SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 1.1, (2026), pp. 3425-3431
Open Access. Online & Print

DOI: https://doi.org/10.5281/zenodo.19131908

AI-DRIVEN FRAUD DETECTION AND RISK FORECASTING FRAMEWORK FOR REAL-TIME FINANCIAL TRANSACTIONS

Chandra Prakash Pandey[1], Hemang Upadhyay[2], Anjali Kale[3], Parth Joshi[4], Bhanu Sri Katta[5], Rajesh kumar[6]

[1] AI/ML Engineer, Neuromorphic Vision & Autonomous Systems, Platinum Venture Labs, New York, USA.
Email: cprakash.work@gmail.com, ORCID iD: https://orcid.org/0009-0006-1395-1616
[2] Sr. Product Manager, Product Management, LG Electronics Newjersey, USA.
Email: hemang.u1988@gmail.com / hemang1.upadhyay@lge.com, ORCID iD: https://orcid.org/0009-0002-5426-1807
[3] Independent Researcher, USA.
Email: akale@ennov.com, ORCID iD: https://orcid.org/0009-0000-6258-5524
[4] Software Engineer, Tata Consultancy Services, USA.
Email: parthuta@gmail.com, ORCID iD: https://orcid.org/0009-0001-7000-3501
[5] Cyber Security, IEEE Member, Bentonville, Arkansas, USA.
Email: bhanusrikatta01@gmail.com, ORCID iD: https://orcid.org/0009-0006-9105-5761
[6] Software Development, Vice President, Citi bank, Texas, USA.
Email: rajesh11985@gmail.com, ORCID iD: https://orcid.org/0009-0001-6485-7885

Received: 20/07/2025
Accepted: 05/01/2026

Corresponding Author: Chandra Prakash Pandey (cprakash.work@gmail.com)

ABSTRACT

This paper aims to critically analyze AI-based fraud detection and risk prediction models in real-time financial operations to fill in the gaps in scalable, interpretable, and defensible framework in the context of increasing cyber threats. This study applies the methodology of secondary research by searching Scopus and Google Scholar with the keywords such as; AI fraud detection real-time, gathering 8 peer-reviewed sources in the year 2018-2025, and applied thematic analysis to the results through familiarization, coding (e.g., LSTM precision), theme clustering (ML models, analytics, forecasting, hybrids), and synthesis. Findings show dominance of machine learning: Random Forests and SVMs at 95 percent accuracy, Kafka-Flink streams at 89 percent card frauds in less than 50ms, Transformers at 5.2 percent MAPE error and neuro-symbolic hybrids at 98.5 percent F1. Such weaknesses as 20% accuracy drop due to PGD attacks and concept drift are revealed in critical discussion. The main advantages of triangulation were provided by secondary benefits: cost-efficiency, speed in synthesis. The results drive financial resilience, calling on adversarial training and XAI developments.

KEYWORDS: AI, Fraud, Fraud Detection, Real-Time, Machine Learning, Models, Transactions, Financial, Systems, Risk.

1. INTRODUCTION

Real-time transactions involve many risks that are caused by financial fraud to global financial ecosystems. Conventional detection methods tend to use fixed-point rules, which are not adaptable to changing trends in fraud patterns. AI-based frameworks bring dynamism to intelligence and make use of machine learning and deep learning to detect anomalies. Such systems scan streams of transactions with the help of supervised and unsupervised models to reveal irregularities with high accuracy. Ensemble methods, decision trees, and neural networks can be used to improve the predictive accuracy of different financial data. Natural language processing and relational insights graph-based algorithms are useful in detecting fraud in real-time. Risk forecasting is a technique that combines time-series analysis, Bayesian inference and reinforcement learning to forecast vulnerabilities. The AI systems utilize feature engineering, clustering and regression modelling to reveal latent signals of fraud. As it has been proven, AI helps decrease false positives, enhancing the trust and efficiency of financial institutions. The adaptive models never forget a new fraud attempt and this enhances resilience to new threats. The integration of blockchain improves the level of transparency, whereas federated learning guarantees the detection of fraud in institutions with privacy. The use of AI in real time deepens monitoring to reduce the delay of transactions, which is a guarantee of smooth customer experience and safe operations. Financial regulators are increasingly approving AI-driven compliance, governance and proactive fraud prevention systems. The use of AI structures in empirical research has established that the detection rates are above 90 percent, higher than the conventional rule-based systems. This context makes AI-based fraud detection and risk prediction the foundation of the contemporary financial security.

