International Journal of Emerging Trends in Computer Science and Information Technology
ISSN: 3050-9246 | https://doi.org/10.56472/ICCSAIML25-110
Eureka Vision Publication | ICCSAIML'25-Conference Proceeding

Original Article

State-of-the-Art Machine Learning approaches for Fraud Detection in Financial Institutions

Yasodhara Varma Rangineeni
Vice President.

Abstract *- Financial fraud poses a significant threat to the viability & the integrity of financial institutions, leading to the significant monetary losses, reputational damage & their regulatory complications. Although conventional rule-based fraud detection systems are more effective, they encounter difficulty in adapting to the evolving & intricate nature of their fraudulent activity. ML has become a useful tool for their fraud detection since it helps to analyze large transactional data, identify complex patterns, and really discover anomalies by timing. This work explores hybrid, unsupervised, and supervised models as advanced ML techniques for fraud detection. Decision trees, random forests & the deep neural networks are the most frequently employed supervised learning methods due to their exceptional accuracy in identifying their fraudulent transactions. Unsupervised models, including autoencoders & clustering algorithms, are particularly adept at identifying emerging fraud trends without the necessity of previously labeled content. Furthermore, hybrid approaches take advantage of both paradigms to improve the detection's performance. The article highlights significant developments in the domains of graph-based fraud detection, FL for the privacy-preserving analysis & explainable artificial intelligence (XAI) integration to increase the model interpretability and their regulatory conformity. The accuracy, adaptability & the efficacy of ML-based fraud detection systems are significantly higher than those of traditional methods, as evidenced by experimental findings & case analyses. Nevertheless, challenges continue to exist, including the need for the continuous model retraining, aggressive assaults & data asymmetry. The results highlight the need of reaching financial fraud reduction by means of the scalable, interpretable, and strong ML models. By offering a thorough analysis of the sophisticated ML algorithms, their practical uses, and the possible research prospects in the field of fraud detection inside the financial institutions, this work adds to the present body of knowledge. Legislators, financial experts, and data scientists working to improve their fraud detection systems in a financial environment going more and more digital and complicated depend on the insights provided.*

Keywords *- Fraud detection, Machine Learning, Financial Institutions, Anomaly Detection, Supervised Learning, Unsupervised Learning, Deep Learning, Reinforcement Learning, Hybrid Models, Credit Card Fraud, Identity Theft, Data Preprocessing, Model Evaluation, Financial Security, Artificial Intelligence.*

1. Introduction

A pervasive & the dynamic threat to economic stability, a reduction in the consumer trust & the significant financial losses is fraud within financial institutions. Deceitful financial transactions & sophisticated their cybercrimes taking advantage of flaws in the digital banking systems constitute part of the fraudulent activity. Given the enormous volumes of sensitive information and the capital under control, financial institutions including banks, credit card firms & the investment organizations are main targets. Even with strict legal systems & security procedures, fraudsters constantly create new plans that call for advanced detection & the prevention mechanisms. Although fraud takes several forms, credit card fraud, identity theft & insider fraud is the most often occurring one. Credit card fraud results from illegal activity carried out with stolen or counterfeit card data. Often using phishing, card skimming & their online data breaches, this type of fraud is Identity theft is the act of someone adopting the identity of another to access their financial accounts, apply for loans, or commit other financial crimes. Through social engineering, hacking or data leaks criminals might get personally identifiable information (PII). Insider fraud is the result of employees of financial companies using their access to confidential information for their own advantage, therefore supporting outside their criminal activity or faulty transactions. Money laundering, wire fraud & the synthetic identity fraud each of which poses significant threats to their financial institutions and their customers are other forms of fraud.

Conventional methods of the fraud detection mostly rely on the established thresholds, hand reviews & their rule-based algorithms to spot questionable transactions. These methods have inherent restrictions even if they provide a minimal degree of their security. Static rule-based systems provide high faulty positives & false negatives since they cannot change with changing fraud methods. While complex fraud schemes may go unnoticed, actual transactions could be mistakenly labeled as fraudulent & cause customer dissatisfaction. Manual evaluations are time-consuming & labor-intensive, so they are not practical for actual time fraud detection in huge volumes of financial transactions. Moreover, fraudsters constantly change their approach to escape their traditional detection techniques, therefore stationary models are useless in changing fraud environments.

