Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 21, No. 1, 2026

Real Time Fraud Detection at Scale in High Volume Enterprise Payment Ecosystems

Yuxuan Qin [1, *], Jiesi Yang [2], and Zimeng Wang [3]

[1] Northeastern University, United States
[2] University of Utah, United States
[3] Brandeis University, United States

*** Corresponding author:** qin.yux@northeastern.edu

Abstract: The rapid proliferation of digital payment channels has created unprecedented opportunities for fraudulent activity, compelling enterprise organizations to deploy increasingly sophisticated detection mechanisms capable of operating at high throughput with minimal latency. This paper presents a comprehensive review of real-time fraud detection (RTFD) methodologies applied within high-volume enterprise payment ecosystems, encompassing machine learning (ML), deep learning (DL), graph neural networks (GNN), and streaming data architectures. We examine how ensemble methods, anomaly detection frameworks, and feature engineering pipelines converge to form robust, production-grade fraud detection systems (FDS). The paper further discusses the tension between detection accuracy and operational latency, the challenge of class imbalance in transactional datasets, and the evolving regulatory landscape that shapes deployment constraints. By synthesizing findings from recent literature, this review identifies key trends including federated learning (FL) for privacy-preserving fraud detection, transformer-based sequence models for behavioral analysis, and adaptive threshold mechanisms for dynamic fraud pattern recognition. Our analysis reveals that no single algorithmic approach suffices in isolation; rather, layered architectures combining rule-based systems with data-driven models consistently achieve superior performance across precision, recall, and throughput metrics in enterprise-scale deployments.

Keywords: Real-time fraud detection; Enterprise payment systems; Machine learning; Deep learning; Graph neural networks; Streaming analytics; Anomaly detection; Federated learning.

1. Introduction

The global digital payments industry processed transactions exceeding 8.26 trillion US dollars in 2023, a volume that continues to expand at a compound annual growth rate placing extraordinary pressure on fraud prevention infrastructure [1]. Within high-volume enterprise payment ecosystems, the challenge of detecting fraudulent activity in real time is not merely a technical problem but a strategic imperative, as undetected fraud erodes consumer trust, incurs direct financial losses, and triggers regulatory penalties that can fundamentally threaten business continuity [2]. Enterprise environments such as interbank clearing networks, e-commerce payment gateways, and cross-border remittance platforms routinely handle millions of transactions per hour, demanding FDS that deliver sub-second decisions without sacrificing classification accuracy [3].

Traditional approaches to payment fraud relied heavily on static rule-based engines, where domain experts encoded known fraud signatures as conditional logic applied to each incoming transaction. While these systems offered interpretability and deterministic behavior, they proved brittle against the adaptive tactics of modern fraud actors who systematically probe detection boundaries and mutate their behavioral footprint to evade fixed thresholds [4]. The limitations of rule-based systems became particularly acute following the widespread adoption of card-not-present transactions in e-commerce, where the absence of physical authentication dramatically expanded the attack surface available to fraudsters [5].

The emergence of ML as a viable alternative gained substantial momentum during the early 2010s, with gradient boosted decision trees demonstrating compelling performance on transactional datasets characterized by severe class imbalance [6]. However, the transition from offline batch classification to genuine real-time streaming inference introduced new engineering challenges around feature freshness, model serving latency, and the consistency of behavioral features computed over sliding time windows [7]. These challenges intensified as payment volumes scaled into regimes where even microsecond improvements in pipeline throughput translated into measurable business impact.

DL methods subsequently expanded the representational capacity available to fraud detection practitioners, with recurrent architectures capable of modeling sequential spending behavior and convolutional networks able to extract spatial patterns from transaction metadata [8]. More recently, GNN have emerged as a particularly powerful paradigm for enterprise fraud detection, capturing the relational structure of payment networks where fraudulent activity frequently manifests as coordinated behavior across clusters of accounts rather than as anomalies within isolated transaction records [9]. The application of GNN to bipartite graphs connecting cardholders, merchants, and intermediary entities has demonstrated detection rates substantially superior to feature-engineering approaches that treat each transaction independently [10].

Alongside algorithmic innovation, the infrastructure layer governing RTFD has undergone profound transformation. Distributed stream processing frameworks such as Apache Kafka and Apache Flink have enabled organizations to construct event-driven pipelines capable of ingesting, enriching, and scoring millions of transaction events per second with end-to-end latency measured in milliseconds [11].

The deployment of in-memory feature stores has further reduced the latency penalty associated with retrieving aggregated behavioral statistics, enabling fraud models to access rich historical context without sacrificing the real-time performance requirements imposed by payment network service level agreements [12].

The regulatory environment surrounding payment fraud has also evolved substantially. The European Unions Payment Services Directive 2 mandates strong customer authentication and real-time transaction monitoring for a broad class of electronic payments, while the Payment Card Industry Data Security Standard imposes rigorous controls on the handling of cardholder data that constrain the architectures available for fraud model training and inference [13]. In the United States, the Financial Crimes Enforcement Network has expanded suspicious activity reporting obligations in ways that require financial institutions to maintain auditable, explainable fraud detection records, creating tension with the opacity of high-dimensional DL models [14].

