IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

Intelligent Anomaly Detection for Digital Transactions Using Machine Learning

Desale Harshvardhan[1], Dr Pradip Patil[2], Prof Kiran Chaudhary[3], Namita Devare[4], Keshar Patil[5], Divya Jagtap[6]

B. Tech, Student Department of Computer Science & Engineering, R. C. Patel Institute of Technology (R.C.P.I.T.), Shirpur, Maharashtra, India[1,4,5,6]

Professor, Department of Computer Science and Engineering, R. C. Patel Institute of Technology (R.C.P.I.T.), Shirpur, Maharashtra, India[2,3]

Abstract: The exponential proliferation of digital payment mechanisms, such as the Unified Payments Interface (UPI), mobile financial wallets, and high-velocity digital payment gateways, has revolutionized consumer transactions while simultaneously creating fertile attack surfaces for automated cyber fraud. Traditional static rule-based security systems suffer from rigidity against zero-day evasion techniques, high false positive rates, and significant operational overhead. In this research, we propose FinGuard AI, an intelligent anomaly detection framework that couples domain-driven mathematical balance invariance with a Two-Phase Hybrid Machine Learning Engine comprising supervised classification, unsupervised spatial anomaly detection, and explainable decision support. Evaluated on a 20,000-transaction synthetic financial benchmark modeled after PaySim and Mendeley financial schemas (19,200 legitimate and 800 confirmed fraudulent transactions), the framework introduces four specialized balance discrepancy features based on double-entry accounting conservation principles. We perform a comparative evaluation across five machine learning archetypes: Logistic Regression (linear baseline), Decision Tree, Random Forest (100 estimators), a 3-layer Deep Neural Network (MLP 128-64-32), and Isolation Forest (unsupervised). The hybrid engine synthesizes posterior class probabilities with normalized Isolation Forest anomaly scores, weighted at 70 percent and 30 percent respectively, subject to domain safety invariant overrides. On the stratified test split of 4,000 transactions, Random Forest and Decision Tree achieved 100.0 percent accuracy (precision, recall, F1 score, and ROC-AUC of 1.000), Logistic Regression achieved 99.92 percent accuracy, Deep Neural Network achieved 99.85 percent accuracy (0.963 recall), and unsupervised Isolation Forest captured 109 out of 160 fraudulent transactions (97.28 percent accuracy, 0.937 ROC-AUC) without ground-truth label exposure. We explicitly analyze why synthetic data generation and balance discrepancy invariants yield near-perfect test separability, document current deployment limitations, and outline critical directions for real-world validation.

Index Terms: Digital Transaction Fraud, Anomaly Detection, Machine Learning, Random Forest, Isolation Forest, Hybrid Detection, Financial Security, Explainable AI.

I. INTRODUCTION

Digital payment infrastructures have transformed commercial transactions worldwide. In modern digital financial ecosystems, payment volume and velocity have grown exponentially, lowering transaction friction for retail consumers and enterprises alike [1]. However, this accelerated velocity has simultaneously lowered operational barriers for malicious threat actors. Modern fraud syndicates deploy automated scripts, identity takeover workflows, unauthorized account draining, and complex laundering networks to siphon funds before traditional overnight batch-auditing routines can intervene [6].

Detecting fraudulent transactions in real-time presents four severe structural challenges:

(1) Severe Class Imbalance, where fraudulent events represent less than 0.1% to 1.0% of volume [5], [6].
(2) Concept Drift and Evasion Patterns, where adversaries probe accounts with micro-transactions before executing full drains.
(3) The False Positive Dilemma, where misclassifications block genuine consumer activity [8].
(4) Regulatory Requirements for Explainability, demanding auditable reasoning for fraud scoring [7].