Providing adaptive, data-driven approaches for anomaly identification & predictive analysis, machine learning (ML) has evolved into a useful tool in the battle against financial crime. Unlike traditional rule-based methods, ML models can evaluate vast amounts of transactional information, detect complex patterns & precisely mark fraudulent activity. Using labeled previous fraud data, supervised learning techniques including decision trees, support vector machines & the neural networks help to train models able to categorize the incoming transactions as fraudulent or legitimate. Techniques for unsupervised learning such as anomaly detection algorithms & clustering are quite good at actual time detection of fresh fraud trends. Moreover, by absorbing the knowledge from new fraudulent events, reinforcement learning helps to constantly improve their fraud detection techniques. Including ML into financial fraud detection systems greatly increases detection efficacy, lowers running prices, and helps to reduce faulty positives. Examining innovative ML techniques for fraud detection in financial institutions, this article evaluates their efficacy, challenges & pragmatic uses. Emphasizing their advantages over traditional approaches, this will look at several ML methodologies including hybrid approaches, ensemble models & deep learning. The article will also examine useful applications, legal considerations & developing patterns in ML-based fraud avoidance. This work intends to provide a whole assessment of advanced ML techniques so as to improve the fraud detection capabilities of financial organizations.

2. Machine Learning Techniques for Fraud Detection

Financial fraud detection has become a major focus of research & the development since ML techniques play a basic role in the recognition of the fraudulent transactions. Conventional statistical models & complex deep learning systems are among the several ML approaches applied to increase their fraud detection efficiency. The main ML techniques used in fraud detection are investigated in this part under supervised learning, unsupervised learning, reinforcement learning & the hybrid approaches.

2.1 Methodologies of supervised learning

Often used in fraud detection, supervised learning uses labeled datasets to classify previous events as either lawful or fraudulent. These models use historical trends to project results using fresh content.

2.1.1 Logistic Model

A basic statistical model used in their binary classification activities including fraud detection is logistic regression. Using previous data, it evaluates the probability that a certain transaction is fraudulent. This approach is quite understandable, which makes it a preferred choice for the financial companies following regulations. Although simple, logistic regression can be powerful when combined with feature engineering techniques including transaction-specific traits, including transaction frequency, geographic location & the abnormal spending behavior. Still, its effectiveness may be limited when dealing with significantly skewed datasets, a common problem in their fraud detection.

2.1.2 Random Forests and Decision Trees

Hierarchical models called "conclusion trees" divide the data based on the features to produce a set of rules leading to a categorization conclusion. Individual decision trees often overfit the data, therefore reducing their generalizability even if they are easy to understand. Random forests, an ensemble learning method, combine many decision trees to maximize their resilience & accuracy in order to meet this restriction. Random forests show lowered sensitivity to noise & effectively manage their complex interactions among numerous variables in the fraud detection. Still, their computational complexity increases with tree count, which makes them less effective for actual time fraud detection systems.

2.1.3 Support Vector Mechanisms (SVM)

A strong classification method called Support Vector Machine (SVM) finds a perfect hyperplane to tell apart non-fraudulent from the fraudulent transactions. It is particularly effective in high-dimensional environments, so it helps to manage these several facts of transactions. By detecting nonlinear connections among the transaction characteristics, kernel-based SVM models improve fraud detection efficacy. Nevertheless, SVMs can be resource-intensive, especially when managing huge financial information, therefore limiting their value in actual time fraud detection systems.

2.1.4 Synthetic Neural Networks

By spotting complex, nonlinear trends in the transactional data, neural networks especially deep learning models—have shown amazing accuracy in their fraud detection. Sequential transaction data can be used by multi-layer perceptron (MLPs) & recurrent neural networks (RNNs) to find anomalies suggestive of fraud. Image-like models of transactional activity have been examined using their convolutional neural networks (CNNs). Although neural networks show great accuracy, their interpretability presents difficulties that call for questions about their regulatory compliance & the explained ability for the financial firms. Furthermore constructing deep learning models requires huge computer resources and lots of labeled data.

2.2 Methods of Unsupervised Learning

In fraud detection when labeled information is either restricted or absent, unsupervised learning techniques prove helpful. These systems find their abnormalities or hidden trends in transactional data without known fraud flags.