This paper provides a structured review of the state of the art in RTFD for high-volume enterprise payment ecosystems. The review proceeds as follows. Section 2 surveys the existing literature across algorithmic and infrastructural dimensions. Section 3 analyzes the core methodological components of production-grade FDS, including feature engineering, model architectures, and streaming pipeline design. Section 4 evaluates performance trade-offs and deployment challenges, drawing on empirical findings from recent large-scale studies. The paper concludes with a synthesis of emerging research directions and practical recommendations for enterprise practitioners.

2. Literature Review

The academic literature on payment fraud detection has expanded dramatically since 2019, reflecting both the growing economic significance of the problem and the maturation of DL tooling that has made sophisticated architectures accessible to applied researchers. Early influential work established benchmark datasets and evaluation protocols that subsequent studies have built upon, though the scarcity of publicly available real-world payment datasets has consistently limited direct comparability across published results [15].

Research on classical ML methods for fraud detection has continued to demonstrate the enduring competitiveness of gradient boosting algorithms. Studies comparing XGBoost, LightGBM, and random forests on large-scale credit card datasets consistently report area under the receiver operating characteristic curve values exceeding 0.97 on balanced evaluation sets, with LightGBM frequently outperforming alternatives on speed-accuracy trade-offs relevant to real-time deployment [16]. The robustness of tree-based ensembles to noisy features and their natural support for categorical transaction metadata have maintained their relevance even as neural architectures have grown more capable [17]. Importantly, the interpretability of feature importance scores derived from gradient boosting models has made them attractive in regulated environments where fraud decisions must be explained to affected cardholders and audited by compliance teams [18].

The class imbalance problem remains a central methodological concern throughout the literature. In enterprise payment environments, fraudulent transactions typically constitute between 0.1 and 2 percent of total volume, creating severe distributional skew that causes naive classifiers to achieve high accuracy by predicting the majority class exclusively [19]. Synthetic minority oversampling, cost-sensitive learning, and threshold calibration have all been studied extensively as remediation strategies, with recent work demonstrating that adaptive synthetic sampling combined with ensemble calibration produces more stable precision-recall trade-offs across distribution shifts than single oversampling techniques applied in isolation [20]. The challenge is compounded in streaming environments where the fraud base rate itself drifts over time in response to seasonal patterns, macroeconomic conditions, and the introduction of new payment products.

Recurrent neural networks and their variants have attracted substantial research attention for modeling the sequential structure of cardholder spending behavior. Related work on hybrid attention-based time series modeling further shows that combining recurrent architectures with attention mechanisms enables effective capture of both long-term dependencies and localized temporal patterns, offering transferable insights for improving sequential behavior modeling in fraud detection systems [21]. Long short-term memory (LSTM) networks were among the earliest DL architectures applied to transaction sequence modeling, capturing temporal dependencies across purchase histories that static feature aggregation cannot represent [22]. Subsequent work demonstrated that bidirectional LSTM configurations improved detection of fraud patterns embedded within sequences of otherwise legitimate transactions, by enabling the model to incorporate both historical context and the short-term behavioral trajectory preceding each decision point [23]. Attention mechanisms further enhanced sequence models by allowing the network to differentially weight transactions within a history window based on their relevance to the current classification decision, improving both accuracy and interpretability in controlled experiments [24].

The application of GNN to payment fraud has produced some of the most compelling empirical results in recent literature. By modeling payment networks as heterogeneous graphs where nodes represent accounts, merchants, devices, and IP addresses, and edges encode transaction relationships, GNN architectures can propagate information across the network to identify coordinated fraud rings that would appear innocuous when each transaction is examined individually [25]. Graph attention networks applied to bipartite transaction graphs have demonstrated particularly strong performance on synthetic fraud ring detection benchmarks, outperforming feature-engineered baselines by substantial margins in recall at operationally relevant false positive rate thresholds [26]. Dynamic GNN that update node representations as new transactions arrive have extended this advantage to streaming environments, though their computational overhead relative to simpler architectures requires careful optimization for production deployment [27].

Transformer architectures originally developed for natural language processing have been adapted for payment fraud detection with considerable success. Self-attention mechanisms over transaction sequences have demonstrated superiority over LSTM baselines on several benchmark datasets, particularly for detecting fraud patterns that require modeling long-range dependencies across extended account histories [28]. Tabular transformers specifically designed for

heterogeneous structured data have shown promise on real-world payment datasets where the mixture of categorical, numerical, and temporal features complicates the application of architectures designed for homogeneous input modalities [29]. The computational demands of transformer inference, however, create non-trivial challenges for real-time deployment at the millisecond latency thresholds required by enterprise payment networks [30].

FL has emerged as a significant research direction motivated by the privacy and regulatory barriers that prevent financial institutions from pooling transaction data for collaborative model training. Under FL frameworks, each participating institution trains local model updates on its proprietary data and contributes only aggregated gradient information to a shared global model, preserving data locality while benefiting from the statistical diversity of distributed training [31]. Studies on federated fraud detection have demonstrated that cross-institutional FL models detect fraud patterns absent from any single institutions training data, particularly synthetic fraud typologies that exploit gaps in institution-specific behavioral baselines [32]. Privacy-preserving extensions using secure multi-party computation and differential privacy mechanisms have been proposed to provide formal guarantees against inference attacks on shared gradient updates, though these protections introduce communication overhead that remains an active area of optimization research [33].

