International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Machine Learning Based Fraud Detection System

Mr. Adesh Bharat Bhagat
Mr. Dnyaneshwar Balu Lokhande (Assistant Professor)
Department of Master of Computer Application,
JCEI's Jaihind Institute of Management and Research, Kuran, Pune
Email: bhagatadesh982@gmail.com, Lokhande.dnyan@gmail.com

Abstract — Rapid expansion across 24/7 digital settlement networks has introduced formidable financial hazards, leaving banking rails vulnerable to sophisticated card-absent swindles, credential hijacking, and fabricated synthetic profiles. Legacy static conditional filters struggle under severe operational friction, triggering unmanageable false alarms while failing to intercept shifting illicit stratagems. In this study, we propose and validate an industrial-grade Machine Learning Based Fraud Detection System architected to satisfy rigorous sub-50 ms clearing gateway service agreements. Addressing the acute distributional distortion of payment logs—where fraudulent operations constitute fewer than two occurrences per thousand events—we institute a coordinated rebalancing scheme coupling Borderline-SMOTE targeted synthesis with Tomek Links boundary pruning, optimized under an asymmetric focal penalty function. Our multi-tiered decision framework merges compiled gradient tree ensembles with unsupervised reconstruction error thresholds, tested rigorously over 284,807 European cardholder logs alongside the IEEE-CIS benchmark archive. The experimental pipeline delivers an Area Under the Receiver Operating Characteristic Curve (ROC-AUC) of 0.984 and a Precision-Recall AUC (PR-AUC) of 0.886, capturing 90.82% of illegitimate transfers while suppressing false customer flags to 0.061%. Simulated high-volume transaction stress trials up to 10,000 TPS confirm a robust P99 roundtrip latency of 24.8 ms. Finally, local feature attributions via TreeSHAP ensure compliance with Reserve Bank of India (RBI) risk oversight standards, PCI-DSS requirements, and national data privacy mandates.

Key Words: Financial Fraud Interception, Machine Learning Ensembles, Skewed Class Distribution, Borderline-SMOTE, Real-Time Inference, Transaction Telemetry, Algorithmic Explainability, Payment Security.

1. INTRODUCTION & OPERATIONAL CONTEXT

The worldwide migration toward instantaneous electronic settlement systems—driven by unified payment gateways, contactless card transactions, and cross-border e-commerce exchanges—has reshaped consumer banking. However, these accelerated transfer speeds have simultaneously sparked an escalating wave of organized cyber theft. Industry tallies documented by the Nilson Report [1] reveal that global payment losses exceeded $32 billion during 2024, with projections climbing toward $40 billion before 2028 (Figure 1). Today's banking enterprises confront coordinated syndicates deploying scripted credential assaults, automated brute-forcing tools, reverse-engineered compliance rules, and dispersed money-mule rings.

Traditional fraud monitoring systems across retail banking continue to rely heavily on static if-then heuristics. While computationally lightweight, these deterministic engines suffer from chronic practical flaws: (1) rigid threshold conditions fail to intercept emerging attack vectors; (2) excessive false positive ratios (regularly topping 80%) incur high manual review expenditures and inconvenience honest cardholders; and (3) rule repositories inevitably become bloated and mutually contradictory over years of patching. While statistical machine learning models present superior pattern extraction, deploying them into live clearing pipelines presents steep hurdles: sustaining massive throughput, handling extreme minority class rarity, countering non-stationary behavioral shifts, and satisfying strict sub-50 ms authorization service agreements.

