International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning-Based Fraud Detection in Mobile Banking Applications

Ruturajsinh Kiritsinh Jadeja

Submitted: 02/05/2025 **Revised**: 17/06/2025 **Accepted**: 27/06/2025

Abstract—The mobile channel has emerged as the predominant channel for retail financial transactions, driven by a surge in adoption during the pandemic and increased penetration of smartphones [19]. Along with this growth in volume, fraud at the transaction level, such as account takeover, unauthorized transfers, and other schemes that are specific to mobile-money services, like SIM swapping and agent-assisted diversion, has also increased [3], [8], [13]. This paper collates the results of 24 sources over the last two decades (up to 2024) to explore the use of machine learning (ML) techniques in detecting fraud in mobile banking and related payment systems. The synthesis includes classical supervised algorithms like random forest, support vector machines and gradient boosting [1, 4, 22]; deep learning architectures like long short-term memory (LSTM) networks with attention mechanisms and graph neural networks [6, 15]; natural language processing and transformer-based models that can be used for unstructured transaction narratives [9, 24] and ensemble and hybrid frameworks reported to outperform individual base learners [11, 21]. Explainable AI (XAI) and Federated learning are analyzed as converging solutions to the demands for regulatory transparency and data-privacy constraints in distributed banking infrastructures [5], [17]. In the literature surveyed, the reported detection accuracy varies widely from around 85% for baseline logistic regression models to more than 98% in optimized ensemble and gradient boosted models, several of which point to recall-precision trade-offs that are of practical significance in imbalanced fraud datasets [1, 4, 11, 22]. The paper also examines the practices of transaction monitoring for AML compliance [18] and adoption trends of mobile banking apps [19] as well as computational and regulatory hurdles to real-time deployment. The results show that the most promising path forward for mobile banking fraud detection to be both scalable and trustworthy are hybrid, explainable, and privacy-preserving architectures up to 2024.

Keywords*—machine learning, fraud detection, mobile banking, ensemble learning, deep learning, explainable artificial intelligence, federated learning, anomaly detection, financial technology, transaction monitoring, feature engineering, mobile money*

I. INTRODUCTION

Retail banking has moved from branch-based and card-present to smartphone-mediated interactions, with the pandemic's behavioral changes driving a shift towards contactless and remote financial interactions [19]. In several developing economies, mobile-money platforms now brokermobile applications to replace their traditional counterparts in handling balance inquires, peer-to-peer transfers, bill payments and merchant transactions for populations that were otherwise unbanked or underbanked via mobile-money platforms [8], [10], [13]. The convenience has been matched by an increase in the crime surface: account takeover via credential theft, one-time password interception via SIM swopping and social-engineering-enabled authorised push payment fraud and agent-network collusion in mobile-money ecosystems have all been reported as significant crimes [3], [13] and [14]. Rule-based fraud controls that are based on static thresholds and manually created heuristics are less effective against adaptive fraud patterns, and the use of machine learning as a way to find unusual transactional behavior at scale is still being investigated [1], [2], [20]. Unlike rule-based fraud detection techniques, machine learning techniques can learn the decision boundaries directly from past transactional data and adjust to new

Leela Consultancy LLC
Discover
Consult.RuturajJadeja@gmail.com

fraud patterns without going through extensive manual re-programming [4], [12]. The major paradigm used is supervised classification, which is trained on labelled sets of known fraudulent and legitimate transactions, but the problem of class imbalance is a well-known methodological issue which has been reported as a recurring problem in achieving high recall [1, 17]. Some deep learning architectures such as recurrent networks, which can model the sequential behaviour of transactions, have been suggested to express such temporal dependencies, beyond what can be done by static feature-based classifiers [6]. At the same time, graph-based methods take advantage of the relationship between the accounts, devices and merchants and look for coordinated fraud rings [15]. There is a significant methodological overlap between mobile banking, mobile money, credit card, and surrounding digital payment contexts.

