data augmentation techniques, such as Synthetic Minority Over-sampling Technique (SMOTE) and K-Conditional Generative Adversarial Network (K-CGAN), have been developed to improve model robustness and enhance fraud detection performance.

SMOTE is a widely used oversampling technique that generates synthetic fraudulent transactions based on existing data points. Aghware *et al*. (2024) found that applying SMOTE to fraud detection models increased Random Forest's accuracy from 98.02% to 99.19%, demonstrating the effectiveness of synthetic sample generation in improving model performance. Similarly, Setiawan *et al.* (2023) observed that combining SMOTE with clustering techniques like HDBSCAN resulted in enhanced precision and recall, ensuring a more balanced fraud detection model.

K-CGAN, on the other hand, represents a more advanced data augmentation approach that generates synthetic fraud cases while maintaining realistic transaction patterns. Strelcenia & Prakoonwit (2023) introduced K-CGAN as an alternative to SMOTE, showing that it provides better generalization in fraud detection models by preserving the relationships between transaction attributes. Their findings indicate that models trained with K-CGAN achieve higher true positive rates while reducing false positives, improving the overall efficiency of fraud detection systems.

Cheah *et al*. (2023) further demonstrated that combining K-CGAN with ensemble learning techniques enhances fraud detection accuracy compared to using SMOTE alone. Their study found that ensemble models trained on K-CGAN-generated data performed significantly better in real-world fraud detection scenarios. Nayyer *et al*. (2023) and Cheah *et al*. (2023) also emphasized that integrating data augmentation methods with ensemble techniques leads to superior fraud detection performance, as these approaches address data imbalance while leveraging the strengths of multiple classifiers.

Overall, data augmentation plays a crucial role in improving ML-based fraud detection by balancing datasets and enhancing model adaptability. The combination of SMOTE, K-CGAN, and ensemble methods offers an effective solution to overcoming the challenges posed by class imbalance, ensuring that fraud detection models remain accurate and reliable in diverse financial environments.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

3.3.3. Optimizing model performance with feature engineering techniques

Feature engineering is a vital component of fraud detection, as selecting the right features significantly influences ML model accuracy. Principal Component Analysis (PCA) and Genetic Algorithms (GAs) are two widely used feature selection techniques that optimize fraud detection models by reducing dimensionality and improving classification efficiency.

PCA is a dimensionality reduction technique that transforms high-dimensional transaction data into a smaller set of principal components while preserving most of the original information. This method helps prevent overfitting and reduces computational costs in ML models. Yang *et al*. (2021) demonstrated that applying PCA to fraud detection models improved SVM classification accuracy by effectively reducing noise in the dataset. Similarly, Qu *et al*. (2021) found that combining PCA with the Adaboost algorithm significantly enhanced fraud detection performance, highlighting its effectiveness in reducing the complexity of financial transaction data.

In contrast, Genetic Algorithms (GAs) use heuristic search mechanisms inspired by natural selection to identify the most relevant fraud detection features. Mosa *et al*. (2024) applied GAs to credit card fraud detection models, demonstrating that optimized feature selection reduces computational costs and improves model accuracy. Their study found that GAs effectively eliminates irrelevant features, ensuring that fraud detection models focus on high-impact transaction attributes.

Yang *et al*. (2021) highlighted the benefits of automatic feature extraction techniques, which minimize manual effort and enhance fraud detection accuracy. Their findings suggest that automated feature engineering methods can generate a broader range of behavioral features, significantly improving ML models' ability to detect fraudulent activities. Furthermore, Ghosh Dastidar *et al*. (2020) introduced a feature aggregation and transformation framework that optimizes feature selection, further enhancing the performance of fraud detection models.

The combination of PCA, GAs, and automated feature selection techniques ensures that ML models are trained on highly relevant fraud detection features, improving efficiency and accuracy. Additionally, ensemble feature selection approaches have been shown to address class imbalance and high dimensionality, making them an essential tool for developing robust fraud detection models (AUTO-INSURANCE FRAUD DETECTION, 2020).