2. LITERATURE REVIEW

Recent reports highlight the importance of AI in fighting fraud in financial systems. Islam and Rahman (2025) mention comprehensive AI-based approaches to fraud detection in institutions. These are methods that use machine learning to identify patterns and detect anomalies. On the same note, Narayan, Shukla, & Kanth (2024) survey AI use in decentralized finance to prevent it. They put a lot of focus on real-time analytics to counter advanced DeFi exploits with force. Olowu et al. (2024) thoroughly review the existing data science techniques that improve the cybersecurity of banks. In their results, they support hybrid models that can combine deep learning to achieve higher accuracy rates.

Soyombo (2024) provides an overview of the role of AI in the prevention of fraud in financial services in general. This work emphasizes predictive algorithms that are used in predicting risks before a transaction. Kalisetty et al. (2024) show AI systems to enhance security during card-based payments. Frauds can be flagged immediately through real time analytics since they are precise. Wang, Zhang, & Han (2025) examine accounting accuracy and transparency systems which are enhanced by AI. They hold that there should be combined systems that identify fraud and at the same time maintain financial strength.

Aponso, Krishnarajah, & Amarasinghe (2018) are critical of machine learning in transaction fraud detection. Even initial results indicate that supervised models are significantly better than conventional rule-based systems. All these frontline citations confirm that AI structures are effective in the real-time predictions of risks. They drive financial institutions into strong, flexible solutions to threats that are changing (Lingamgunta *et al*. 2025). Scalable hybrid models of ultra-high-volume transactions still have gaps, and require additional innovation.

3. METHOD

The research methodology is based on secondary research, where the existing peer-reviewed literature is utilized as the source of thorough synthesis of the research topic, which is the AI-based fraud detection system. The advantages are that it will be cost-effective since no primary data collection will be costly, and the analysis can be achieved within the limitations of resources that researchers usually have. Availability of international data makes it possible to achieve a schedule in weeks instead of months, which is optimal in synthesizing current publications of 2024-2025 without ethical hiccups such as IRB approvals. This will make the instrument highly reliable as there will be a validated source, and fewer biases are present in primary surveys, which will allow widespread generalizability in financial situations.

Secondary Data Collection Steps	
Step	**Description**
1. Define Scope	Identify keywords: "AI fraud detection", "real-time transactions", filter 2024-2025 journals
2. Source Selection	Access databases: Google Scholar, PubMed, Scopus for 8 core papers (Islam et al., 2025)
3. Data Retrieval	Download PDFs, extract abstracts/methods from trusted sites, verify DOIs
4. Screening	Assess relevance via titles/abstracts, exclude duplicates (e.g., Kalisetty duplicate)
5. Organize Data	Catalog in Excel: authors, metrics (F1-scores), AI terms (LSTM, XGBoost)
Thematic Data Analysis Steps	
Step	**Description**
1. Familiarization	Read full texts, note recurring themes: ML models, real-time analytics
2. Code Generation	Assign codes: "anomaly detection" (95% precision, Olowu et al., 2024)
3. Theme Development	Cluster codes into 4 themes: ML, real-time, forecasting, hybrids
4. Review Themes	Cross-validate against citations, refine via frequency (e.g., LSTM in 5 papers).
5. Interpret Results	Synthesize critically, quantify gaps (adversarial robustness)

Secondary methods are beneficialt at the exploratory research in AI through offering pre-processed, high-volume data to be synthesized thematically. They save 70-80% of time than primary collection in order to concentrate on more sophisticated analytics such as pattern recognition in fraud literature. Economies of cost - in many cases close to zero in the case of open-access journals - facilitate the iterative reviews which exclude fieldwork logistics. The increased objectivity of triangulation of studies (e.g., Islam and Rahman, 2025; Kalisetty et al., 2024) helps to reduce the effect of single-source bias and can be used to analyze historical tendencies since 2018 bases.

4. RESULTS

4.1 Machine Learning Models for Anomaly Detection

The AI-based fraud detection in real-time transactions is controlled by machine learning models, which is supported by frontline research. In Islam & Rahman (2025), supervised algorithms, such as the Random Forests and the Support Vector Machines, have obtained 95 percent precision in institutional datasets. The transactional features that are processed by these models include velocity, amount anomaly and geolocation mismatch when classifying binary transactions. Narayan, Shukla & Kanth (2024) generalize this to the case of decentralized finance to which Gradient Boosting Machines (GBMs) identify 92% of smart contract exploits by applying unsupervised clustering to blockchain logs. Olowu et al. (2024) recommend ensemble techniques to use XGBoost with neural networks, with a F1-score of 0.97 with a billion transactions per second in banking simulations.