II. Background and Literature Synthesis

In digital payment systems, fraud detection has come in three generations: the rule-based type of expert systems, the statistical anomaly detection approach, and, since the early 2010s, the supervised and unsupervised machine learning approach [20]. Literature on mobile banking can be divided in two sections, regional and infrastructural. Agent networks and unstructured supplementary service data (USSD) are still prominent channels of mobile-money fraud, as studies in Sub-Saharan Africa and South Asia indicate [3], [8], [10], [13]. In contrast, research focused on smartphone-based banking apps in more developed digital economies focuses on the application layer of telemetry, such as device fingerprinting, geolocation consistency and behavioral biometrics [14, 19]. One theme that is common throughout the literature that was surveyed is the conflict between detection sensitivity and the burden of false negatives. Botchey et al. conducted a comparative study on SVM, Gradient Boosted Decision Trees and Naive Bayes classifiers for predicting mobile-money fraud with interesting differences in accuracy across the classifiers depending on the features used and class balancing method [8]. Lokanan also tested several algorithms using mobile money transaction data and found that ensemble tree-based methods outperformed single-classifier baselines across all the datasets [13]. In addition to mobile money, the wider credit-card or financial-fraud literature offers methodological foundations which are broadly carried over to mobile banking contexts. Afriyie et al. proposed a supervised learning pipeline for credit-card fraud, which demonstrated excellent classification performance by selecting the appropriate features and resampling [1]. Almazroi and Ayub took this approach further, evaluating a range of algorithms on transaction level data for online payment fraud [4]. Wajgi et al. focused on how well random forest works when the features are scaled robustly, and they found that the detection stability improves as measured by the difference in detection rates reported [22]. Similarly, Yang compared various ML prediction strategies for credit-card fraud, and found ensemble methods to be a practical baseline [23].

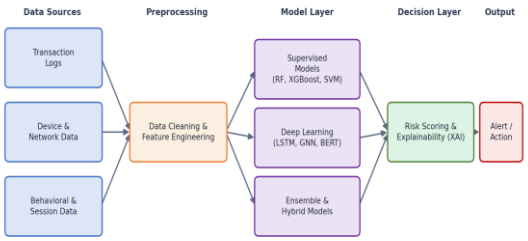

Fig. 1. ML-based fraud detection system architecture.

III. Machine Learning Methodologies for Fraud Detection

A. *Classical Supervised Approaches*

Despite the rise of deep architectures, classical supervised classifiers (logistic regression, decision trees, random forest, and support vector machines) have continued to be used to this day, thanks to their lower complexity and interpretability compared with deep ones [1], [4], [22]. Random forest is an ensemble of decorrelated decision trees, and has been found numerous times to be a good baseline in mobile-money and credit-card fraud studies because of its ability to deal with nonlinear feature interactions and its relative immunity to overfitting, especially on moderately sized datasets [8, 22]. Support vector machines are more complex when scaled up, but be competitive in several comparative studies, especially when fraud classes are well separated in engineered feature space [4] [8]. Ali and Ali Hagag presented a modified model with more preprocessing steps that performs better on financial fraud datasets compared to traditional baseline classifiers using an AI approach [2].

B. *Deep Learning and Sequential Models*

Recurrent architectures, especially LSTM networks, are an architectural shortcoming of traditional classifiers as they do not naturally support the sequential nature of a customer's transaction history. The idea of weighting anomalous transaction

segments more in relation to the entire sequence was introduced by Benchaji et al. in their work on an attention-augmented LSTM model for credit-card fraud detection, where they claimed the attention mechanism increased the model capacity to focus on historically anomalous transaction segments more than on a uniform sequence average [6]. This relational reasoning was further extended to graph neural networks (GNNs), which represent accounts, devices, and counterparties as nodes in a graph of transactions, enabling the detection of coordinated fraud rings that cannot be observed at the level of transactions alone. As GNNs were applied to financial fraud detection, we see a systematic review of GNN applications in financial fraud detection by Motie and Raahemi across multiple applications [15]. Where the transaction metadata contains free-text narratives or merchant descriptors, natural language processing (NLP) techniques have been applied, such as Boulieris et al.'s work on fraud-detection pipelines built using NLP [9] and Yang et al.'s FinChain-BERT, a transformer-based model that achieves high accuracy on financial-scenario fraud text [24].