These findings underscore the importance of feature engineering in fraud detection, as optimized feature selection improves classification accuracy, enhances computational efficiency, and reduces false positives. By integrating advanced feature selection techniques with ML-based fraud detection models, financial institutions can build more accurate and scalable fraud prevention systems that adapt to evolving fraud tactics.

The role of additional data sources and feature engineering is fundamental in improving credit card fraud detection. The integration of behavioral and contextual transaction data enhances model accuracy, while data augmentation techniques like SMOTE and K-CGAN improve fraud detection in imbalanced datasets. Feature selection techniques such as PCA and Genetic Algorithms further optimize model performance, ensuring that fraud detection models operate efficiently and effectively.

As fraud tactics continue to evolve, the future of fraud detection will rely on integrating diverse data sources, improving data augmentation methods, and optimizing feature selection processes. By leveraging advanced data analytics and ML-driven feature engineering, financial institutions can develop highly accurate fraud detection systems that adapt to emerging fraud trends while minimizing operational risks.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

3.4. Practical Challenges and Future Research Directions

3.4.1. Challenges in implementing machine learning-based fraud detection system

The deployment of machine learning (ML)-based fraud detection systems in real-world financial environments presents several challenges, primarily concerning data privacy, regulatory compliance, and the evolving tactics of fraudsters. One of the most pressing concerns is data privacy, as financial institutions must handle highly sensitive customer information. The need for ML models to process vast amounts of transaction data raises concerns about privacy breaches and potential misuse of personal information. Federated learning has emerged as a promising approach to mitigate privacy risks by allowing ML models to train on decentralized data, preventing the need for direct data sharing (Silva *et al*., 2024). However, this technique introduces complexities related to model synchronization and communication efficiency, which may reduce the overall effectiveness of fraud detection systems (Nicholls *et al*., 2021).

Regulatory compliance is another major obstacle in the practical deployment of ML-based fraud detection. Different regions have varying data protection laws, such as the General Data Protection Regulation (GDPR) in Europe, which imposes strict guidelines on data collection, storage, and automated decision-making (Nicholls *et al*., 2021). These regulatory constraints require ML-driven fraud detection models to be explainable and transparent, yet many deep learning approaches, such as neural networks, function as "black boxes" with limited interpretability. Additionally, the dynamic nature of regulations forces financial institutions to continuously adapt their fraud detection strategies, which can be challenging when using static ML models that require frequent retraining.

Another significant challenge is the continuous evolution of fraud tactics, which demands adaptive and self-learning fraud detection models. Fraudsters leverage sophisticated techniques, including AI-generated fraudulent transactions and adversarial attacks, to bypass traditional detection mechanisms (Odeyemi *et al*., 2024). To counter this, ML models must continuously learn and adapt to emerging fraud patterns, necessitating frequent algorithm updates and retraining (Potla, 2023). However, the imbalanced nature of fraud datasets—where legitimate transactions vastly outnumber fraudulent ones—complicates the learning process, often leading to high false-negative rates, where fraudulent transactions are misclassified as legitimate.

Additionally, data quality and availability significantly impact the performance of fraud detection models. High-quality labeled datasets are essential for training ML models effectively, but financial institutions often struggle to obtain sufficient fraud-related data due to privacy concerns and proprietary restrictions (Silva *et al*., 2024; Farabi *et al*., 2024). Poor data quality, missing values, or noisy data can reduce model accuracy, necessitating robust data preprocessing and feature engineering techniques (Pan, 2024).

Despite these challenges, ML remains a powerful tool for fraud detection, offering the ability to detect fraud in real time, reduce false positives, and enhance fraud detection efficiency. However, financial institutions must implement innovative solutions such as federated learning, continuous model retraining, and advanced feature selection techniques to effectively address these challenges while ensuring data security and regulatory compliance.

3.4.2. The ethical and collaborative aspects of fraud detection

Beyond technical challenges, the ethical implications of ML-based fraud detection systems require careful consideration. While ML models enhance fraud detection accuracy, their reliance on automated decision-making raises concerns about bias, fairness, and

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

transparency. If not properly managed, algorithmic biases may lead to unfair classifications, disproportionately affecting certain customer groups (Nicholls *et al*., 2021). The lack of interpretability in deep learning models further complicates regulatory compliance, as customers and regulators demand clear explanations for fraud-related decisions. Future research must prioritize explainable AI (XAI) techniques that enhance model transparency and accountability.