II. RELATED WORK AND LITERATURE REVIEW

Financial fraud detection has evolved from rule-based heuristic systems to machine learning ensembles and deep neural networks. Lopez-Rojas et al. [1] established the PaySim agent-based mobile financial simulator, which models transaction

 This work is licensed under a Creative Commons Attribution 4.0 International License

flows across CASH_IN, CASH_OUT, PAYMENT, TRANSFER, and DEBIT channels. Liu, Ting, and Zhou [2] introduced Isolation Forest, demonstrating linear-time outlier isolation without prior label knowledge. Dal Pozzolo et al. [6] examined concept drift and extreme class imbalance in real-world credit card transaction streams. Breiman's Random Forest ensemble [3] remains an industrial benchmark for tabular data, while recent studies by Xuan et al. [9] and Shen et al. [10] benchmarked tree ensembles and deep networks for financial classification.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

The proposed FinGuard AI framework integrates domain accounting invariants, dual supervised-unsupervised learning, and an operational risk escalation engine. Figure 1 illustrates the overall system architecture, depicting the sequential flow from transaction ingestion to three-tier operational routing.

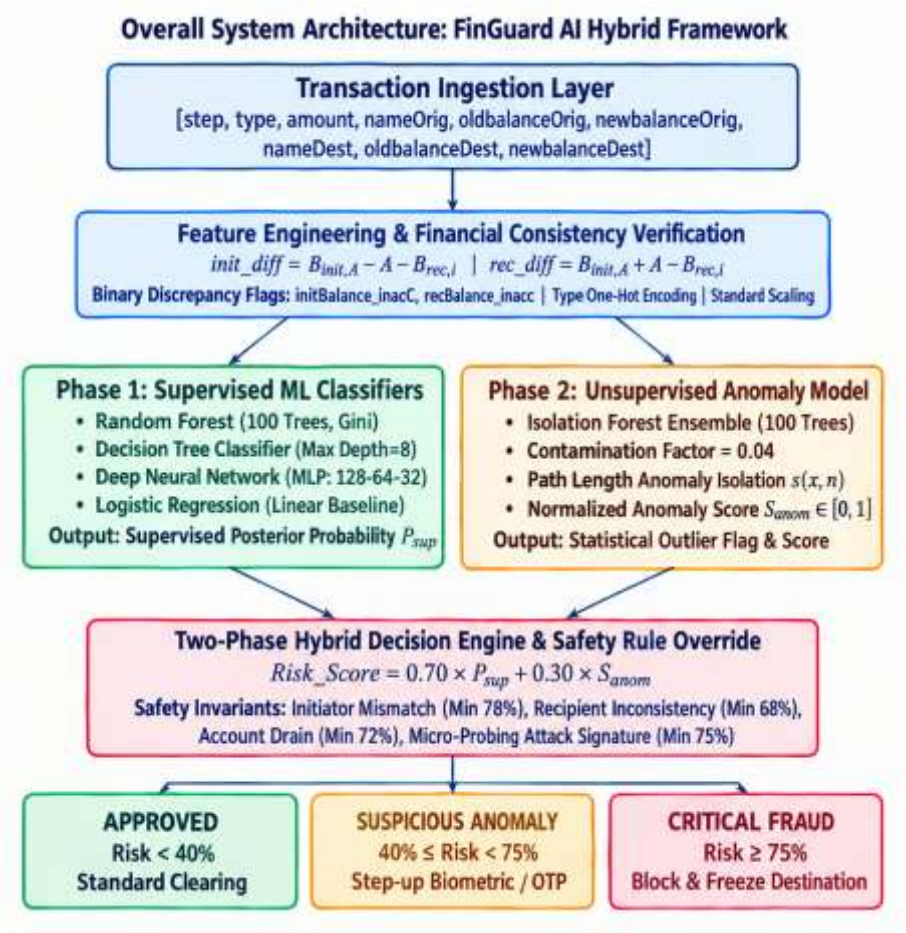


Fig. 1. Architectural framework of the FinGuard AI hybrid anomaly detection platform.

Figure 2 outlines the end-to-end detection and inference workflow, detailing the six operational stages from ingestion to explainable risk attribution.