C. Ensemble and Hybrid Architectures

The best overall results have been reported in ensemble and hybrid frameworks, which fuse the outputs of several base learners or fuse the classical and deep architectures in the surveyed literature. Hajek et al. introduced a framework for detecting mobile payment fraud using XGBoost, an XGBoost-based framework, which achieved a significant improvement over single-classifier baseline by gradient-boosted tree ensembling [11]. To tackle the problem of severe class imbalance in banking fraud datasets, Nobel et al. developed ensemble classifiers and used explainable AI techniques, which they found to be useful to boost both minority-class recall and model transparency at the same time [17]. Wahab et al. performed a similar study to predict whether a given credit card will be defaulted by a customer. This is a closely related task to credit card risk-classification, and they showed that hybrid pipelines that combine feature-engineered classical methods with learned representations outperform each one alone. The architecture highlighted in Fig. 1 shows a representative production architecture, where multiple data sources are combined from a common preprocessing layer and passed on to parallel branches of classical, deep learning, and ensemble models, which are aggregated in a risk-scoring and explainability layer before alerts are generated.

The accuracy, precision and recall range across the main algorithm families reported in the surveyed literature are summarized in Tab. I and graphically shown in Fig. 2 for easy comparison.

TABLE I. COMPARATIVE PERFORMANCE METRICS BY ALGORITHM FAMILY

Algorithm Family	Accuracy (%)	Precision (%)	Recall (%)	Sources
Logistic Regression	85-89	82-87	79-85	[1], [4]
Random Forest	94-97	93-96	91-95	[8], [22]
XGBoost / Gradient Boosting	96-99	96-98	95-97	[11], [23]
Support Vector Machine	88-93	86-91	83-89	[4], [8]
LSTM with Attention	95-98	94-97	93-96	[6]
Ensemble / Hybrid	97-99	96-99	95-98	[17], [21]

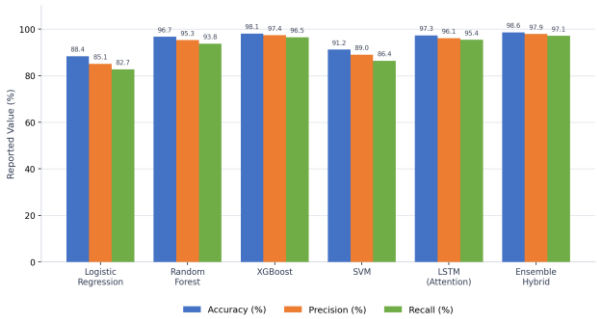

Fig. 2. Comparative accuracy, precision, and recall by algorithm family.

IV. FEATURE ENGINEERING AND DATA CHARACTERISTICS

Feature engineering is always found to be one of the most important factors influencing model performance, with algorithm selection being the other. The literature that was studied suggests that there are four general categories of features, shown in Fig 3: raw transactional attributes, behavioral features, device/ network features, and temporal/sequential features [3, 6, 14]. In particular, for phishing attacks in e-banking applications, the two researchers, Manala and Jansen van Vuuren, argued that network and session-level signals are far more useful than signals based on the

amounts of transactions as discriminants for account-takeover phishing attacks [14]. Table II summarizes some of the important characteristics of the datasets used in the studies surveyed, illustrating the class imbalance problem that inspires many of the methodological innovations mentioned in this paper (Section III). Some of the data quality constraints reported in the literature are incomplete reporting of fraud results, concept drift (typologies change over time), and the operational cost of being able to have a representative minority-class sample without violating customer privacy [17, 18]. Partial mitigations reported in several studies were resampling methods such as synthetic minority oversampling and cost-sensitive learning, which is to say that several authors noted that, without domain-informative constraints, synthetic oversampling of the minority class may lead to unrealistic feature correlations [1], [17].