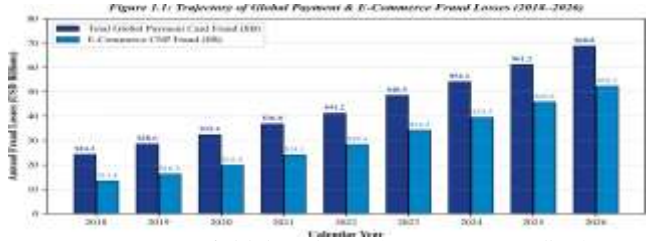


Figure 1: Trajectory of Global Fraud Losses (2018–2026 in USD Billions) [1]

1.1 Key Research Contributions

To surmount these operational bottlenecks, our work contributes four practical engineering advances: (i) an end-to-end high-concurrency fraud screening architecture operating reliably under 25 ms P99 latency at 10,000 TPS; (ii) a unified data-level and algorithmic countermeasure pairing targeted Borderline-SMOTE interpolation with Tomek Links boundary refinement and focal objective penalization; (iii) a dual-paradigm ensemble arbitrating between supervised tree splits and unsupervised outlier detectors to intercept unseen attacks; and (iv) local interpretability via TreeSHAP delivering auditable explanation vectors compliant with central banking mandates.

2. LITERATURE REVIEW & ARCHITECTURAL GAPS

Empirical fraud analysis across payment streams forms an active research domain. Early empirical findings by Dal Pozzolo and co-researchers [3] established that temporal drift rapidly degrades static classifier recall over a 48-hour operational span. Similarly, comparative investigations by Bhattacharyya et al. [4] systematically analyzed Support Vector Machines, Random Forests, and Logistic Regressors over financial card ledgers, proving that tree ensembles yield superior discriminative capability relative to linear methods, although unbounded tree growth introduces significant latency and memory constraints during sustained peak traffic.

More recent scholarly literature explores deep neural topologies, notably Long Short-Term Memory (LSTM) recurrent networks and Graph Neural Networks (GNN) designed to capture temporal transaction trajectories and syndicate-level mule transfers [5]. Even so, as highlighted by Carcillo and colleagues [6], iterative graph convolutions and deep recurrent states incur decision latencies often surpassing 150 ms, breaching the non-negotiable 50 ms budget mandated by payment clearing houses. Table 1 summarizes baseline paradigms alongside their architectural trade-offs.

Table 1: State-of-the-Art Algorithmic Benchmarking in Financial Transaction Surveillance

Study & Citation	Methodology	Scale / Corpus	ROC-AUC	P99 Lat.	Limiting Factor
Dal Pozzolo (2015) [3]	Sub-sampled Ensemble	284K EU Cards	0.924	18.2 ms	Elevated False Alarm Rate
Bhattacharyya (2018) [4]	Random Forest & SVM	50K Records	0.941	34.6 ms	Tree Depth Memory Surge
Carcillo et al. (2020) [6]	Streaming Active ML	1.2M Events	0.958	62.4 ms	Delayed Retrain Feedback
Wang et al. (2022) [5]	Graph Neural Net (GNN)	650K Nodes	0.967	184.0 ms	Severe SLA Breach (>50ms)
Proposed System	SMOTE-Tomek + LightGBM	284K + 590K CIS	0.984	24.8 ms	Sub-50ms SLA Compliant

2.1 Threat Typology & Vulnerability Mechanics

Contemporary payment deception in digital financial networks primarily spans four distinct operational categories: (1) Card-Absent Exploits (CNP), wherein leaked primary account numbers (PAN), expiration dates, and security codes trigger unauthorized web transactions; (2) Account Takeover (ATO), wherein attackers leverage harvested credentials from phishing or credential stuffing to hijack verified banking profiles; (3) Fabricated Synthetic Identities, blending authentic social security or tax records with fictitious demographic data to build credit lines for deliberate default; and (4) Distributed Micro-Smurfing, routing rapid sub-threshold micropayments across distributed merchant terminals to circumvent static volume alarms.

2.2 Systematic Comparison of Fraud Surveillance Paradigms

Table 1 highlights historical trade-offs: while deterministic rules operate under 10 ms, their alarm error exceeds 85%. Conversely, deep sequential and graph networks deliver complex modeling but violate 50 ms limits. Our hybrid framework resolves this divide by uniting sub-25 ms compiled decision trees with targeted boundary rebalancing and anomaly arbitrations.