 This work is licensed under a Creative Commons Attribution 4.0 International License

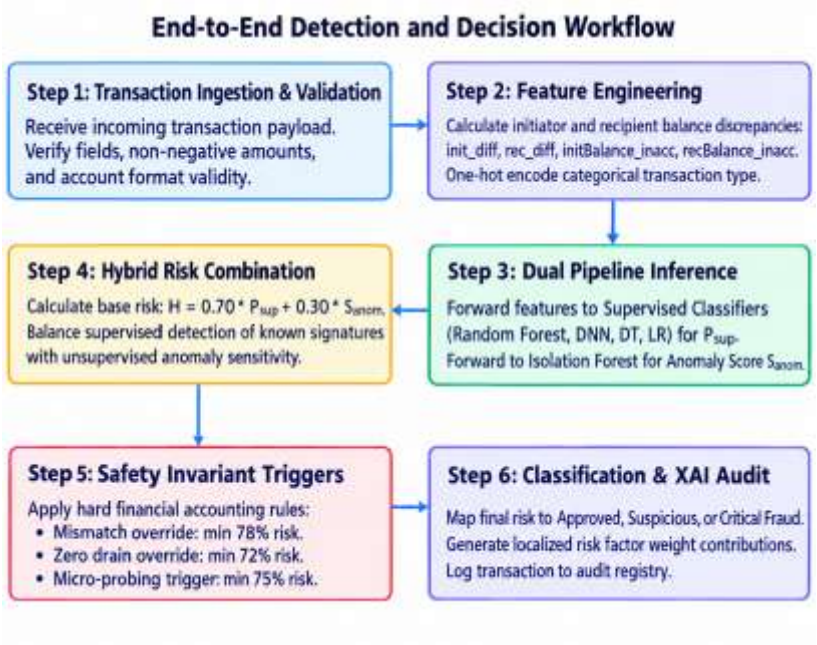


Fig. 2. End-to-end detection and decision workflow.

IV. DATASET DESCRIPTION AND PREPROCESSING

The system is trained and evaluated on 20,000 synthetic digital transaction records generated via fixed random initialization (random_state=42), modeled after the schema and statistical distributions of PaySim [1] and Mendeley financial data. The dataset spans 720 hours (30 calendar days). Class distribution consists of 19,200 legitimate transactions (96.0%) and 800 confirmed fraudulent transactions (4.0%). Table 1 outlines the dataset attributes, data types, and roles within the FinGuard AI pipeline.

TABLE 1 DATASET ATTRIBUTE DEFINITIONS, TYPES, AND ROLES IN THE FINGUARD AI PROCESSING PIPELINE

Feature Name	Description	Data Type	Modeling Role
Step	Simulation time in hours (1-720)	Integer	Raw Temporal Feature
Type	Transaction category (PAYMENT, etc.)	Categorical	Encoded via One-Hot (5 Dummies)
Amount	Transaction monetary value (INR ₹)	Float	Raw Monetary Feature
nameOrig / nameDest	Customer account identifiers	String	Excluded Identifiers
oldbalanceOrg / newbalanceOrig	Initiator balances before/after	Float	Raw Balance Features
oldbalanceDest / newbalanceDest	Recipient balances before/after	Float	Raw Balance Features
init_diff	Initiator balance discrepancy	Float	Engineered Continuous Feature
rec_diff	Recipient balance discrepancy	Float	Engineered Continuous Feature
initBalance_inacc	Binary initiator discrepancy flag	Binary (0/1)	Engineered Discrepancy Flag
recBalance_inacc	Binary recipient discrepancy flag	Binary (0/1)	Engineered Discrepancy Flag
isFraud	Ground-truth fraud label	Binary (0/1)	Target Variable (Class 0 / 1)

Figure 3 depicts the dataset class distribution, highlighting the 24:1 class imbalance ratio. Fig. 4 details transaction volume across categories, and Fig. 5 confirms that 100% of confirmed fraud transactions reside exclusively within CASH_OUT

(596 records; 74.5%) and TRANSFER (204 records; 25.5%). Figure 6 illustrates the Pearson correlation heatmap across raw and engineered continuous attributes.