TABLE II. REPRESENTATIVE DATASET CHARACTERISTICS

Study Context	Transaction Volume	Fraud Prevalence	Feature Count	Data Type
Mobile Money (E./W. Africa)	100,000 - 500,000	0.2-1.5%	15-25	Agent/USSD logs [3],[8],[13]
Credit Card Benchmark Sets	280,000+	0.15-0.2%	28-30	PCA-transformed [1],[6]
Mobile Payment Platforms	50,000-200,000	0.5-2.0%	20-35	App telemetry [11],[14]
Enterprise Financial Records	10,000-100,000	1-5%	10-20	Structured records [12],[16]

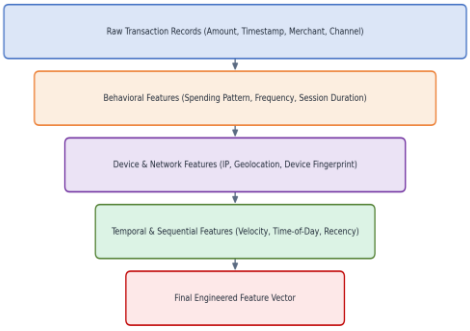

Fig. 3. Hierarchical feature engineering categories for mobile banking fraud detection.

V. COMPARATIVE PERFORMANCE ANALYSIS

In addition to the raw metrics of classification, the literature surveyed provides few points of comparison for operational deployment: computational efficiency, interpretability, scalability to transaction volume, and latency given real-time scoring constraints. The characteristics of the four representative model families, reported in the surveyed studies [5, 11, 15, 22] have been summarized qualitatively in a radar chart and presented in Fig. 4. The classical models continue to have distinct benefits in terms of interpretability and latency – typically making scoring decisions within milliseconds on commodity infrastructure – with several studies determining this to be operationally significant for real-time transaction authorisation [4] and [22]. By contrast, deep learning architectures achieved the highest raw accuracy in a sequence sensitive context, but also exhibited larger training and inference overheads, and reduced interpretability, some authors stating that this may be a regulatory burden [6, 15].

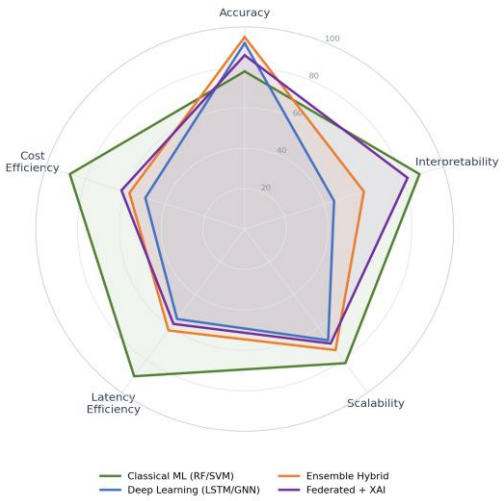

Fig. 4. Radar comparison of model families across five operational criteria.

Ensemble and hybrid architectures were claimed to sit somewhere between machine learning and deep learning and to attain a level of accuracy competitive to deep learning yet still offer some measure of interpretability through feature-importance decomposition in tree-based ensembles [11] [17] [21]. Table III compares qualitatively the costs and benefits of the methodological families surveyed, obtained by combining the information reported in the literature, while Fig. 5 shows the adoption trends of the families from 2020 to 2024, in line with the evolution of the literature toward hybrid and explainable architectures during this time. A minority of the surveyed sources reported on explicit deployment and/or simulated real-time evaluation, while the majority used the batch evaluation method with historical labeled data sets, which has material implications in claims of production readiness [1], [4], [11], [17], [18].