2.2.1 An anomaly detection system Using Clustering Methods (DBSCAN, K-means)

Often used for fraud detection by grouping their related transactions are clustering techniques including K-means & Density-Based Spatial Clustering of the Applications with Noise. K-means organizes data based on their feature similarity, thereby pointing out the transactions that deviate from any cluster as possible fraud. K-means assumes that clusters are spherical & of similar scale, which does not always match actual fraud tendencies. Conversely, DBSCAN searches for dense areas in the data and labels points in sparse areas as anomalies. Since fraudulent transactions usually show as outliers, DBSCAN is more successful in the fraud detection. DBSCAN is more flexible for the financial fraud detection than K-means since it does not need a prior cluster number of specifications. Still, its efficiency depends on the parameter choice, which can affect its ability to spot their dishonest behavior.

2.2.2 Autoencoders

Because they can develop effective representations of the transactional information, a class of neural networks called autoencoders is often applied in the fraud detection. These models consist of a decoder to reconstruct input information & an encoder compressing it. Most usually an anomaly is a transaction that significantly differs from the reconstructed version. Particularly good at handling high-dimensional financial information, autoencoders can also identify subtle fraud trends. Still, they require careful hyperparameter tuning & rely on an unsupervised approach, therefore faulty positives could arise should usual transaction variations be mistakenly detected as fraudulent.

2.2.3 Separative Forest

By means of recursive dataset partitioning, isolation forest is a tree-based method for the anomaly detection that isolates their fraudulent transactions. Because of their frequency, fraudulent transactions are found more quickly than the normal ones. With high-dimensional information, this method is computationally efficient & the effective, hence suitable for actual time fraud detection. Still, as with other anomaly detection methods, it could run across challenges with complex fraud schemes that pass for the natural behavior, therefore producing possible faulty negatives.

2.3 Reinforcement Learning for Detection of Fraud

Comparatively with traditional supervised & unsupervised approaches, reinforcement learning (RL) provides a fresh approach in their fraud detection. In reinforcement learning, an agent learns to respond in an environment by getting the rewards or penalties, therefore improving its decision-making over time. Reinforcement learning agents are taught to recognize their fraudulent transactions in financial fraud detection by means of the constant method modification in response to evolving fraud trends. Unlike supervised learning, which relies on the previous labels, reinforcement learning models adaptably absorb information from fresh information, hence increasing their resilience to changing their fraudulent techniques. Optimizing actual time transaction monitoring systems is one way reinforcement learning helps to prevent fraud. Reward-based feedback allows reinforcement learning models to maximize their fraud detection accuracy while restricting faulty positives, hence preserving consumer delight by means of the optimal accuracy. Furthermore, reinforcement learning can be combined with other ML techniques as policy gradient approaches or deep Q-networks (DQNs) to raise their fraud detection efficacy in the financial institutions. Notwithstanding its promise, reinforcement learning requires careful design of reward structures and huge computational resources to ensure best performance in fraud detection.

2.4 Combining Strategies

Integration of several ML techniques reduces their faulty positives & improves their detection accuracy, hence strengthening hybrid fraud detection models. Combining supervised & unsupervised learning techniques helps hybrid models to use their labeled information & spot developing fraud trends. Commonly used in a hybrid approach to find suspicious transactions which are subsequently identified by a supervised model, such as a random forest or a neural network are anomaly detection methods including their autoencoders & clustering. This helps the system to quickly identify both known & unidentifiable fraud events. Hybrid fraud detection systems benefit much from ensemble learning approaches including stacking & boosting. Weak classifiers are combined in the models as XGBoost & LightGBM to create a more robust fraud detection system. By combining the advantages of many models & reducing their individual flaws, these ensemble techniques improve forecast accuracy. By combining many ML techniques, hybrid models offer a more complete approach for the fraud detection, hence enhancing the security & integrity of financial transactions. Still, they require careful model choice & tuning to ensure best performance while avoiding too complex processing.