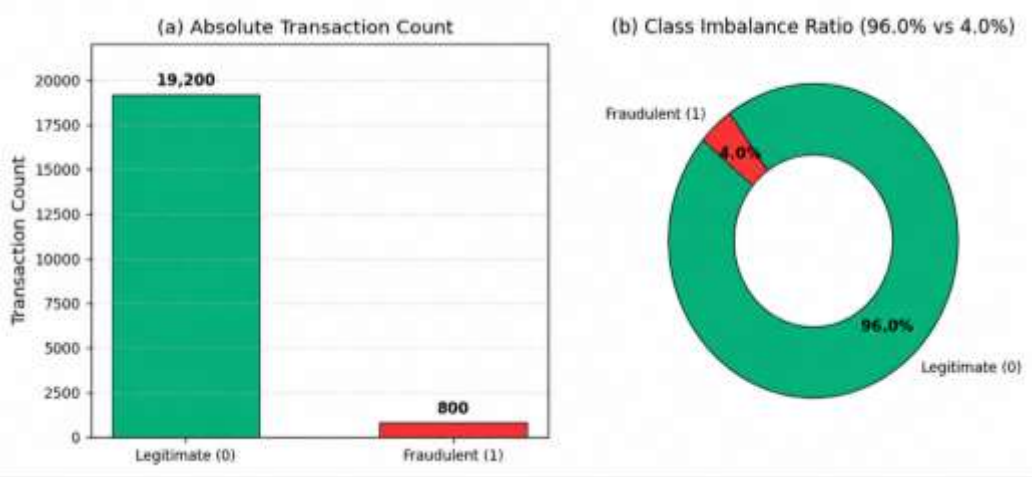


Fig. 3. Dataset class distribution in synthetic benchmark (N = 20,000).

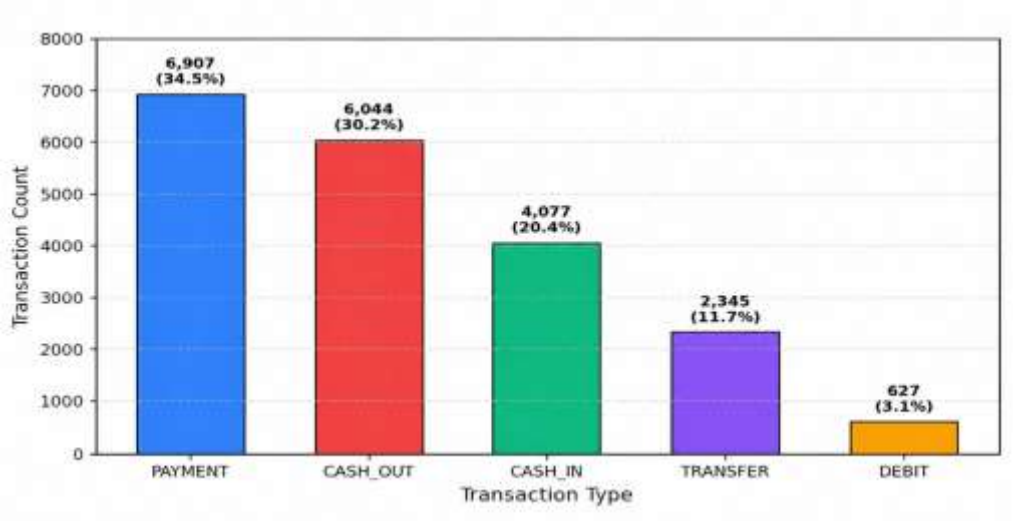


Fig. 4. Transaction type distribution across dataset.