Collaboration between financial institutions and AI researchers is essential for improving fraud detection models. The sharing of anonymized fraud-related data across financial institutions could enhance fraud detection accuracy, as fraudsters often operate across multiple financial networks. However, data-sharing limitations due to privacy regulations restrict such collaborative efforts. One potential solution is secure multi-party computation (MPC), which enables institutions to analyze fraud patterns collectively without exposing sensitive customer data (Silva *et al*., 2023).

Moreover, ethical considerations extend to the impact of fraud detection on customer experience. Overly aggressive fraud detection models may lead to a high rate of false positives, causing legitimate transactions to be flagged as fraudulent. This can result in customer frustration, transaction delays, and reputational damage for financial institutions. Research must focus on developing balanced fraud detection models that maintain high accuracy while minimizing disruptions to legitimate users (Silva *et al*., 2024).

Overall, the ethical deployment of AI in fraud detection must strike a balance between fraud prevention, customer rights, and regulatory compliance. Future AI-driven fraud detection systems should emphasize fairness, accountability, and transparency, ensuring that customers receive clear explanations for fraud-related decisions while maintaining high detection efficiency.

3.4.3. Future research directions in ML-based fraud detection

To address the evolving nature of fraud, future research must explore advanced machine learning architectures such as deep learning, reinforcement learning, and cross-domain data integration. Deep learning has demonstrated significant potential in fraud detection due to its ability to automatically extract hidden patterns from large datasets. Recent studies highlight the effectiveness of Generative Adversarial Networks (GANs) in enabling semi-supervised learning, allowing models to learn from both labeled and unlabeled data to detect emerging fraud patterns (Saranya *et al*., 2023). Additionally, deep learning models such as Long Short-Term Memory (LSTM) networks have proven effective in detecting sequential fraud patterns, further improving fraud detection accuracy (Jan, 2021; Sharma *et al*., 2022).

Reinforcement learning (RL) is another promising research avenue, particularly in real-time fraud prevention. RL-based models, such as Deep Q-Networks (DQN), can dynamically adjust fraud detection thresholds, optimizing fraud detection based on historical transaction outcomes (El-Toukhy *et al*., 2023). This adaptive learning capability allows ML models to respond to emerging fraud patterns in real-time, reducing false positives and improving overall detection performance (Njoku *et al*., 2024).

Furthermore, cross-domain data integration is gaining traction in fraud detection research. By incorporating behavioral and contextual data from multiple sources, ML models can develop a holistic understanding of fraudulent activity (Ejiofor, 2023). Federated learning techniques further facilitate secure cross-domain data sharing, allowing institutions to collaborate on fraud detection without compromising sensitive information (Silva *et al*., 2023).

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Another crucial research direction is hyperparameter tuning and model optimization in deep learning architectures. Studies indicate that fine-tuning hyperparameters significantly impacts fraud detection performance, particularly in terms of model accuracy and sensitivity (Sulaiman *et al*., 2024). Advanced techniques such as Autoencoders and transformers are being explored to enhance the accuracy of fraud detection models (Dehkordi *et al*., 2025).

Lastly, the deployment of ML-based fraud detection systems in different geographical and regulatory environments remains a key challenge. Fraud patterns vary across regions, demographics, and financial ecosystems, requiring ML models to be customized for different jurisdictions. Future research must focus on adaptive fraud detection models that can dynamically adjust to regional fraud patterns while complying with local regulations.

Future advancements in fraud detection will rely on deep learning, reinforcement learning, and cross-domain data integration to create more accurate, adaptable, and secure fraud detection systems. These research directions will not only improve fraud detection efficiency but also ensure that AI-driven fraud detection remains ethical, transparent, and regulatory-compliant.