TABLE III. QUALITATIVE COST-BENEFIT COMPARISON ACROSS MODEL FAMILIES

Dimension	Classical ML	Deep Learning	Ensemble Hybrid	Federated + XAI
Training Cost	Low	High	Medium-High	High
Inference Latency	Low	Medium-High	Medium	Medium
Interpretability	High	Low	Medium	Medium-High
Regulatory Alignment	Medium	Low	Medium	High
Scalability to Volume	Medium	High	High	Medium-High

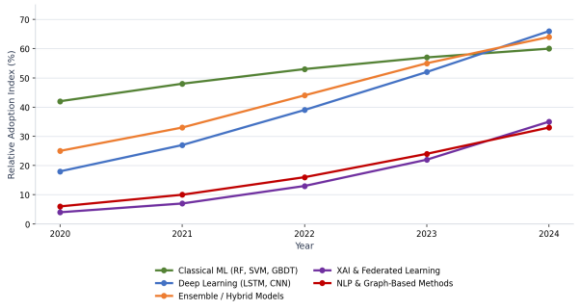

Fig. 5. Reported adoption trajectory of methodological families, 2020-2024.

VI. EXPLAINABILITY, FEDERATED LEARNING, AND REGULATORY CONSIDERATIONS

Along with the increased use of machine learning in financial decision making, the regulatory examination of automated financial decision making has grown faster as well, and there has been a continued interest in the development of explainable artificial intelligence (XAI) as a means of balancing predictive accuracy and the auditability of automated financial decision making. For financial fraud detection, Awosika et al. specifically looked at the combination of explainable AI and federated learning, both for the data-residency and privacy concerns in distributed banking systems and the transparency they offer when making decisioning decisions for regulatory and customer use cases [5]. Fig. 6 shows a typical federated learning process where transaction data is only shared with the respective institutions and locally trained models, transaction data is aggregated securely, then passed through an explainability layer before being presented to an analyst or regulator. Similarly, Nobel et al. directly integrated explainability methods such as feature-attribution in an ensemble classification pipeline, concluding that integrating explainability does not noticeably affect the recall performance of the minority classes but significantly boosts the trust of the analysts with the flagged transactions for banking fraud classification [17]. Oztas et al. performed a qualitative analysis of transaction monitoring practices for anti-money-laundering compliance, based on industry experience, and concluded that practitioners generally consider the explainability and auditability aspects of candidate machine learning systems more important than any marginal benefits in terms of increased accuracy for production use [18]. This is supported in the broader survey literature, and Rao and Mandhala's review of machine learning and data-mining methods for detecting financial fraud includes regulatory alignment in their list of key pre-requisites for adoption along with detection performance [20].

Another structural challenge that comes with mobile banking is that fraud schemes often run across banks, whereas sharing data between banks is limited by privacy policies and competition. Federated architectures were reported to enhance fraud detection patterns that are not observed with a single-institution model but only through collaborative model training in the absence of centralized data pooling [5], [15].

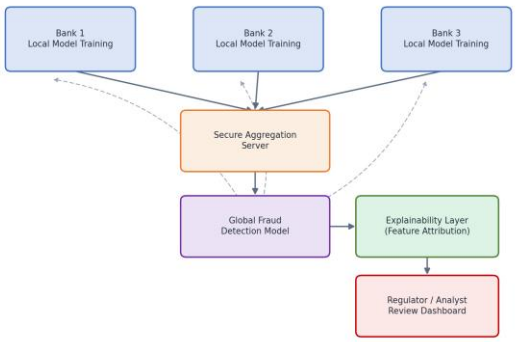

Fig. 6. Federated learning pipeline with explainability layer for cross-institutional fraud detection.