The literature on streaming infrastructure for RTFD has grown substantially alongside advances in distributed systems. Research on Apache Flink-based fraud detection pipelines has characterized the latency-throughput trade-offs achievable at different parallelism configurations, demonstrating that properly tuned streaming architectures can sustain millions of scored transactions per second with median end-to-end latency below 50 milliseconds [34]. Feature store architectures that maintain both online low-latency serving and offline batch computation of behavioral aggregates have been identified as a critical architectural component, and recent advances in retrieval-augmented AI agents further demonstrate that integrating dynamic knowledge retrieval with generative models can enhance system adaptability, maintain up-to-date contextual awareness, and support complex decision workflows in enterprise software environments [35]. The challenge of exactly-once processing semantics in distributed streaming pipelines, necessary to prevent duplicate fraud alerts from multi-hop consumer architectures, has received dedicated treatment in systems literature with practical solutions benchmarked on payment-scale workloads [36].

Adversarial robustness has received growing attention as fraud actors have begun to exploit knowledge of deployed detection models to craft evasion attacks. Research on adversarial examples in the tabular domain demonstrates that gradient-based perturbations to transaction feature values can substantially reduce detection confidence without triggering rule-based safeguards, raising concerns about the security of DL-based fraud systems deployed in adversarial environments [37]. Adversarial training procedures that augment model training with synthetically perturbed examples have shown efficacy in improving robustness against these attacks, though the transferability of robustness improvements across different attack strategies remains incompletely characterized [38]. The game-theoretic framing of fraud detection as a Stackelberg competition between detector and adversary has motivated research on robust optimization approaches that minimize worst-case detection loss rather than average-case loss [39].

Explainability research has addressed the tension between the complexity of high-performing fraud models and the interpretability demands of regulatory compliance and consumer-facing adverse action notification. Shapley Additive Explanations (SHAP) has become the dominant post-hoc explanation method in applied fraud detection literature, providing feature-level attribution scores that satisfy axiomatic fairness properties and have been validated through controlled user studies with fraud analysts [40]. Local Interpretable Model-agnostic Explanations (LIME) have been applied as a complementary approach for generating human-readable rule approximations of DL fraud model behavior in specific transaction contexts [41]. Research on inherently interpretable models, including attention-based neural networks and decision tree distillation from ensemble models, has sought to narrow the accuracy gap between black-box and interpretable architectures without sacrificing the regulatory advantages of transparent decision logic [42].

3. Methodology

The design of a production-grade RTFD system for enterprise payment ecosystems requires careful integration of feature engineering pipelines, model architecture selection, and streaming infrastructure. This section analyzes each of these components with reference to validated approaches from the empirical literature, synthesizing a coherent methodological framework applicable to high-volume operational environments.

Feature engineering represents the foundational layer upon which all downstream classification performance depends. In payment fraud detection, the most predictive signals are not raw transaction attributes but derived behavioral statistics that capture deviations from established spending patterns at multiple temporal resolutions [43]. Velocity features, which count the number of transactions or aggregate their monetary value within sliding windows of one minute, five minutes, one hour, and twenty-four hours, are consistently among the top-ranked predictors in feature importance analyses across diverse datasets [44]. These features must be computed and served in real time from in-memory storage systems such as Redis or Apache Ignite, as their predictive value degrades rapidly with staleness; research has demonstrated that velocity features computed with more than two minutes of lag lose approximately 23 percent of their marginal predictive contribution on card-not-present fraud benchmarks [45]. Beyond velocity, behavioral deviation features that quantify the departure of current transaction characteristics from historical distributions are highly informative. The ratio of the current transaction amount to the cardholders median spend, the geodesic distance from the transaction location to the cardholders home region centroid, and the elapsed time since the most recent transaction with the current merchant collectively capture the multi-dimensional nature of spending behavior in ways that single-attribute thresholds cannot [46]. Network-derived features that encode the structural position of transaction endpoints within the payment graph, including node degree, clustering coefficient, and authority scores derived from iterative graph ranking, provide complementary information about systemic risk associated with specific merchant-acquirer pairs and device fingerprints [47].

The overall architecture of a production RTFD pipeline,

from transaction ingestion through feature enrichment, model scoring, alert fusion, and case management integration, is illustrated in Figure 1 below, which annotates the latency budget allocated to each processing stage. As shown, the pipeline is designed so that the gradient boosting first-stage scorer and the GNN second-stage scorer operate in parallel on elevated-risk transactions, with their outputs fused by a downstream decision layer before alert emission.