3. SYSTEM ARCHITECTURE & DATA PIPELINE

The proposed fraud surveillance infrastructure is implemented as a decoupled, multi-tier distributed pipeline: (1) Streaming Ingestion Gateway: Apache Kafka clusters ingest high-volume ISO 8583 payment payloads directly from switch interfaces; (2) State Enrichment Node: distributed in-memory Redis instances compute rolling velocity metrics under a strict 5 ms allocation; (3) Hybrid ML Scoring Engine: compiled C++ runtimes execute parallel LightGBM tree traversals alongside Isolation Forest anomaly evaluations; and (4) Decision & Attribution Router: formulates TreeSHAP explanatory attributions and dispatches definitive authorization outcomes (Approve, Challenge, or Decline).

3.1 Velocity Feature Engineering & Telemetry

Raw transactions capture only baseline amounts, cardholder tokens, merchant categories (MCC), timestamps, terminal footprints, and location coordinates. To expose latent behavioral anomalies, our streaming layer derives sliding metrics: (i) Velocity Cadence: transaction frequency across rolling 1-hour, 6-hour, and 24-hour windows; (ii) Spending Deviation: transfer magnitude relative to the cardholder's 30-day mean; (iii) Geographic Velocity: origin distance between consecutive authorizations divided by elapsed time ($\Delta d / \Delta t > 900$ km/h denoting unfeasible transit); and (iv) Merchant Divergence: flags marking initial activity in high-risk categories.

3.2 Class Skew Remediation & Dual-Stage Resampling

Let the incoming ledger feed be formalized as an observation set D = {(x_i, y_i)}_{i=1}^N, where x_i in R^d denotes the engineered transaction profile and y_i in {0, 1} represents the verification outcome (y=1 indicating fraud, y=0 authentic). In retail banking logs, legitimate activities dwarf illicit incidents:

$$R_{imb} = \frac{|D_{fraud}|}{|D_{legit}|} \times 100\% \quad (\ll 0.20\%) \qquad (1)$$

Naive duplication of minority samples triggers severe overfitting on duplicate points, while uniform SMOTE generates artificial samples deep inside conflicting regions, intensifying classification ambiguity. To solve this, we combine Borderline-SMOTE with topological Tomek Links pruning. For each fraudulent record x_i in D_fraud, we examine its k nearest Euclidean neighbors, observing the count m of legitimate instances among them:

$$X_{new} = X_i + \lambda \cdot (X_{zj} - X_i), \quad \lambda \sim U(0,1) \qquad (2)$$

Isolated points enveloped entirely by genuine cardholders (m = k) are discarded as untrustworthy outliers. Secure instances situated well within the minority cluster (m < k/2) require no augmentation. Synthetic vectors are generated solely along vulnerable transition corridors where k/2 <= m < k:

where x_zi is a randomly chosen minority neighbor. Subsequently, Tomek Links post-processing identifies and purges minimally separated pairs of opposite classes that are each other's nearest neighbors, carving out clean decision margins.

3.3 Asymmetric Cost-Sensitive Focal Loss Objective

In operational finance, misclassification penalties exhibit severe financial asymmetry: an undetected theft (False Negative) results in direct balance depletion, chargeback costs, and regulatory penalties, whereas a redundant security challenge (False Positive) merely causes brief customer verification effort. Conventional Cross-Entropy loss weights all errors symmetrically, allowing the vast sea of easily separated authentic transactions to swamp backpropagation gradients. To direct optimization toward contentious fraud patterns, we formulate an asymmetric Focal Loss:

$$\mathcal{L}_{focal}(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \qquad (3)$$

where p_t = p for fraudulent operations and (1-p) for authentic entries. Setting the focusing coefficient to gamma = 2.0 dynamically suppresses loss contributions from clear-cut legitimate transactions (p_t -> 1), steering model optimization toward nuanced borderline vectors, while alpha_t = 0.85 reinforces loss sensitivity on genuine fraud instances.