4. CONCLUSION

Machine learning has emerged as a powerful tool for detecting credit card fraud, significantly outperforming traditional rule-based methods. The ability of ML models to analyze large transaction datasets, adapt to evolving fraud patterns, and minimize false positives has made them indispensable in financial fraud prevention. Among the most effective algorithms, Random Forest, LightGBM, and Artificial Neural Networks have demonstrated superior accuracy and recall in identifying fraudulent transactions. Additionally, techniques such as Synthetic Minority Over-sampling Technique (SMOTE) and K-Conditional Generative Adversarial Network (K-CGAN) have been instrumental in addressing data imbalance issues, ensuring that ML models generalize effectively to real-world scenarios.

Despite these advancements, challenges remain in implementing ML-based fraud detection systems, particularly regarding data privacy, regulatory compliance, and the continuous evolution of fraud tactics. Financial institutions must navigate complex data protection laws while maintaining the transparency and explainability of automated fraud detection decisions. Federated learning and secure multi-party computation (MPC) have been proposed as solutions to privacy concerns, but their widespread adoption still requires further research and technological refinement.

Future research should focus on deep learning architectures, reinforcement learning, and cross-domain data integration to improve fraud detection adaptability and efficiency. Additionally, the ethical considerations surrounding AI-driven fraud detection—particularly fairness, interpretability, and accountability—must be further explored. As fraud tactics continue to evolve, financial institutions must invest in self-learning, real-time fraud detection systems that dynamically adjust to emerging threats while ensuring compliance with global regulations.

5. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

6. REFERENCES

Aghware, F. O., Ojugo, A. A., Adigwe, W., Odiakaose, C. C., Ojei, E. O., Ashioba, N. C., and Geteloma, V. O. (2024). Enhancing the random forest model via synthetic minority oversampling technique for credit-card fraud detection. *Journal of Computing Theories and Applications*, *1*(4), 407-420.

Ahmed, F., and Shamsuddin, R. (2021). A comparative study of credit card fraud detection using the combination of machine learning techniques with data imbalance solution. In *2021 2nd International Conference on Computing and Data Science, 2021,* 112-118.

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., and Ahmed, M. (2022). Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. *IEEE Access*, *10*, 39700-39715.

Al-dahasi, E. M., Alsheikh, R. K., Khan, F. A., and Jeon, G. (2025). Optimizing fraud detection in financial transactions with machine learning and imbalance mitigation. *Expert Systems*, *42*(2), e13682.

Al-Dahidi, S., Madhiarasan, M., Al-Ghussain, L., Abubaker, A. M., Ahmad, A. D., Alrbai, M., and Zio, E. (2024). Forecasting solar photovoltaic power production: A comprehensive review and innovative data-driven modeling framework. *Energies*, *17*(16), 4145.

Al-Hashedi, K. G., and Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, *40*, 100402.

Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., and Egwuonwu, M. N. (2023). Analysing the impact of advanced analytics on fraud detection: a machine learning perspective. *European Journal of Computer Science and Information Technology*, *11*(6), 103-126.

Chaquet-Ulldemolins, J., Gimeno-Blanes, F. J., Moral-Rubio, S., Muñoz-Romero, S., and Rojo-Álvarez, J. L. (2022). On the black-box challenge for fraud detection using machine learning (ii): nonlinear analysis through interpretable autoencoders. *Applied Sciences*, *12*(8), 3856.

Cheah, P. C. Y., Yang, Y., and Lee, B. G. (2023). Enhancing financial fraud detection through addressing class imbalance using hybrid SMOTE-GAN techniques. *International Journal of Financial Studies*, *11*(3), 110.

Chen, W., Liu, C., Xing, F., Peng, G., and Yang, X. (2022). Establishment of a maturity model to assess the development of industrial AI in smart manufacturing. *Journal of Enterprise Information Management*, *35*(3), 701-728.

Choi, D., and Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. *Security and Communication Networks*, *2018*(1), 5483472.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Choi, D., and Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. *Security and Communication Networks*, *2018*(1), 5483472.

Compagnucci, I., Corradini, F., Fornari, F., Polini, A., Re, B., and Tiezzi, F. (2023). A systematic literature review on IoT-aware business process modeling views, requirements and notations. *Software and Systems Modeling*, *22*(3), 969-1004.