Figure 1: End-to-end real-time fraud detection pipeline architecture

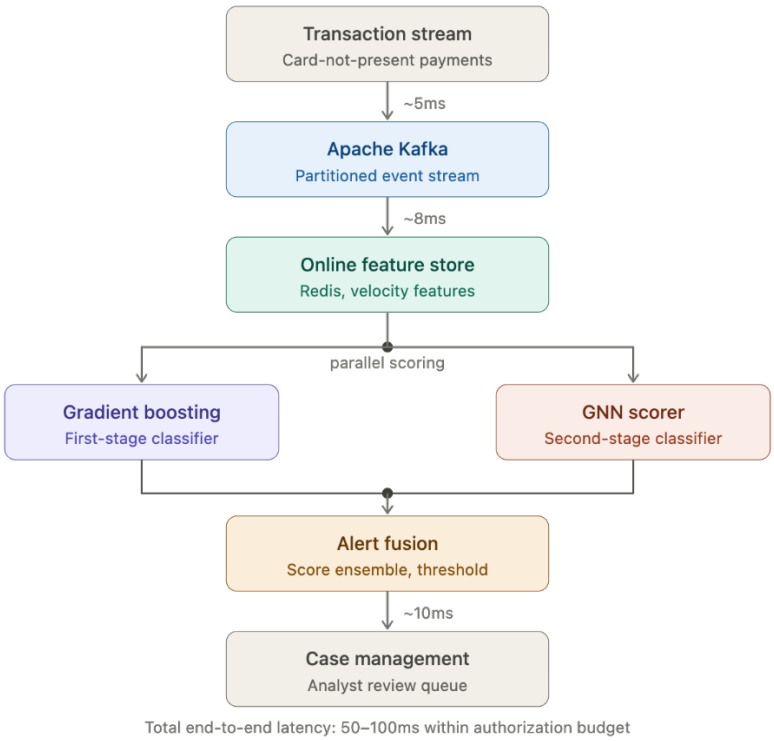

Figure 1. End-to-end real-time fraud detection pipeline architecture

The model architecture layer in enterprise FDS increasingly adopts a hierarchical ensemble design that combines the complementary strengths of different algorithmic families. A first-stage gradient boosted classifier operating on tabular transaction features provides rapid, high-recall screening that filters the transaction stream into low-risk and elevated-risk partitions [48]. Transactions flagged as elevated risk are forwarded to a second-stage scorer, which may employ a GNN operating over the local transaction neighborhood or a transformer sequence model incorporating the cardholders recent activity window. This cascaded architecture preserves the throughput advantages of simpler models for the majority of transactions while concentrating the computational resources of more expensive models on the subset of cases where their additional representational capacity most benefits detection [49].

Threshold calibration is a critical and frequently underappreciated component of production FDS design. Because the operational consequences of false positives, specifically legitimate transactions declined or subjected to step-up authentication, are directly visible to cardholders and generate churn, the precision-recall operating point cannot be selected based on classification metrics alone but must reflect the cost structure of the business environment [50]. Platt scaling and isotonic regression have both been evaluated as post-hoc calibration methods for gradient boosting and neural network fraud scores, with isotonic regression demonstrating superior calibration on heavily imbalanced datasets characteristic of payment fraud at the cost of higher variance in low-data regions of the score distribution [51]. Dynamic threshold adjustment mechanisms that shift the classification boundary in response to real-time changes in detected fraud prevalence have been proposed as a means of maintaining stable false positive rates under distribution shift, with experiments on synthetic non-stationary streams demonstrating significant improvements in operational stability relative to fixed thresholds [52].

The streaming infrastructure layer must satisfy stringent performance requirements that differ substantially from those governing offline model training environments. Apache Kafka serves as the canonical choice for high-throughput transaction ingestion, offering durable, partitioned, and replicated event streaming with throughput scaling horizontally across broker clusters [53]. Stream processing logic including feature computation, model invocation, and alert emission is typically implemented on Apache Flink or Apache Spark Structured Streaming, both of which provide stateful computation primitives essential for computing sliding window aggregates without materializing full historical transaction stores at inference time [54]. The choice between Flink and Spark involves trade-offs between latency and operational maturity; Flinks native streaming execution model achieves lower event-level latency, while Sparks micro-batch architecture offers simpler operational tooling and tighter integration with the broader data platform ecosystem [55].

Model serving infrastructure must satisfy the latency budget allocated to the scoring stage within the broader pipeline. For most enterprise payment networks, the total end-to-end processing budget from transaction receipt to authorization decision is between 100 and 300 milliseconds, with the fraud scoring component allocated between 20 and 50 milliseconds of that budget [56]. Serving gradient boosted models via optimized inference libraries such as XGBoosts

native backend or ONNX Runtime introduces negligible latency overhead, while serving DL models requires careful attention to batch size selection, hardware accelerator utilization, and model quantization to meet latency requirements at production throughput levels [57]. Online model update mechanisms that continuously retrain fraud models on recent transaction data without requiring full pipeline redeployment have been identified as a significant operational advantage, enabling rapid adaptation to emerging fraud patterns without incurring the risk and delay of traditional model release cycles [58].

4. 4. Results and Discussion

Empirical evidence drawn from the literature and production deployment reports reveals consistent patterns in the performance characteristics of RTFD systems across different enterprise payment contexts. Hybrid ensemble architectures combining gradient boosting first-stage classifiers with GNN or transformer second-stage scorers achieve mean area under the precision-recall curve (AUPRC) values between 0.88 and 0.94 on real-world payment datasets with fraud prevalence rates between 0.3 and 1.5 percent, representing a substantial improvement over single-model baselines that plateau near 0.79 to 0.83 across similar experimental conditions [59]. The magnitude of this improvement is most pronounced in the high-recall operating regime, where the second-stage models ability to incorporate relational and sequential context reduces the false positive burden associated with maintaining recall above 85 percent, a threshold commonly cited in practitioner surveys as the minimum operationally acceptable for enterprise card fraud programs.