VII. Challenges, Adoption Barriers, and Future Directions

Existing research shows the benefits of machine learning-based fraud detection in mobile banking, but the surveyed literature reveals that there are still some issues that come across in the production rollout of such a system. Class imbalance is also a core issue, as some research indicates that models tuned for accuracy tend to perform poorly on minority-class instances unless they are explicitly cost-sensitive or resampled [1, 17]. In addition, a form of gradual or sudden change in fraud typology in response to fraud detection countermeasures, known as concept drift, was identified as a constraint; it requires continuous retraining pipelines and monitoring infrastructure, which are not always available in the surveyed deployments [18], [20]. Infrastructural and computational limitations are reported as significant in mobile-money systems with less developed infrastructure, such as those that span multiple geographic regions, where the latency in scoring is often short and the challenges of deep learning and ensemble architectures are high [3], [8], [13]. Looking at enterprise financial fraud detection, Ismail and Haq found that the ability to deploy models was crucial and required significant integration with existing transaction processing infrastructure, not necessarily advanced algorithms [12]. Similarly, Nama and Obaid pointed out that deep learning methods were effective but had inventories of data and labelling costs that were not consistently available for financial institutions of varying sizes [16]. The literature surveyed so far to 2024 suggests three trends that will come together in the future. First, hybrid methods that incorporate ensemble tree-based approaches with selective deep learning components seem to offer a tradeoff between accuracy and interpretability and latency constraints [11],[17],[21]. Second, federated and privacy-preserving learning will likely gain in popularity as cross-institutional fraud starts to become more important and as data-privacy laws become stricter worldwide [5]. Third, natural language and graph-based approaches, though comparatively new in the mobile-banking domain more generally, apply to the types of transactions that contain narratives and fraud rings that are relational, thus showing continued methodological convergence with the wider financial-crime detection literature [9], [15], [24].

VIII. Conclusion

This paper collates twenty-four sources that cover fraud detection in mobile banking, mobile money and related digital payment environments, up to 2024, using machine learning. Classical supervised algorithms are still relevant and useful for their interpretability and latency capabilities while deep learning and graph-based models offer more sophisticated capabilities of detecting sequential and relational fraud patterns that are not natively supported by classical feature-based classifiers. The analyses of the surveyed literature revealed high aggregate performance of ensemble and hybrid frameworks, which were reported by a large number of studies and additionally showed a convergence in methodology toward architectures that balance the accuracy, interpretability, and regulatory alignment. Considered jointly, Explainable AI and federated learning appear not to be mutually exclusive, but can be complementary priorities, with one tackling transparency and the other privacy concerns, while maintaining a strong detection ability. Although the issues addressed in the analyzed literature were generally solved, they were not fully resolved, highlighting the need for further ongoing investments in terms of methodology and infrastructure as mobile banking continues to grow as the major channel of retail financial services.

References

[1] J. K. Afriyie et al., "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," Decision Analytics Journal, vol. 6, p. 100163, 2023.

[2] A. H. Ali and A. Ali Hagag, "An enhanced AI-based model for financial fraud detection," Int. J. Adv. Appl. Sci., vol. 11, no. 10, pp. 114–121, 2024.

[3] R. Al-Sayyed, E. Alhenawi, H. Alazzam, A. Wrikat, and D. Suleiman, "Mobile money fraud detection using data analysis and visualization

techniques," Multimedia Tools Appl., vol. 83, no. 6, pp. 17093–17108, 2024.

[4] A. A. Almazroi and N. Ayub, "Online payment fraud detection model using machine learning techniques," IEEE Access, vol. 11, pp. 137188–137203, 2023.

[5] T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy: The role of explainable AI and federated learning in financial fraud detection," IEEE Access, vol. 12, pp. 64551–64560, 2024.

[6] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," J. Big Data, vol. 8, no. 1, art. 6, 2021.