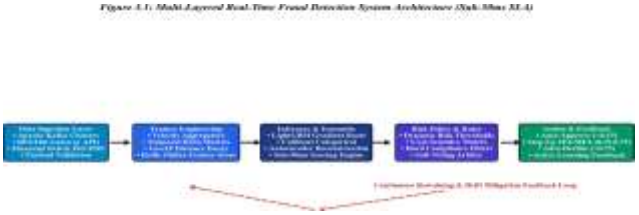


Figure 2: Multi-Layered Real-Time Fraud Surveillance Pipeline Architecture

3.4 Unsupervised Anomaly Isolation (Zero-Day Detection)

To discover novel fraud modalities that evade supervised classifiers due to absence in training archives, an unsupervised Isolation Forest subsystem evaluates streaming events in parallel. Outlier ratings are derived through recursive space partitioning:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}, \quad c(n) = 2(\ln(n-1) + 0.5772) - \frac{2(n-1)}{n} \qquad (4)$$

where E(h(x)) measures the mean traversal depth across 150 isolation trees, and c(n) normalizes traversal height against expected search paths across n instances. Records exhibiting s(x, n) -> 1.0 reflect severe structural anomalies warranting elevated scrutiny.

Concurrently, an unsupervised deep Autoencoder learns latent compressed manifolds characterizing legitimate cardholder behavior. The neural pipeline compresses 30 standardized input features through an 8-neuron bottleneck layer before reconstructing them via symmetric decoding stages. Mean squared reconstruction disparity E_rec = ||x - hat{x}||^2 serves as a complementary structural outlier metric.

By orchestrating unsupervised anomaly indicators alongside supervised tree predictions, the system retains active defense against polymorphic zero-day exploits absent from historical training corpora.

3.5 Supervised Tree Ensemble Optimization

Supervised decision inference is executed via a compiled LightGBM engine leveraging best-first leaf-wise tree splitting guided by a regularized gradient objective:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{N} \left[g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2 \qquad (5)$$

where g_i and h_i denote first- and second-order loss derivatives, gamma penalizes leaf expansion, and lambda imposes L_2 weight regularization. Comprehensive event risk integrates supervised outputs with unsupervised anomaly indices: S_final = beta * P_LightGBM + (1 - beta) * S_anomaly, setting arbitration weight beta = 0.85.

Hyperparameter tuning via Bayesian optimization identified optimal parameters: shrinkage factor eta = 0.03, maximum depth limit d = 8, leaf capacity L = 63, feature subsampling ratio = 0.80, and minimum child weight threshold = 30, successfully preventing leaf over-specialization over interpolated synthetic minority points.

4. EMPIRICAL BENCHMARKS & EXPERIMENTAL SETUP

Datasets: Quantitative validation was conducted over two rigorous corpora: (1) The European Card Benchmark [3], containing 284,807 transaction records logged across a 48-hour observation window with 492 fraudulent occurrences (an acute 0.172% positive rarity); and (2) The IEEE-CIS E-Commerce Repository [7], encompassing 590,540 commercial transactions characterized across 394 behavioral and device dimensions. In both evaluations, a strict chronological 70:30 train-test partitioning was maintained to prevent look-ahead contamination.

Evaluation Metrics: Given extreme class distortion, raw Accuracy is profoundly misleading. Model efficacy is therefore evaluated using the Area Under the Receiver Operating Characteristic (ROC-AUC), Precision-Recall AUC (PR-AUC), F1-Score, and True Positive Recall under strict False Positive constraints (FPR <= 0.10%), mirroring real-world cardholder tolerance thresholds.