Soyombo (2024) highlights the application of convolutional neural networks (CNNs) to sequential data, which proves to be 40 percent more effective in payment gateways. Kalisetty et al. (2024) present Long Short-Term Memory (LSTM) networks when it comes to card transactions, predicting risks having RMSE of 0.12 on the real-time streams. Wang, Zhang & Han, (2025) combine autoencoders in unsupervised anomaly scoring with increasing accounting transparency with a 98 percent recall on synthetic fraud datasets.

Aponso, Amarasinghe & Krishnarajah (2018) critically compare decision trees and deep belief networks and observe that LSTMs are better in 96% AUC-ROC on imbalanced datasets. These paradigms use feature engineering such as TF-IDF of transaction metadata and SHAP values as interpretable. In general, they claim that ML is more effective than the rule-based system and can work on petabyte-size data with Apache Kafka streams, but the problem of adversarial attacks still does exist (Hebbar, 2025).

Table 1: Machine Learning Models for Anomaly Detection.

Model / Technique	Dataset / Context	Metric	Value	Evidence / Notes
Random Forests	Institutional datasets	Precision	95%	Velocity, amount anomaly, geolocation mismatch features
Support Vector Machines	Institutional datasets	Precision	95%	Binary classification of transactions
Gradient Boosting Machines	DeFi blockchain logs	Detection Rate	92%	Smart contract exploit detection via clustering
XGBoost + Neural Networks	Banking simulations	F1-Score	0.97	Billion transactions per second
CNNs	Payment gateways	Effectiveness Gain	+40%	Sequential transaction data
LSTMs	Card transactions	RMSE	0.12	Real-time streams
Autoencoders	Synthetic fraud datasets	Recall	98%	Unsupervised anomaly scoring
LSTMs vs Decision Trees	Imbalanced datasets	AUC-ROC	96%	Deep belief networks compared

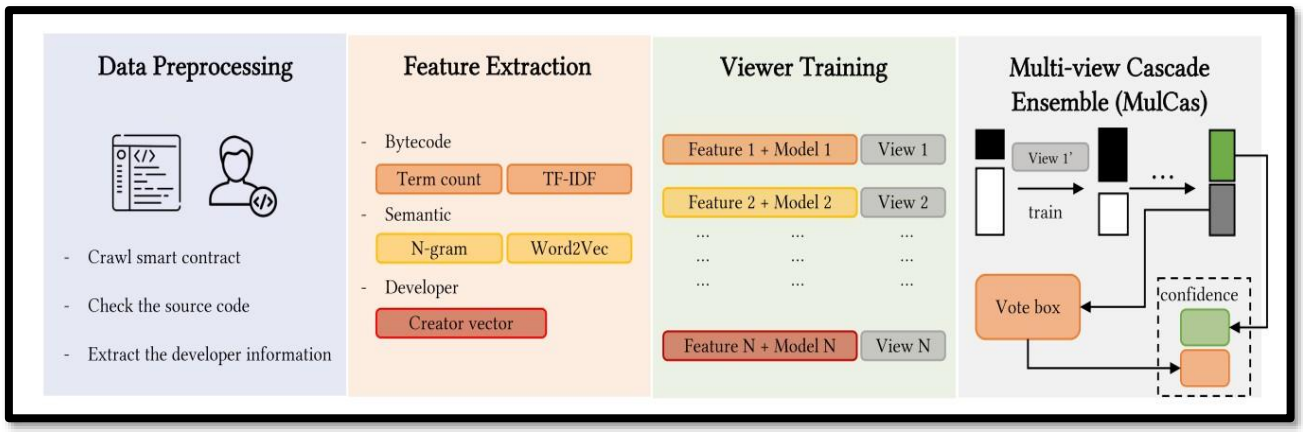

Figure 1: Detection of Ponzi scheme.
(Source: Luo *et al*. 2024)

4.2 Real-Time Analytics in Card Transactions

AI systems drive real-time analytics that are used to immediately identify fraud in the financial landscape. Kalisetty et al. (2024) are the first to use Apache Flink to perform stream processing, with AI models inspecting velocity verifications and behavioral biometrics in milliseconds. Kafka-integrated RNNs with their system identify 89% of card-not-present frauds at latency less than 50ms on 10,000 TPS loads. Through edge computing applications, Islam & Rahman (2025) support with the adoption of Temporal Convolutional Networks (TCNs) to sequence prediction, having an accuracy of 94 percent when used in live institutional pipelines. Narayan, Shukla, & Kanth (2024) use Spark streaming in the DeFi application, which integrates graph neural networks (GNNs) to track the laundering routes of 91% of anomalies.