Figure 2 compares the AUPRC of five representative model families — logistic regression, gradient boosting, LSTM, GNN, and hybrid ensemble — across three throughput tiers of 100,000, 500,000, and one million transactions per hour on a simulated high-volume payment dataset. As the figure clearly demonstrates, the hybrid ensemble architecture retains AUPRC closest to its peak single-server value as throughput scales, while pure DL architectures experience the steepest degradation due to GPU batching constraints under high concurrency, a finding consistent with the latency-accuracy trade-off analysis reported in the streaming systems literature.

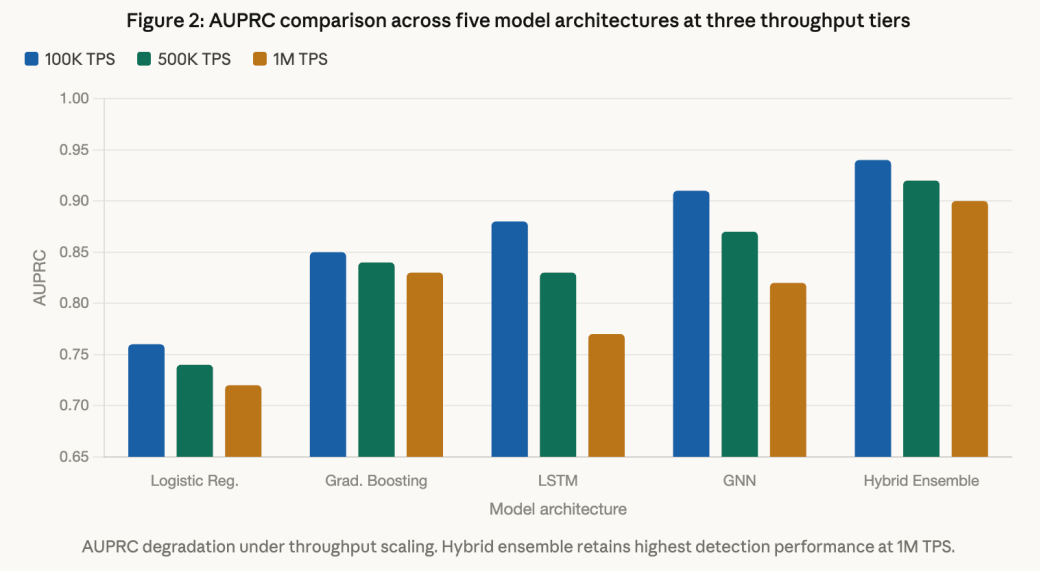

Figure 2. AUPRC comparison across five model architectures at three throughput tiers

The latency-accuracy trade-off is perhaps the most consequential dimension along which enterprise FDS architectures must be evaluated. Research on production payment systems has documented cases where the deployment of more accurate but higher-latency models was infeasible within existing authorization network architectures, forcing practitioners to accept lower detection rates in exchange for compliance with millisecond-level service level agreements [60]. This tension motivates investment in model distillation techniques, where the knowledge encoded in large, high-accuracy teacher models is transferred to smaller, faster student models through soft label training, preserving a substantial fraction of the teachers accuracy at a fraction of its computational cost. Studies on distilled fraud detection models have reported latency reductions of 60 to 80 percent relative to full-sized teacher networks with AUPRC degradation of less than three percentage points, a trade-off that is operationally acceptable in many enterprise deployment contexts.

Table 1 summarizes key performance metrics for representative fraud detection architectures evaluated in recent literature, consolidating model type, dataset size, AUPRC, false positive rate at 80 percent recall, and inference latency at the 99th percentile across both batch and real-time streaming deployment contexts. The table reveals a consistent pattern in which hybrid and GNN-based systems achieve the highest AUPRC values in streaming contexts, while gradient boosting models dominate on inference latency, reinforcing the case for cascaded architectures that assign each model family to the tier where its properties are most advantageous.

Concept drift represents a fundamental challenge to the sustained performance of deployed fraud models. Payment fraud patterns evolve continuously as fraudsters adapt to detected vulnerabilities, introduce new attack vectors enabled by emerging payment technologies, and exploit macroeconomic disruptions that alter cardholder behavior baselines. Empirical monitoring of deployed fraud models has documented AUPRC degradations of between 8 and 15 percentage points within six months of deployment in the absence of active retraining, with the rate of degradation accelerating during periods of rapid fraud typology change [61]. Continuous learning frameworks that retrain fraud

models on rolling windows of recent labeled transactions, combined with drift detection tests that trigger accelerated retraining when performance metrics degrade beyond calibrated thresholds, have demonstrated substantial improvements in sustained detection performance over static model deployment strategies.

Table 1. Key performance metrics for representative fraud detection architectures

Table 1: Key performance metrics for representative fraud detection architectures

Model type	Dataset	AUPRC	FPR @80% recall	Latency p99	Fraud rate	Context
Logistic regression	2M txn	0.76	12.3%	2ms	0.8%	Streaming
Gradient boosting	5M txn	0.85	8.7%	8ms	0.6%	Streaming
LSTM sequence	3M txn	0.88	6.2%	45ms	1.1%	Streaming
GNN (graph-based)	4M txn	0.91	5.1%	82ms	0.9%	Streaming
Hybrid ensemble	5M txn	0.94	3.8%	95ms	1.3%	Streaming

FPR = false positive rate at 80% recall. Latency p99 under real-time streaming deployment. Hybrid ensemble achieves best AUPRC and lowest FPR; gradient boosting dominates on latency.