Da'u, A., and Salim, N. (2020). Recommendation system based on deep learning methods: a systematic review and new directions. *Artificial Intelligence Review*, *53*(4), 2709-2748.

Dehkordi, S. B., Nasri, S., and Dami, S. (2025). Unveiling anomalies: transformative insights from transformer-based autoencoder models. *International Journal of Computers and Applications*, *47*(1), 29-44.

Dube, L., and Verster, T. (2023). Enhancing classification performance in imbalanced datasets: A comparative analysis of machine learning models. *Data Science in Finance and Economics*, *3*(4), 354-379.

Ejiofor, O. E. (2023). A comprehensive framework for strengthening USA financial cybersecurity: integrating machine learning and AI in fraud detection systems. *European Journal of Computer Science and Information Technology*, *11*(6), 62-83.

El-Hassani, F. Z., Amri, M., Joudar, N. E., and Haddouch, K. (2024). A new optimization model for MLP hyperparameter tuning: modeling and resolution by real-coded genetic algorithm. *Neural Processing Letters*, *56*(2), 105.

El-Toukhy, A. T., Mahmoud, M. M., Bondok, A. H., Fouda, M. M., and Alsabaan, M. (2023). Countering evasion attacks for smart grid reinforcement learning-based detectors. *IEEE Access*, *11*, 97373-97390.

Farabi, S. F., Prabha, M., Alam, M., Hossan, M. Z., Arif, M., Islam, M. R., and Biswas, M. Z. A. (2024). Enhancing credit card fraud detection: A comprehensive study of machine learning algorithms and performance evaluation. *Journal of Business and Management Studies*, *6*(3), 252-259.

Ghosh Dastidar, K., Jurgovsky, J., Siblini, W., and Granitzer, M. (2022). NAG: neural feature aggregation framework for credit card fraud detection. *Knowledge and Information Systems*, *64*(3), 831-858.

He, Y. (2022). Machine learning methods for credit card fraud detection. *Highlights in Science, Engineering and Technology*, *23*, 106-110.

Hilal, W., Gadsden, S. A., and Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, *193*, 116429.

Hossain, M. N., Hassan, M. M., and Monir, R. J. (2022). Analyzing the classification accuracy of deep learning and machine learning for credit card fraud detection. *Asian Journal for Convergence in Technology (AJCT)*, *8*(3), 31-36.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Ilori, O., Nwosu, N. T., and Naiho, H. N. N. (2024). Advanced data analytics in internal audits: A conceptual framework for comprehensive risk assessment and fraud detection. *Finance & Accounting Research Journal*, *6*(6), 931-952.

Jan, C. L. (2021). Detection of financial statement fraud using deep learning for sustainable development of capital markets under information asymmetry. *Sustainability*, *13*(17), 9879.

Kenyon, W., and Tilton, P. D. (2012). Potential red flags and fraud detection techniques. *A Guide to Forensic Accounting Investigation*, 231-269.

Kraus, S., Breier, M., Lim, W. M., Dabić, M., Kumar, S., Kanbach, D., and Ferreira, J. J. (2022). Literature reviews as independent studies: guidelines for academic practice. *Review of Managerial Science*, *16*(8), 2577-2595.

Kumar, S., Ahmed, R., Bharany, S., Shuaib, M., Ahmad, T., Tag Eldin, E., and Shafiq, M. (2022). Exploitation of machine learning algorithms for detecting financial crimes based on customers' behavior. *Sustainability*, *14*(21), 13875.

Maniraj, S. P., Saini, A., Ahmed, S., and Sarkar, S. (2019). Credit card fraud detection using machine learning and data science. *International Journal of Engineering Research*, *8*(9), 110-115.

Mienye, I. D., and Sun, Y. (2023). A machine learning method with hybrid feature selection for improved credit card fraud detection. *Applied Sciences*, *13*(12), 7254.

Mohsen, O. R., Nassreddine, G., and Massoud, M. (2023). Credit card fraud detector based on machine learning techniques. *Journal of Computer Science and Technology Studies*, *5*(2), 16-30.