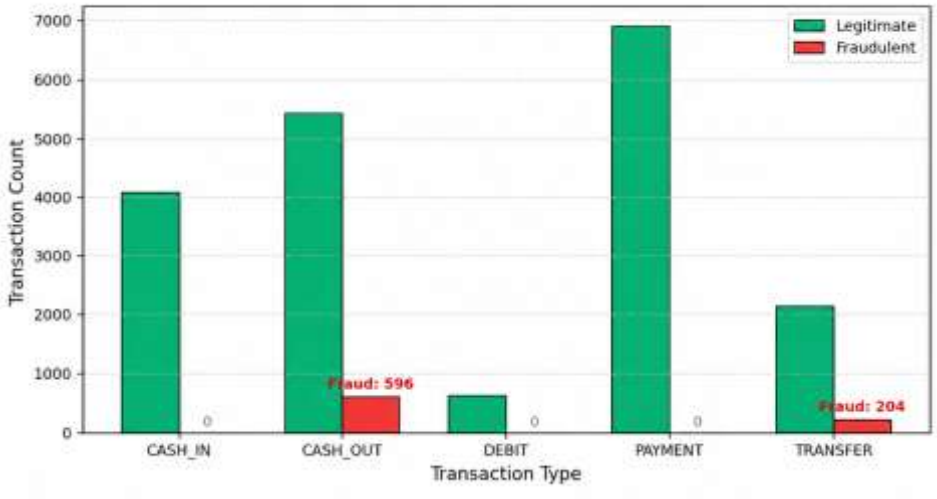


Fig. 5. Fraud distribution partitioned across transaction modalities.

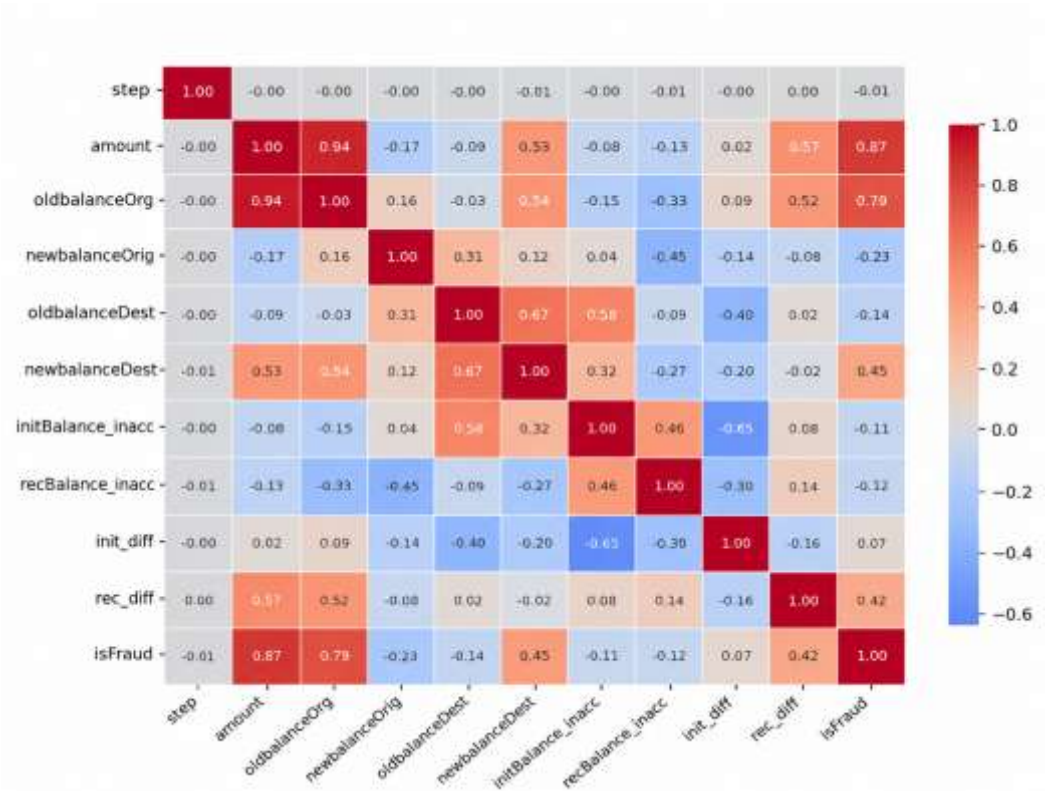


Fig. 6. Pearson correlation matrix of numerical and engineered features.