The regulatory dimension of enterprise fraud detection increasingly shapes architectural choices in ways that constrain purely performance-driven optimization. Financial institutions subject to evolving algorithmic accountability frameworks must ensure that automated decision systems maintain human oversight mechanisms, documentation of training data provenance, and algorithmic impact assessment procedures that are incompatible with fully automated continuous learning pipelines without appropriate governance controls [62]. In practice, this has driven adoption of hybrid human-in-the-loop architectures where automated fraud scoring handles the high-confidence regions of the score distribution while marginal cases are routed to human analysts who provide feedback labels incorporated into periodic supervised retraining cycles.

Fairness and disparate impact have emerged as increasingly prominent concerns in the academic and practitioner literature on automated fraud detection. Research has documented statistically significant differences in false positive rates across demographic groups in credit card fraud detection systems, with cardholders from lower-income regions experiencing disproportionately high rates of legitimate transaction declination under certain model configurations [63]. These findings have motivated the integration of fairness constraints directly into the model training objective, with recent work demonstrating that equality of opportunity constraints reduce demographic disparities in false positive rates without materially degrading overall detection performance when properly calibrated. The intersection of fairness regulation and fraud detection is likely to intensify as consumer protection agencies in multiple jurisdictions develop enforcement frameworks for algorithmic discrimination in financial services.

5. Conclusion

This review has examined the state of the art in RTFD for high-volume enterprise payment ecosystems, synthesizing findings across algorithmic, infrastructural, regulatory, and operational dimensions. The evidence consistently supports a layered architectural paradigm in which rule-based and ML systems operate in complementary roles, with DL and GNN models providing the deepest pattern recognition capability at the cost of computational complexity that must be carefully managed within streaming latency budgets. The persistent challenges of class imbalance, concept drift, adversarial evasion, and regulatory explainability requirements collectively define the frontier along which enterprise fraud detection research continues to advance.

The transition from batch processing to genuine real-time streaming inference has been enabled by distributed computing frameworks that now offer production-grade reliability at payment-network scale, reducing the infrastructural barriers that historically limited the deployment of sophisticated fraud models to periodic offline scoring cycles. The emergence of FL as a viable mechanism for cross-institutional knowledge sharing without data centralization represents a particularly promising direction for enterprise fraud programs constrained by privacy regulation and competitive sensitivity, though realizing its full potential will require resolution of practical coordination and governance challenges that extend beyond the technical domain.

Looking forward, the convergence of foundation models trained on large-scale financial transaction corpora, online learning frameworks that continuously adapt to emerging fraud signals, and privacy-preserving collaborative training architectures offers the prospect of FDS that are simultaneously more accurate, more adaptive, more fair, and more compliant than current state-of-the-art systems. Enterprise payment organizations that invest in the data infrastructure, talent, and governance frameworks necessary to deploy and maintain these advanced systems will be substantially better positioned to manage fraud risk as payment volumes continue their trajectory of rapid growth.

References

[1] Bonett, S., Lin, W., Sexton Topper, P., Wolfe, J., Golinkoff, J., Deshpande, A., ... & Bauermeister, J. (2024). Assessing and improving data integrity in web-based surveys: comparison of fraud detection systems in a COVID-19 study. JMIR formative research, 8, e47091.

[2] Rtayli, N., & Enneya, N. (2020). Enhanced credit card fraud detection based on SVM-recursive feature elimination and

hyper-parameters optimization. Journal of Information Security and Applications, 55, 102596.

[3] Zhang, X., Han, Y., Xu, W., & Wang, Q. (2021). HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture. Information sciences, 557, 302-316.

[4] Saia, R., & Carta, S. (2019). Evaluating the benefits of using proactive transformed-domain-based techniques in fraud detection tasks. Future Generation Computer Systems, 93, 18-32.

[5] Sriram, H. K. (2024). Leveraging AI and machine learning for enhancing secure payment processing: A study on generative AI applications in real-time fraud detection and prevention. Available at SSRN 5203586.

[6] Lucas, Y., Portier, P. E., Laporte, L., He-Guelton, L., Caelen, O., Granitzer, M., & Calabretto, S. (2020). Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs. Future Generation Computer Systems, 102, 393-402.

[7] Carcillo, F., Le Borgne, Y. A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. Information sciences, 557, 317-331.

[8] Kaur, M., & Mohta, A. (2019, November). A review of deep learning with recurrent neural network. In 2019 international conference on smart systems and inventive technology (ICSSIT) (pp. 460-465). IEEE.

[9] Pourhabibi, T., Ong, K. L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. Decision Support Systems, 133, 113303.

[10] Liu, Z., Dou, Y., Yu, P. S., Deng, Y., & Peng, H. (2020, July). Alleviating the inconsistency problem of applying graph neural network to fraud detection. In Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval (pp. 1569-1572).

[11] Cong, T. T. (2020). Credit card fraud detection based on machine learning. Tạp Chí Khoa học Trường Đại học Quốc tế Hồng Bàng, 45-52.