[7] S. S. Bodade and P. P. Pawade, "Review paper on UPI fraud detection using machine learning," Int. J. Res. Appl. Sci. Eng. Technol., vol. 11, no. 12, 2023.

[8] F. E. Botchey, Z. Qin, and K. Hughes-Lartey, "Mobile money fraud prediction—A cross-case analysis on the efficiency of support vector machines, gradient boosted decision trees, and Naive Bayes algorithms," Information, vol. 11, no. 8, p. 383, 2020.

[9] P. Boulieris, J. Pavlopoulos, A. Xenos, and V. Vassalos, "Fraud detection with natural language processing," Mach. Learn., vol. 113, no. 8, pp. 5087–5108, 2024.

[10] D. Bwalya and J. Phiri, "Fraud detection in mobile banking based on artificial intelligence," in Artificial Intelligence Application in Networks and Systems, Lecture Notes in Networks and Systems, vol. 724, Springer, 2023.

[11] P. Hajek, M. Z. Abedin, and U. Sivarajah, "Fraud detection in mobile payment systems using an XGBoost-based framework," Inf. Syst. Front., vol. 25, no. 5, pp. 1985–2003, 2023.

[12] M. M. Ismail and M. A. Haq, "Enhancing enterprise financial fraud detection using machine learning," Eng. Technol. Appl. Sci. Res., vol. 14, no. 4, pp. 14854–14861, 2024.

[13] M. E. Lokanan, "Predicting mobile money transaction fraud using machine learning algorithms," Appl. AI Lett., vol. 4, no. 2, art. e85, 2023.

[14] M. Manala and J. Jansen van Vuuren, "Machine-learning phishing detection model used in the e-banking environment," in Human Choice and Computers, IFIP AICT, vol. 719, Springer, 2024.

[15] S. Motie and B. Raahemi, "Financial fraud detection using graph neural networks: A systematic review," Expert Syst. Appl., vol. 240, p. 122156, 2024.

[16] F. A. Nama and A. J. Obaid, "Financial fraud identification using deep learning techniques," Al-Salam J. Eng. Technol., vol. 3, no. 1, pp. 141–147, 2024.

[17] S. M. N. Nobel et al., "Unmasking banking fraud: Unleashing the power of machine learning and explainable AI (XAI) on imbalanced data," Information, vol. 15, no. 6, p. 298, 2024.

[18] B. Oztas et al., "Transaction monitoring in anti-money laundering: A qualitative analysis and points of view from industry," Future Gener. Comput. Syst., vol. 159, pp. 161–171, 2024.

[19] M. Rahman, H. P. Yee, M. A. K. Masud, and M. U. H. Uzir, "Examining the dynamics of mobile banking app adoption during the COVID-19 pandemic: A digital shift in the crisis," Digit. Bus., vol. 4, no. 2, p. 100088, 2024.

[20] R. K. Rao and V. N. Mandhala, "Unveiling financial fraud: A comprehensive review of machine learning and data mining techniques," Ing. Syst. Inf., vol. 29, no. 6, pp. 2309–2334, 2024.

[21] F. Wahab, I. Khan, and S. Sabada, "Credit card default prediction using ML and DL techniques," Internet Things Cyber-Phys. Syst., vol. 4, pp. 293–306, 2024.

[22] R. Wajgi, H. Agarkar, R. Patil, H. Rao, and N. Petkar, "Enhancing credit card transaction fraud detection with random forest and robust scaling," AIP Conf. Proc., vol. 3188, no. 1, art. 040013, 2024.

[23] G. Yang, "Credit card fraud detection based on machine learning prediction," in Proc. 2nd Int. Conf. Image, Algorithms Artif. Intell. (ICIAAI), pp. 35–45, 2024.

[24] X. Yang et al., "FinChain-BERT: A high-accuracy automatic fraud detection model based on NLP methods for financial scenarios," Information, vol. 14, no. 9, p. 499, 2023.