Mosa, D. T., Sorour, S. E., Abohany, A. A., and Maghraby, F. A. (2024). CCFD: Efficient credit card fraud detection using meta-heuristic techniques and machine learning algorithms. *Mathematics*, *12*(14), 2250.

Nayyer, N., Javaid, N., Akbar, M., Aldegheishem, A., Alrajeh, N., AND Jamil, M. (2023). A new framework for fraud detection in bitcoin transactions through ensemble stacking model in smart cities. *IEEE Access*, *11*, 90916-90938.

Nicholls, J., Kuppa, A., and Le-Khac, N. A. (2021). Financial cybercrime: A comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape. *IEEE Access*, *9*, 163965-163986.

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

Odeyemi, O., Mhlongo, N. Z., Nwankwo, E. E., and Soyombo, O. T. (2024). Reviewing the role of AI in fraud detection and prevention in financial services. *International Journal of Science and Research Archive*, *11*(1), 2101-2110.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Ogundokun, R. O., Misra, S., Maskeliunas, R., and Damasevicius, R. (2022). A review on federated learning and machine learning approaches: categorization, application areas, and blockchain technology. *Information*, *13*(5), 263.

Pan, E. (2024). Machine learning in financial transaction fraud detection and prevention. *Transactions on Economics, Business and Management Research*, *5*, 243-249.

Parvin, H., MirnabiBaboli, M., and Alinejad-Rokny, H. (2015). Proposing a classifier ensemble framework based on classifier selection and decision tree. *Engineering Applications of Artificial Intelligence*, *37*, 34-42.

Patel, S. K., and Panday, D. (2024). Optimizing credit card fraud detection: A genetic algorithm approach with multiple feature selection methods. *Advances in Distributed Computing and Artificial Intelligence Journal*, *13*, e31533-e31533.

Pitsane, M. Y., Mogale, H., and van Rensburg, J. J. (2022). Improving accuracy of credit card fraud detection using supervised machine learning models and dimension reduction. In *International Conference on Intelligent and Innovative Computing Applications*, *2022,* 290-301.

Potla, R. T. (2023). AI in fraud detection: Leveraging real-time machine learning for financial security. *Journal of Artificial Intelligence Research and Applications*, *3*(2), 534-549.

Qu, Z., Liu, H., Wang, Z., Xu, J., Zhang, P., and Zeng, H. (2021). A combined genetic optimization with AdaBoost ensemble model for anomaly detection in buildings electricity consumption. *Energy and Buildings*, *248*, 111193.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M. J. O. G. R. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, *71*, 804-818.

Saranya, N., Devi, M. K., and Mythili, A. (2023). Data science and machine learning methods for detecting credit card fraud. *The Scientific Temper*, *14*(03), 840-844.

Serenko, A. (2021). A structured literature review of scientometric research of the knowledge management discipline: a 2021 update. *Journal of knowledge management*, *25*(8), 1889-1925.

Setiawan, R., Tjahjono, B., Firmansyah, G., AND Akbar, H. (2023). Fraud detection in credit card transactions using HDBSCAN, UMAP and SMOTE methods. *International Journal of Science, Technology & Management*, *4*(5), 1333-1339.

Sharma, M. A., Raj, B. G., Ramamurthy, B., and Bhaskar, R. H. (2022). Credit card fraud detection using deep learning based on auto-encoder. *In ITM Web of Conferences*, *50*, 01001.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056

Silva, P. R., Vinagre, J., and Gama, J. (2023). Towards federated learning: An overview of methods and applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *13*(2), e1486.

Strelcenia, E., and Prakoonwit, S. (2023). Improving classification performance in credit card fraud detection by using new data augmentation. *AI*, *4*(1), 172-198.

Sulaiman, S. S., Nadher, I., and Hameed, S. M. (2024). Credit card fraud detection using improved deep learning models. *Computers, Materials & Continua*, *78*(1), 1050-1069.

Yang, F., Liu, S., Dobriban, E., and Woodruff, D. P. (2021). How to reduce dimension with PCA and random projections?. *IEEE Transactions on Information Theory*, *67*(12), 8154-8189.

DOI:
p-ISSN: 2828-9064 e-ISSN: 2828-9056