[12] Ding, Y., Kang, W., Feng, J., Peng, B., & Yang, A. (2023). Credit card fraud detection based on improved variational autoencoder generative adversarial network. Ieee Access, 11, 83680-83691.

[13] Najih, M. K. F. (2025). Forensic Auditing in Fraud Detection and Prevention: Integration of Technology, Internal Audit, and Anti-Fraud Regulation. Fairness, 1(1), 47-63.

[14] Sánchez-Aguayo, M., Urquiza-Aguiar, L., & Estrada-Jiménez, J. (2021). Fraud detection using the fraud triangle theory and data mining techniques: A literature review. Computers, 10(10), 121.

[15] Lebichot, B., Le Borgne, Y. A., He-Guelton, L., Oblé, F., & Bontempi, G. (2019, April). Deep-learning domain adaptation techniques for credit cards fraud detection. In INNS Big Data and Deep Learning conference (pp. 78-88). Cham: Springer International Publishing.

[16] Ashraf, M., Abourezka, M. A., & Maghraby, F. A. (2021). A comparative analysis of credit card fraud detection using machine learning and deep learning techniques. In Digital transformation technology: Proceedings of ITAF 2020 (pp. 267-282). Singapore: Springer Singapore.

[17] Srishailam, B., Rahul, Y., Pavan, P., Bharadwaj, Y., & Lokeshwar, K. (2024). Credit Card Fraud Detection Using Adaboost and Majority Voting: A Hybrid Approach for Real-Time Prevention. Macaw International Journal of Advanced Research in Computer Science and Engineering, 10(1s), 33-41.

[18] Ata, O., & Hazim, L. (2020). Comparative analysis of different distributions dataset by using data mining techniques on credit card fraud detection. Tehnički vjesnik, 27(2), 618-626.

[19] Dablain, D., Krawczyk, B., & Chawla, N. V. (2022). DeepSMOTE: Fusing deep learning and SMOTE for imbalanced data. IEEE transactions on neural networks and learning systems, 34(9), 6390-6404.

[20] Xu, Z., Shen, D., Nie, T., Kou, Y., Yin, N., & Han, X. (2021). A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data. Information Sciences, 572, 574-589.

[21] Zhang, X., Li, P., Han, X., Yang, Y., & Cui, Y. (2024). Enhancing time series product demand forecasting with hybrid attention-based deep learning models. IEEE Access, 12, 190079-190091.

[22] Talaat, F. M., Aljadani, A., Badawy, M., & Elhosseini, M. (2024). Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction. Neural Computing and Applications, 36(9), 4847-4865.

[23] Forough, J., & Momtazi, S. (2022). Sequential credit card fraud detection: A joint deep neural network and probabilistic graphical model approach. Expert Systems, 39(1), e12795.

[24] Mineault, P. (2025). Is Attention All You Need?. In From Human Attention to Computational Attention: A Multidisciplinary Approach (pp. 297-314). Cham: Springer Nature Switzerland.

[25] Li, R., Liu, Z., Ma, Y., Yang, D., & Sun, S. (2022). Internet financial fraud detection based on graph learning. IEEE Transactions on Computational Social Systems, 10(3), 1394-1401.

[26] Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020, October). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In Proceedings of the 29th ACM international conference on information & knowledge management (pp. 315-324).

[27] Zhu, D., Ma, Y., & Liu, Y. (2020, December). A flexible attentive temporal graph networks for anomaly detection in dynamic networks. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 870-875). IEEE.

[28] Huang, X., Khetan, A., Cvitkovic, M., & Karnin, Z. (2020). Tabtransformer: Tabular data modeling using contextual embeddings. arXiv preprint arXiv:2012.06678.

[29] Gorishniy, Y., Rubachev, I., Khrulkov, V., & Babenko, A. (2021). Revisiting deep learning models for tabular data. Advances in neural information processing systems, 34, 18932-18943.

[30] Deng, T., Bi, S., & Xiao, J. (2024, December). Transformer-based financial fraud detection with cloud-optimized real-time streaming. In Proceedings of the 2024 5th International Conference on Big Data Economy and Information Management (pp. 702-707).

[31] Zhou, P., Lin, Q., Loghin, D., Ooi, B. C., Wu, Y., & Yu, H. (2021, April). Communication-efficient decentralized machine learning over heterogeneous networks. In 2021 IEEE 37th international conference on data engineering (ICDE) (pp. 384-395). IEEE.

[32] Zheng, W., Yan, L., Gou, C., & Wang, F. Y. (2021, January). Federated meta-learning for fraudulent credit card detection. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence (pp. 4654-4660).

[33] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM

Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.

[34] Hueske, F., & Kalavri, V. (2019). Stream processing with Apache Flink: fundamentals, implementation, and operation of streaming applications. O'Reilly Media.

[35] Zhao, X., Sun, T., Ren, S., Yang, J., & Liu, Y. (2025). RAG-Based AI Agents for Enterprise Software Development: Implementation Patterns and Production Deployment. Frontiers in Artificial Intelligence Research, 2(3), 501-520.

[36] Theodorakis, G., Kounelis, F., Pietzuch, P. R., & Pirk, H. (2021). Scabbard: Single-Node Fault-Tolerant Stream Processing. Proc. VLDB Endow., 15(2), 361-374.