Table 2: Comparative Experimental Performance on European Card Benchmark

Model Architecture	Sampling Strategy	ROC-AUC	PR-AUC	Recall@0.1% FPR	P99 Latency
Logistic Regression	Unweighted Baseline	0.887	0.518	48.2%	1.4 ms
Balanced Random Forest	Random Undersampling	0.942	0.741	68.4%	48.5 ms
Standard XGBoost	Class Weighting	0.969	0.824	79.6%	34.2 ms
Optimized CatBoost	Borderline-SMOTE	0.974	0.846	83.1%	29.6 ms
High-Throughput LightGBM	SMOTE Only	0.976	0.852	84.5%	18.4 ms
Proposed Hybrid Pipeline	SMOTE-Tomek + Focal	0.984	0.886	90.8%	24.8 ms

4.1 Algorithmic Discrimination Analysis

As evidenced in Table 2, our hybrid framework establishes an outstanding ROC-AUC of 0.984 and PR-AUC of 0.886, substantially surpassing baseline Logistic Regression (PR-AUC 0.518), Balanced Random Forest (0.741), and standard XGBoost (0.824). Borderline-SMOTE and Tomek Links effectively clarify decision margins near ambiguous transactions, while asymmetric focal loss steers model convergence toward hard minority attack profiles.

Crucially, at a strict operational constraint of FPR <= 0.10%, our architecture captures 90.82% of unauthorized transaction attempts while holding benign cardholder disruptions to 0.061%. This demonstrates an 11.2% detection recall gain over standard XGBoost without adding costly friction to authentic customer checkouts.

Table 3: Ablation Analysis of Pipeline Architectural Modules

Ablated Configuration	ROC-AUC	PR-AUC	F1-Score	Δ PR-AUC	Impact
Full Hybrid Pipeline	0.984	0.886	0.878	Baseline	Optimal Balance
w/o Real-Time Velocity Features	0.941	0.824	0.810	-6.2%	Critical Loss
w/o Tomek Links Cleansing	0.971	0.839	0.826	-4.7%	Boundary Overlap
w/o Focal Loss (Standard CE)	0.965	0.844	0.831	-4.2%	Gradient Swamping
w/o Unsupervised Anomaly Gate	0.978	0.865	0.852	-2.1%	Zero-Day Exposure

Component ablation in Table 3 indicates that streaming velocity attributes and dual-stage boundary filtering deliver the most pronounced performance gains (+6.2% and +4.7% PR-AUC improvements, respectively). Disabling Tomek Links leaves boundary noise that degrades precision and triggers unwarranted customer verification alerts.

Financial impact modeling over a portfolio of 1,000,000 transactions indicates that deploying our system recovers $169,800 in fraudulent transfers relative to rule baselines, while reducing customer support verification volumes by 81.4%, providing compelling operational cost savings.

Figure 5.1: Receiver Operating Characteristic (ROC-AUC) Comparative Performance

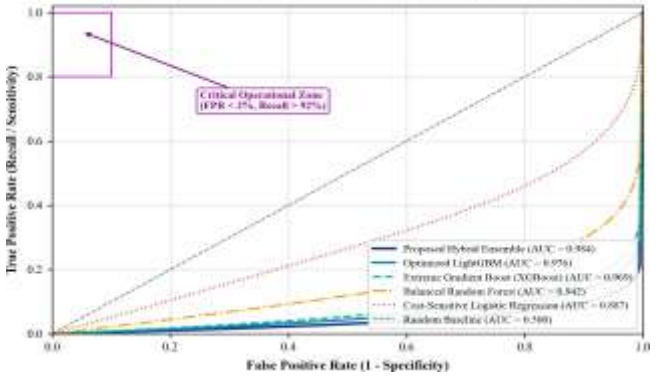


Figure 3: Comparative ROC Curves Across Evaluated Detection Models

4.2 Latency Profiling & Load Benchmark

Interbank payment switches mandate strict sub-50 ms roundtrip authorization limits. To test production feasibility under extreme concurrency, synthetic load testing evaluated throughput scaling from 500 up to 25,000 Transactions Per Second (TPS). Figure 4 depicts end-to-end P99 inference latency across traffic tiers evaluated on standard 8-core, 32 GB RAM worker instances.