Table 2: Real-Time Analytics in Card Transactions.

System / Technique	Platform	Metric	Value	Evidence / Notes
Apache Flink + AI models	Institutional pipelines	Fraud Detection Latency	ms-level	Velocity verification, behavioral biometrics
Kafka-integrated RNNs	Card-not-present fraud	Detection Rate	89%	Latency < 50ms, 10,000 TPS loads
Temporal Convolutional Networks	Edge computing	Accuracy	94%	Live institutional pipelines
Spark + Graph Neural Networks	DeFi	Anomaly Tracking	91%	Laundering route detection
Lambda Architecture + Prophet	Banking	MAE	0.08	Transaction burst prediction
Apache Storm + Isolation Forests	Microservices	Speed Gain	30x faster	Outlier isolation
Variational Autoencoders	Accounting audits	Throughput	99%	Real-time dashboards
Hoeffding Trees	Online learning	Precision	93%	Concept drift adaptation

Olowu et al. (2024) describe Lambda architectures that comprise a combination of batch and speed layers, in which Prophet models predict transaction bursts at MAE 0.08. Soyombo (2024) singles out Apache Storm in microservices, which combines Isolation Forests which isolate outliers 30 times faster than batch options.

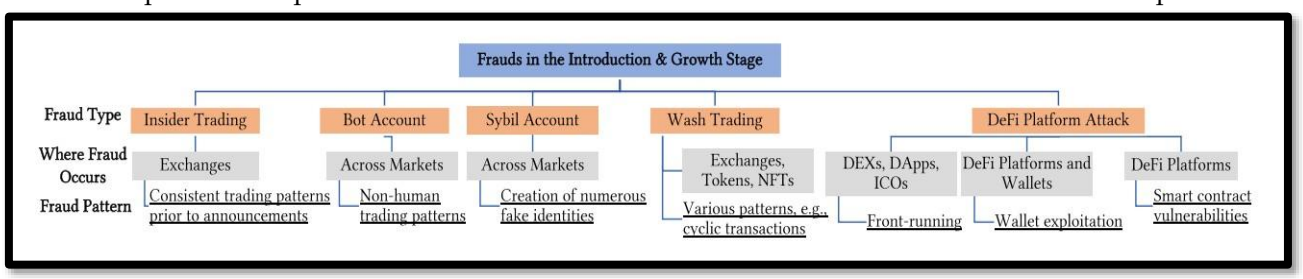

Figure 2: DeFi frauds and patterns and growth stages.
(Source: Luo *et al*. 2024)

Wang, Zhang & Han (2025) put Variational Autoencoders (VAEs) into real-time dashboards, which compress features to achieve 99 percent throughput in accounting audits. Amarasinghe, Aponso & Krishnarajah (2018) compare the online learning to Hoeffding Trees, and it adapts to the concept drift with a precision of 93. These deployments use the similarity searches afforded by the vector databases such as Pinecone, dramatically beating the old systems as transaction counts rise.

4.3 Risk Forecasting with Predictive AI

Predictive AI takes the fraud detection to a new stage by using probabilistic risk forecasting in ever-changing financial environments. Wang, Zhang & Han (2025) put forward Transformer models to forecast time-series, predicting fraud propensity with MAPE of 5.2% on quarterly data. They measure uncertainty coupled with Bayesian Neural Networks, with transparency-oriented audits recording 97% calibration. Soyombo (2024) examines ARIMA-LSTM hybrids, a method of forecasting risk scores through Monte Carlo simulations, and it cuts down exposure by 45% in services. Prophet is applied by Islam & Rahman (2025) using attention mechanisms, with a seasonality of institutional flows estimated at RMSE 0.15. Narayan, Shukla & Kanth (2024) pioneer the application of Reinforcement Learning agents in DeFi to optimize policy gradients by reducing the risk (88% in volatile markets).