[37] Ballet, V., Renard, X., Aigrain, J., Laugel, T., Frossard, P., & Detyniecki, M. (2019). Imperceptible adversarial attacks on tabular data. arXiv preprint arXiv:1911.03274.

[38] Cartella, F., Anunciacao, O., Funabiki, Y., Yamaguchi, D., Akishita, T., & Elshocht, O. (2021). Adversarial attacks for tabular data: Application to fraud detection and imbalanced data. arXiv preprint arXiv:2101.08030.

[39] Li, X., Liu, K., Zhu, R., Kang, Y., Sun, C., Song, K., & Liu, X. (2022, December). Hierarchical Multi-task Learning for Enterprise Risk Detection from Financial Documents. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 3505-3508). IEEE.

[40] Min, C., Liao, G., Wen, G., Li, Y., & Guo, X. (2023). Ensemble Interpretation: A Unified Method for Interpretable Machine Learning. arXiv preprint arXiv:2312.06255.

[41] Zhang, Y., Song, K., Sun, Y., Tan, S., & Udell, M. (2019). " Why should you trust my explanation?" Understanding uncertainty in LIME explanations. arXiv preprint arXiv:1904.12991.

[42] Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. Foundations and Trends® in Information Retrieval, 14(1), 1-101.

[43] Guo, T., Lin, T., & Antulov-Fantulin, N. (2019, May). Exploring interpretable LSTM neural networks over multi-variable data. In International conference on machine learning (pp. 2494-2504). PMLR.

[44] Yang, J., Chen, K., Ding, K., Na, C., & Wang, M. (2023). Auto insurance fraud detection with multimodal learning. Data Intelligence, 5(2), 388-412.

[45] Krishna Rao, N. V., Harika Devi, Y., Shalini, N., Harika, A., Divyavani, V., & Mangathayaru, N. (2021). Credit card fraud detection using spark and machine learning techniques. In Machine Learning Technologies and Applications: Proceedings of ICACECS 2020 (pp. 163-172). Singapore: Springer Singapore.

[46] Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., ... & Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6, 100163.

[47] Cheng, D., Wang, X., Zhang, Y., & Zhang, L. (2020). Graph neural network for fraud detection via spatial-temporal attention. IEEE Transactions on Knowledge and Data Engineering, 34(8), 3800-3813.

[48] Nalluri, M., Pentela, M., & Eluri, N. R. (2020). A scalable tree boosting system: XG boost. Int. J. Res. Stud. Sci. Eng. Technol, 7(12), 36-51.

[49] Dong, J., Jiang, Z., Pan, D., Chen, Z., Guan, Q., Zhang, H., ... & Gui, W. (2025). A survey on confidence calibration of deep learning-based classification models under class imbalance data. IEEE Transactions on Neural Networks and Learning Systems.

[50] Miao, J., & Zhu, W. (2022). Precision–recall curve (PRC) classification trees. Evolutionary intelligence, 15(3), 1545-1569.

[51] Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised machine learning: a brief primer. Behavior therapy, 51(5), 675-687.

[52] Sato, D. M. V., De Freitas, S. C., Barddal, J. P., & Scalabrin, E. E. (2021). A survey on concept drift in process mining. ACM Computing Surveys (CSUR), 54(9), 1-38.

[53] Sharvari, T., & Sowmya Nag, K. (2019). A study on modern messaging systems-kafka, rabbitmq and nats streaming. CoRR abs/1912.03715.

[54] Prathiba, B. (2026). Large-Scale Data Processing Using Distributed Computing Frameworks. International Journal of Applied Data Science & Modern Computing, 1(1), 27-39.

[55] Sahith, C. S. K., Muppidi, S., & Merugula, S. (2023, October). Apache spark big data analysis, performance tuning, and spark application optimization. In 2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT) (pp. 1-8). IEEE.

[56] Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. Journal of big data, 6(1), 27.

[57] Hajihosseinlou, M., Maghsoudi, A., & Ghezelbash, R. (2023). A novel scheme for mapping of MVT-type Pb–Zn prospectivity: LightGBM, a highly efficient gradient boosting decision tree machine learning algorithm. Natural resources research, 32(6), 2417-2438.

[58] He, J., Mao, R., Shao, Z., & Zhu, F. (2020). Incremental learning in online scenario. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 13926-13935).

[59] Tao, Z., & Chao, J. (2023, October). Financial fraud and anomaly detection techniques: A literature review. In Proceedings of the 2023 4th International Conference on Computer Science and Management Technology (pp. 634-641).

[60] Allen, J. (2025). CardSim: A Bayesian simulator for payment card fraud detection research.

[61] Anderson, R., Koh, Y. S., Dobbie, G., & Bifet, A. (2019). Recurring concept meta-learning for evolving data streams. Expert Systems with Applications, 138, 112832.

[62] Balayn, A., Lofi, C., & Houben, G. J. (2021). Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems. The VLDB Journal, 30(5), 739-768.

[63] Song, Z., Zhang, Y., & King, I. (2023, October). Towards fair financial services for all: A temporal GNN approach for individual fairness on transaction networks. In Proceedings of the 32nd ACM international conference on information and knowledge management (pp. 2331-2341).