3. Materials and Methods

3.1 Dataset and Data Preprocessing

This work applied anonymized actual world transaction information from financial institutions when feasible & financial fraud detection datasets acquired from publicly available sources like Kaggle. Public datasets including Kaggle's "Credit Card Fraud Detection" dataset provide anonymized transactional records tagged as either fraudulent or non-fraudulent, therefore enabling researchers in their effective testing & validation of fraud detection systems. Moreover, financial institutions routinely compile thorough transactional records covering consumer demographics, transaction history, device metadata, geolocation data & so provide important information for the fraud detection. Ensuring better inputs for ML models depends on the first data preparation. Before training, the raw datasets usually show missing values, repeated transactions & inconsistent entries that call for correction. For numerical variables, imputation techniques including mean, median, or mode substitution; for categorical information, missing values are usually handled in the most common category. To prevent bias in the model training, duplicate transactions are minimized; data normalization is used to normalize their feature distributions therefore guaranteeing interoperability among several ML models.

In fraud detection, feature engineering is essential since plain transaction data might not sufficiently capture complex fraudulent trends. Different domain-specific traits are developed: user transaction frequency, departure of transaction amounts from the average & location-based anomaly identification. Furthermore generated to increase the accuracy of fraud detection are time-series characteristics including session-specific activity & expenditure patterns over several temporal periods. Depending on the parameters of the model, one-hot encoding or target encoding codes categorical variables including transaction category & merchant type. Recursive Feature Elimination (RFE) & mutual information-based selection help to maintain their most relevant features by means of the unnecessary or non-informative variables being eliminated.

3.2 Model Assessment Metrics and Training

Using labeled datasets to separate fraudulent from actual world transactions, the model training process uses supervised ML techniques Among the several advanced ML models used are logistic regression, decision trees, random forests, gradient boosting machines (e.g., XGBoost, LightGBM) & DL frameworks including artificial neural networks (ANNs) and their long short-term memory (LSTM). Hyperparameter tuning techniques grid search, random search & Bayesian optimization help to improve these models to produce their best prediction performance. Usually split into training, validation & test sets to evaluate their model generalization, the training data Usually using a normal 70-20-10 split, 70% of the data goes for training, 20% for validation & 10% for testing. Moreover, resampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) & Adaptive Synthetic Sampling (ADASYN) are used to handle class imbalance issues due of the clear disparity in the fraud detection datasets characterized by a significant predominance of legitimate transactions over fraudulent ones. Alternatively, cost-sensitive learning approaches are included into model training to impose larger penalties for the misclassification of fraudulent events compared to non-fraudulent situations.

Model performance is investigated using several evaluation criteria. Basic assessments used to assess the effectiveness of the fraud detection systems are precision, recall & also the F1-score. Precision True Positives / (True Positives + False Positives) measures the proportion of the precisely identified fraudulent transactions to the total expected fraudulent transactions. Recall, sometimes known as True Positives / (True Positives + False Negatives), gauges the model's ability to spot the fraudulent cases among the overall count of the actual fraud events. Precision & recall must be balanced in the fraud detection; consequently, the F1-score the harmonic mean of these two measures is a consistent indicator. A complete performance measure, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) evaluates the balance between the true positive rate & also faulty positive rate under several classification thresholds. Whereas values near 0.5 indicate random guessing, AUC-ROC values around 1 indicate extraordinary model performance.

Reliability of the model evaluation is raised via cross-valuation methods. Often with k=5 or k=10, K-fold cross-validation ensures that the model is trained & assessed on the several subsets of the dataset, therefore reducing bias from the data partitioning. Because stratified K-fold cross-valuation preserves the original class distribution in every fold, so offering a balanced representation of the fraudulent & actual world transactions, it is preferred for their fraud detection activities. Furthermore used for the financial records showing temporal correlations is time-series cross-valuation, which ensures that the previous events are used to project future fraudulent activity & thereby prevents data leakage. This work aims to provide an efficient ML framework by means of extensive data preparation, feature engineering, model training & strict evaluation techniques for the detection of fraudulent conduct in the financial transactions.