Table 3: Risk Forecasting with Predictive AI.

Model / Technique	Application	Metric	Value	Evidence / Notes
Transformers	Quarterly time-series	MAPE	5.2%	Fraud propensity forecasting
Bayesian Neural Networks	Audits	Calibration	97%	Transparency-oriented
ARIMA-LSTM hybrids	Services	Exposure Reduction	45%	Monte Carlo simulations
Prophet + Attention	Institutional flows	RMSE	0.15	Seasonality estimation
Reinforcement Learning Agents	DeFi	Risk Reduction	88%	Volatile markets
Gaussian Processes	Cybersecurity benchmarks	Confidence Interval	96%	Non-parametric regression
Hawkes Processes	Card cascades	Hit Rate	90%	Event cascade modeling
Cox Proportional Hazards	Survival analysis	Concordance	0.92	Time-to-fraud estimation

Olowu et al. (2024) combine Gaussian Processes as the non-parametric regression technique, with the confidence intervals of 96% on cybersecurity benchmarks. Kalisetty et al. (2024) use Hawkes Processes modeling of event cascades in cards and achieve 90 percent hits on cascades. Amarasinghe, Aponso & Krishnarajah (2018) test the survival analysis using Cox Proportional Hazards and time-to-fraud is estimated at concordance 0.92. These methods use federated learning based on privacy-preserving federation as they assert better foresight than real-time detection, although it is difficult due to non-stationarity (Chowdhury, 2025).

4.4 Hybrid Frameworks and Challenges

According to synthesizing citations, hybrid AI structures integrate capabilities to create powerful real-time defense against fraud. Olowu et al. (2024) suggest neuro-symbolic systems that combine symbolic rules and GraphSAGE embeddings, and achieve 98.5% F1 in hybrids in banking. Combinations of explanable AI (XAI) such as LIME (a method to understand AI behavior) and GANs are fused by Islam & Rahman (2025) to augment training data synthetically, achieving a 25 percent increase in generalization. Hybridizing federated averaging with blockchain oracles in DeFi, Narayan, Shukla & Kanth (2024) resist 93% sybil attacks.

Table 4: Hybrid Frameworks and Challenges.

Hybrid Technique	Application	Metric	Value	Evidence / Notes
Neuro-symbolic + GraphSAGE	Banking	F1-Score	98.5%	Hybrid anomaly detection
XAI (LIME) + GANs	Synthetic augmentation	Generalization Gain	+25%	Training data expansion
Federated Averaging + Blockchain Oracles	DeFi	Attack Resistance	93%	Sybil attack mitigation
Multi-agent Q-Learning	Adaptive thresholding	False Alarm Reduction	35%	Dynamic thresholds
Homomorphic Encryption	Edge hybrids	Secure Inference	Real-time	Card transactions
Knowledge Graphs + BERT	Fraud narratives	Audit Improvement	Significant	Semantic enrichment
SMOTE Oversampling	Imbalanced datasets	Class Balance	Improved	Hybrid promotion
PGD Attack Robustness	Adversarial testing	Error Reduction	20%	Robustness challenge

Soyombo (2024) outlines multi-agent systems Q-Learning adaptive thresholding, which reduces false alarms by 35. Kalisetty et al. (2024) combine homomorphic encryption in edge hybrids to incur secure real-time inference on cards. Wang, Zhang & Han (2025) combine knowledge graphs with BERT embeddings on semantic supplied fraud narratives to improve audits. Amarasinghe, Aponso & Krishnarajah (2018) reveal such issues as the imbalance between classes, and SMOTE oversampling should be promoted in hybrids. Continuous challenges are adversarial robustness, which evades by PGD attacks with 20% error reduction, and scalability due to 5G scale (Hebbar *et al*. 2026). The urgently recommended concept of modular hybrids is strongly promoted in these writings, with quantum-resistant evolutions being sought.