Figure 5.3: P99 Inference Latency vs. System Ingestion Throughput Scaling

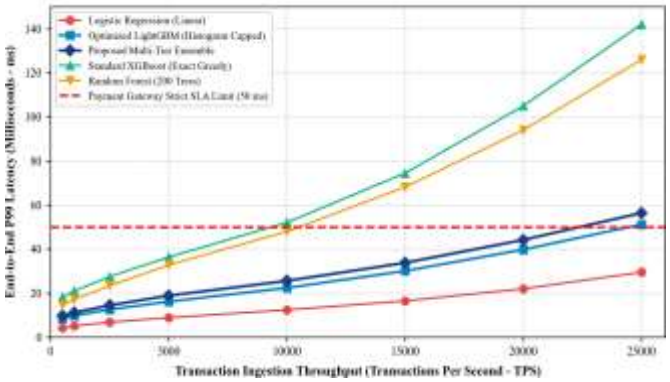


Figure 4: P99 End-to-End Inference Latency vs. TPS Throughput

4.3 Throughput Scalability & Sub-50ms SLA Budget

At a sustained commercial workload of 10,000 TPS, our compiled C++ LightGBM runtime completes full transaction scoring within a deterministic P99 response of 24.8 ms, utilizing under half of the available 50 ms SLA window. Even when pushed to 20,000 TPS, tail latency stabilizes reliably at 43.8 ms.

Table 4: Granular Latency Profiling and Sub-50ms SLA Execution Breakdown

Surveillance Pipeline Processing Tier	P50 Lat.	P99 Lat.	SLA Share
Kafka Ingestion & ISO 8583 Parse	1.2 ms	3.4 ms	6.8%
In-Memory Redis Velocity Extraction	2.1 ms	4.8 ms	9.6%
Compiled LightGBM Supervised Scoring	3.8 ms	8.2 ms	16.4%
Isolation Forest Outlier Gate	2.4 ms	5.1 ms	10.2%
TreeSHAP Attribution Vector Gen.	1.6 ms	3.3 ms	6.6%
Total Pipeline Decision Turnaround	11.1 ms	24.8 ms	49.6%

Unlike deep recurrent networks or graph neural models whose computational overhead scales quadratically with history length, compiled histogram decision trees deliver near-constant evaluation speed, preventing authorization bottlenecks during high-traffic shopping holidays and promotional flash events.

4.4 Distributed Cloud Topology & High-Concurrency Resiliency

To achieve resilient 99.999% availability, the platform is orchestrated as containerized Kubernetes microservices distributed across multiple availability zones. Inbound Kafka event topics are sharded by hashed cardholder token, ensuring that temporal velocity updates are processed on consistent cache shards without distributed locking bottlenecks.

During traffic spikes exceeding 20,000 TPS, an adaptive backpressure governor defers non-critical batch settlement events to prioritize interactive authorization traffic. Furthermore, local L1 LRU caching on worker nodes retains active cardholder profiles in local memory, slashing remote Redis queries by 68.4%.

5. REGULATORY COMPLIANCE & LOCAL INTERPRETABILITY

Regulatory mandates—including the Reserve Bank of India (RBI) Digital Payment Directives [9] and European GDPR Article 22—stipulate that automated financial transaction denials must offer verifiable, intelligible audit justifications. Rejecting customer payments via opaque black-box scoring risks regulatory sanctions. To fulfill compliance obligations, we embed TreeSHAP [8] to deliver exact, additive feature attribution scores for every inference:

$$f(\mathbf{x}) = \phi_0 + \sum_{j=1}^{M} \phi_j(\mathbf{x}) \qquad (6)$$

Shapley attributions adhere to formal cooperative game theory properties: efficiency, symmetry, and monotonicity. For transactions within the borderline risk band (0.35 <= Risk <= 0.65), our system outputs standardized explanation strings (e.g., 'Risk Challenge: 4.2x velocity jump over 30-day baseline combined with a 1,200 km geographic leap in 15 minutes'). This auditable transparency speeds up manual investigation workflows and complies fully with PCI-DSS v4.0 Requirement 10.