4. Results and Discussion

4.1 Comparative Analysis of ML Models

By increasing accuracy, efficiency & the adaptability, ML has greatly improved fraud detection in the financial firms. We have extensively evaluated a variety of ML techniques including supervised, unsupervised & hybrid models for their ability to detect their fraudulent behavior. Usually, accuracy, precision, recall, F1-score & the area under the receiver operating characteristic (ROC-AUC) curve define the general efficacy of these models. Because they can learn from labeled prior fraud content, supervised learning algorithms logistic regression, decision trees, random forests, gradient boosting machines (GBMs) are widely applied. Among these, Gradient Boosting Machines (GBMs), such XGBoost & LightGBM, show improved prediction performance since their power to these model complex feature interactions & manage their imbalanced datasets via weighting & boosting techniques. Although they often require huge processing resources, random forests offer durability & clarity. Although simple & understandable, logistic regression is not suitable for these modeling nonlinear interactions, so it is less suitable for these complex fraud detection applications.

Particularly artificial neural networks (ANNs) & the convolutional neural networks (CNNs), DL models have attracted much interest recently. These models can independently extract high-level features from the unprocessed transaction data, hence reducing the need for hand feature engineering. A kind of the recurrent neural networks (RNNs), long short-term memory (LSTM) networks are rather good in spotting bogus trends in the sequential financial transactions. Still, the main drawbacks of DL models are their restricted interpretability & significant computing price, which complicate regulatory conformity & the dependability. Using unsupervised learning techniques clustering algorithms (e.g., k-means, DBSCAN) and anomaly detection models (e.g., Isolation Forest, One-Class SVM, Autoencoders) we have identified fraud in settings with little labeled information. These models are rather good in their spotting unusual fraud trends by spotting the departures from usual transaction behavior. Still, they often have high faulty positive rates since unusual but legal transactions may be mistakenly labeled as fraud. Furthermore, it is still tough to optimize these models to lower faulty alarms.

Fraud detection has benefited much from hybrid approaches integrating supervised & unsupervised methods. Combining an anomaly detection system with a deep learning classifier in an ensemble model helps to use both approaches. By means of many points of view, these models can enhance their accuracy & lower faulty positives by use of their forecasts. Still, hybrid methods can demand very huge computer resources & careful hyperparameter adjustment to strike a balance between recall & the accuracy. Although ML-based fraud detection has advanced, no one model always outperforms others in all conditions. Several factors determine the best suitable model: data availability, computing constraints, legal requirements & the financial institution tolerance of faulty positives. As a result, financial institutions are using adaptive learning methods & group approaches increasingly to raise their effectiveness of fraud detection.

4.2 Empirical Case Research

Many financial institutions have successfully put ML-based fraud detection systems into use to lower the faulty transactions. With deep learning & anomaly detection techniques applied to monitor millions of the transactions in actual time, PayPal is a shining example. Using previous transaction information, user behavior research & the feature engineering, PayPal's fraud detection technology precisely detects maybe fraudulent transactions. By continually improving its algorithms with fresh transaction information, the company has drastically reduced their fraudulent losses while maintaining their perfect customer experience. ML techniques are used by JPMorgan Chase to spot their money laundering & the credit card fraud activity. Combining network analysis techniques with supervised learning models including GBMs & their neural networks the institution finds unusual transaction trends across their several accounts. Using risk assessment tools driven by AI, JPMorgan Chase has improved their fraud detection rates & lowered their faulty positives. These useful applications underline the effectiveness of ML in combating fraud in their financial institutions, therefore stressing the need of implementing their analytics-driven approaches.

4.3 Roadblocks and Restraints

Though ML-based fraud detection offers several benefits, it still has several fundamental constraints & the challenges. Data imbalance is a major challenge since only a small fraction of the whole dataset are their fraudulent transactions. This discrepancy often produces biased models that give the majority (non-fraudulent) class top priority, therefore compromising the recall rates for their fraud detection. Often used techniques including oversampling, under sampling & the synthetic data generation (e.g., SMote) help to solve this issue; still, careful application is necessary to avoid their overfitting. The interpretability of ML models especially those derived from DL approaches is a clear challenge. Financial companies have legal obligations to follow rules requiring openness in their procedures of decision-making. Deep neural networks among other black-box models provide difficulties in clearly explaining why a certain transaction is labeled as fraudulent. Techniques including LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations) have been developed to clarify model predictions and help to overcome this. Still, reaching balance between model interpretability and efficiency is an ongoing challenge.