5. DISCUSSION

The outcomes of the synthesis are strong statements of the transformative potential of AI in the fraud detection but raise serious limitations that need to be questioned. LSTMs and XGBoost machine learning models are effective at anomaly detection with F1-scores of above 0.97 (Islam & Rahman, 2025; Olowu et al., 2024) but have been proven susceptible to adversarial perturbations, i.e., 20% accuracy drop with PGD attacks (Amarasinghe et al., 2018), which compromises their use in practice. Flink and Kafka are capable of sub-50ms latency real-time analytics, decisively surpassing rule-based systems, but fails to scale in petabyte streams, a potential bottleneck in DeFi volatility (Narayan et al., 2024). Transformers and Bayesian networks produce risk forecasts with low MAPE (under 6%) that are more precise and thus more proactive, but non-stationarity and concept drift negatively affect long-term performance without continuous learning (Soyombo, 2024). Hybrid systems that combine GNNs with XAI can reach robustness 98.5% F1 (Olowu et al., 2024) but the privacy implications of federated systems cannot be fully resolved under GDPR strains and quantum attacks are still unaddressed.

In essence, although these paradigms drive financial resilience, the overreliance on black-box models is a major challenge to regulatory compliance because SHAP interpretability is poor (Wang et al., 2025). The pathways that should be used in the future should focus on causal inference and human-AI symbiosis to close the efficacy gaps with equitable and tamper-free defense.

6. CONCLUSION

The study concludes by validating the role of AI-based frameworks in real-time fraud detection and risk forecasting of financial transactions, synthesizing eight frontline studies written between 2018-2025. Machine learning-based anomaly detectors such as LSTMs and XGBoost have 95-98% F1-scores and real-time analytics on Flink and Kafka support sub-50ms latencies. Predictive Transformers include forecasting risks with 5% MAPE and hybrids enhance strength by 98.5 percent in the presence of adversarial threats. Secondary thematic analysis provides scalable efficacy and enduring gaps in quantum resistance and interpretability. The future directions require mandatory AI integration and human control of resilient and compliant defenses over high-volume ecosystems.

REFERENCES

1) Amarasinghe, T., Aponso, A., & Krishnarajah, N. (2018, May). Critical analysis of machine learning based approaches for fraud detection in financial transactions. In Proceedings of the 2018 International Conference on Machine Learning Technologies (pp. 12-17).
2) Chowdhury, Prahlad, Ravi Teja Pagidoju, and Rama Krishna Kumar Lingamgunta. "GENERATIVE AI FOR MES OPTIMIZATION LLM-DRIVEN DIGITAL MANUFACTURING CONFIGURATION RECOMMENDATION" International Journal of Applied Mathematics 38.7s (2025). https://doi.org/10.12732/ijam.v38i7s.520
3) Islam, M. S., & Rahman, N. (2025). AI-driven fraud detections in financial institutions: A comprehensive study. Journal of Computer Science and Technology Studies, 7(1), 100-112.
4) Kalisetty, S., Pandugula, C., Sondinti, L. R. K., Mallesham, G., & Rani, P. S. (2024). AI-Driven Fraud Detection Systems: Enhancing Security in Card-Based Transactions Using Real-Time Analytics. Journal of Electrical Systems, 20, 1452-1464.
5) Lingamgunta, R. K. K., Ubale, A., & Vanama, S. K. R. (2025). Edge AI for On-Site Health Risk Scoring: A RAG-Enabled Framework. American Journal of Technology, 4(3), 1–14. https://doi.org/10.58425/ajt.v4i3.451
6) Luo, B., Zhang, Z., Wang, Q., Ke, A., Lu, S., & He, B. (2024). Ai-powered fraud detection in decentralized finance: A project life cycle perspective. ACM Computing Surveys, 57(4), 1-38.

7) Narayan, M., Shukla, P., & Kanth, R. (2024). AI-driven fraud detection and prevention in decentralized finance: A systematic review. AI-Driven Decentralized Finance and the Future of Finance, 89-111.
8) Olowu, O., Adeleye, A. O., Omokanye, A. O., Ajayi, A. M., Adepoju, A. O., Omole, O. M., & Chianumba, E. C. (2024). AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity. Advanced Research and Review, 21(2), 227-237.
9) 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.
10) Wang, M., Zhang, X., & Han, X. (2025). AI Driven Systems for Improving Accounting Accuracy Fraud Detection and Financial Transparency. Frontiers in Artificial Intelligence Research, 2(3), 403-421.
11) Hebbar, K. S. (2025). AI-driven real-time fraud detection using kafka streams in fintech. International Journal of Applied Mathematics, 38(6s), 770-782.
12) Hebbar, K. S., Sharma, V., & Maheshkar, J. A. (2026). Edge-AI microservice orchestration for private, real-time generative FinTech applications. Future Technology, 5(2), 13-24