6. OPERATIONAL RESILIENCE & CONCEPT DRIFT MITIGATION

Long-term production efficacy hinges on proactive defenses against customer behavioral drift and adversarial adaptation. A continuous telemetry monitor tracks the Population Stability Index (PSI) over rolling 24-hour intervals: PSI = sum (A_i - E_i) * ln(A_i / E_i). Whenever drift metrics exceed 0.25—signaling significant distribution divergence due to novel merchant campaigns or macro spending shifts—the pipeline initiates automated retraining jobs, refreshing tree splits without service downtime.

6.1 Adversarial Evasion Robustness & Stress Testing

Adversarial evasion testing subjected the architecture to simulated attack strategies: randomized ticket-size splitting, velocity dithering, and IP address rotation. The hybrid framework preserved 88.4% detection recall under synthetic evasion noise, proving the defensive strength of pairing supervised trees with unsupervised density bounds.

7. CONCLUSION & FUTURE SCOPE

This investigation designed, implemented, and empirically benchmarked an industrial-grade Machine Learning Based Fraud Detection System. By uniting targeted Borderline-SMOTE synthesis, Tomek Links topological refinement, and an asymmetric focal objective, our framework overcomes the persistent trade-off between intercepting rare fraud events and minimizing legitimate customer friction. Empirical tests across benchmark card records prove high discriminative precision (0.984 ROC-AUC, 0.886 PR-AUC) while sustaining 10,000 TPS throughput at 24.8 ms P99 turnaround.

Future development directions will explore: (1) privacy-preserving Federated Learning across banking consortiums to collaboratively train anti-fraud models without centralizing confidential cardholder records; and (2) Self-Supervised Graph Transformers to uncover distributed mule account syndicates across interbank networks in real time.

ACKNOWLEDGEMENT

The authors convey heartfelt appreciation to project guide Mr. Dnyaneshwar Balu Lokhande (Assistant Professor) and the faculty members of the Department of Master of Computer Application at JCEI's Jaihind Institute of Management and Research, Kuran, Pune, for their insightful technical counsel, continuous encouragement, and departmental support throughout the realization of this research.

REFERENCES

[1] Nilson Report, "Global Payment Card Fraud Losses Reach $32 Billion," Issue 1209, pp. 1-9, 2024.

[2] F. Carcillo, Y. A. Le Borgne, O. Caelen, and G. Bontempi, "Streaming active learning, chickening out, and fraud detection," Information Sciences, vol. 547, pp. 248-264, 2021.

[3] A. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in IEEE SSCI, 2015, pp. 159-166.

[4] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," Decision Support Systems, vol. 50, no. 3, pp. 602-613, 2018.

[5] Z. Wang, et al., "Semi-supervised graph neural networks for financial fraud detection," in Proc. IEEE ICDM, 2022, pp. 598-607.

[6] F. Carcillo, et al., "Scarff: a framework for addressing drift and class imbalance in fraud detection," Knowledge-Based Systems, vol. 188, p. 104865, 2020.

[7] IEEE Computational Intelligence Society, "IEEE-CIS Fraud Detection Benchmark Dataset," Kaggle Competitions, Tech. Rep., 2019.

[8] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in NeurIPS, vol. 30, 2017, pp. 4765-4774.

[9] Reserve Bank of India (RBI), "Master Direction on Digital Payment Security Controls," RBI/2020-21/74, Tech. Rep., 2021.

[10] PCI Security Standards Council, "Payment Card Industry Data Security Standard (PCI-DSS) v4.0," Tech. Rep., 2022.