Machine learning-based fraud detection suffers much from ethical and privacy concerns. Using vast transaction data raises questions about consumer rights, data privacy, and possibly predictive model biases. Biased models might negatively impact some demographic groups, leading to unfair treatment of customers. Financial institutions have to build thorough data governance systems and follow laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) to help to reduce these hazards. Although ML-based fraud detection has transformed financial security, its efficacy depends on addressing data imbalance, interpretability, and ethical issues, so improving financial security even if it has revolutionized it. Incorporating ethical AI ideas and using explainable artificial intelligence technologies would help financial institutions always improve their machine learning approaches to support compliance and confidence.

5. Conclusion and Future Directions

By allowing more exact, efficient & the flexible identification of fraudulent operations, ML has been revolutionized from fraud detection in their financial firms. By means of a comparative analysis of ML models, including supervised learning models such as gradient boosting machines (GBMs), DL techniques like long short-term memory networks (LSTMs) & unsupervised anomaly detection methods, clarifies the advantages & constraints of various approaches. With plenty of labeled data, supervised learning models attain great accuracy; deep learning techniques excel in feature extraction but have problems in interpretability. Although they are good in spotting emerging fraud patterns, unsupervised techniques usually run high false positive rates. Although hybrid models combining several approaches have shown improved fraud detection efficacy, they still need careful fine-tuning & huge computational resources. Case studies of financial firms as PayPal & JPMorgan Chase show how well ML-based fraud detection lowers their financial crime rates. Still, important problems such as data imbalance, model interpretability & ethical considerations have to be addressed if we are to improve their effectiveness of fraud detection.

Future ML-based fraud detection studies should give top priority to improving their model interpretability, lowering faulty positive rates & enhancing flexibility to newly developing fraudulent techniques. Following legal guidelines & boosting confidence in ML-based fraud detection systems depend on the development of explainable artificial intelligence (XAI) approaches. Moreover, development in the FL could enable institutional cooperation on their fraud prevention, thereby preserving data privacy. Graph-based models and their reinforcement learning techniques taken together might increase the ability to spot their complex fraud activities in actual time. Beyond fraud detection, AI will play a role in financial security in their proactive risk assessment, anomaly prediction & the automated financial crime investigation as it develops. By using AI-driven tactics that combine ethical & the transparent AI practices with efficient fraud detection, financial institutions have to remain proactive to guarantee their security & the compliance in a financial environment going increasingly digitized.

References:

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

[2] Shahana, T., Lavanya, V., & Bhat, A. R. (2023). State of the art in financial statement fraud detection: A systematic review. *Technological Forecasting and Social Change*, *192*, 122527.

[3] 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.

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

[5] Kamuangu, P. (2024). A Review on Financial Fraud Detection using AI and Machine Learning. *Journal of Economics, Finance and Accounting Studies*, *6*(1), 67-77.

[6] Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2022). Fraud detection in banking data by machine learning techniques. *Ieee Access*, *11*, 3034-3043.

[7] 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.

[8] Mienye, I. D., & Jere, N. (2024). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. *IEEE Access*.

[9] Jovanovic, D., Antonijevic, M., Stankovic, M., Zivkovic, M., Tanaskovic, M., & Bacanin, N. (2022). Tuning machine learning models using a group search firefly algorithm for credit card fraud detection. *Mathematics*, *10*(13), 2272.

[10] Alghofaili, Y., Albattah, A., & Rassam, M. A. (2020). A financial fraud detection model based on LSTM deep learning technique. *Journal of Applied Security Research*, *15*(4), 498-516.

[11] 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.

[12] Goecks, L. S., Korzenowski, A. L., Gonçalves Terra Neto, P., de Souza, D. L., & Mareth, T. (2022). Anti-money laundering and financial fraud detection: A systematic literature review. *Intelligent Systems in Accounting, Finance and Management*, *29*(2), 71-85.

[13] Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud.

[14] *Decision Support Systems*, *139*, 113421.

[15] Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & 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.

[16] S. S. Nair, G. Lakshmikanthan, J.ParthaSarathy, D. P. S, K. Shanmugakani and B.Jegajothi, ""Enhancing Cloud Security with Machine Learning: Tackling Data Breaches and Insider Threats,"" 2025 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2025, pp. 912-917, doi: 10.1109/ICEARS64219.2025.10940401.

[17] Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in post-pandemic era. *The Innovation*, *2*(4).