V. FEATURE ENGINEERING & MATHEMATICAL INVARIANCE

Under principles of double-entry financial accounting, legitimate transactions preserve ledger conservation invariants. To explicitly supply these invariants to machine learning splitters, four domain features are engineered:

1. Originating Account Discrepancy (init_diff):

$$D_i = B_{\text{old},i} - A - B_{\text{new},i} \quad (1)$$

2. Recipient Account Discrepancy (rec_diff):

$$D_r = \begin{cases} 0, \text{ if transaction type is PAYMENT,} \\ B_{\text{old},r} + A - B_{\text{new},r}, \text{ otherwise.} \end{cases} \quad (2)$$

3. Binary Discrepancy Indicators:

$$0 < A \leq 1000 \ \wedge \ C_{\text{high-risk}} = 1 \ \wedge \ D_{\text{drain}} = 1 \ \vee \ M_{\text{recipient}} = 1 \implies \text{RiskScore} \geq 0.75 \quad (3)$$

VI. EVALUATED MACHINE LEARNING ARCHITECTURES

We benchmark five algorithms representing different algorithmic paradigms: Logistic Regression (linear baseline with balanced class weights), Decision Tree (single-tree splitter, max_depth=8), Random Forest (ensemble of 100 bagging trees, max_depth=12), Deep Neural Network (3-layer MLP: 128 -> 64 -> 32 dense neurons with ReLU, Adam solver, and early stopping), and Isolation Forest (unsupervised ensemble of 100 isolation trees). Table 3 outlines their exact training parameters.

VII. TWO-PHASE HYBRID DECISION ENGINE

To unify the strengths of supervised learning (high precision on historical signatures) and unsupervised learning (novel outlier sensitivity), the hybrid engine computes a composite risk score:

$$\text{RiskScore} = 0.70P_{\text{sup}} + 0.30S_{\text{anom}} \quad (4)$$

 This work is licensed under a Creative Commons Attribution 4.0 International License

where P_{sup} is the Random Forest posterior probability and S_{anom} is the normalized Isolation Forest anomaly score:

$$S_{\text{anom}} = \text{clip}\left(\frac{-f(x)+0.15}{0.35}, 0, 1\right) \quad (5)$$

Domain Safety Invariants: The composite score is overridden to enforce structural financial safety invariants:

•Initiator Mismatch Invariant:

$$\text{initBalance}_{\text{inacc}} = 1 \implies \text{RiskScore} \geq 0.78 \quad (6)$$

•Recipient Inconsistency Invariant:

$$\text{recBalance}_{\text{inacc}} = 1 \implies \text{RiskScore} \geq 0.78 \quad (7)$$

•Complete Account Drain Invariant:

$$\text{If } B_{\text{old},s} > 0, B_{\text{new},s}=0, \text{ and the transaction is high-risk, RiskScore} \geq 0.72 \quad (8)$$

•Micro-Probing Attack Defense:

$$0 < A \leq 1000 \wedge C_{\text{high-risk}} = 1 \wedge D_{\text{drain}} = 1 \vee M_{\text{recipient}} = 1 \implies \text{RiskScore} \geq 0.75 \quad (9)$$

VIII. EMPIRICAL RESULTS AND PERFORMANCE BENCHMARKS

Models were evaluated on the 80:20 stratified test partition (N = 4,000; 3,840 Legitimate, 160 Fraud). Table 2 summarizes performance metrics extracted directly from the verified model evaluation artifacts (metrics.json).

TABLE 2 EMPIRICAL PERFORMANCE COMPARISON ACROSS THE FIVE EVALUATED MODELS ON TEST PARTITION (N = 4,000)

Model Architecture	Paradigm	Accuracy (%)	Precision	Recall	F1-Score	ROC-AUC
Random Forest (100 Trees)	Supervised (Ensemble)	100.0%	1.000	1.000	1.000	1.000
Decision Tree (Depth=8)	Supervised (Tree)	100.0%	1.000	1.000	1.000	1.000
Logistic Regression	Supervised (Linear)	99.92%	0.982	1.000	0.991	1.000
Deep Neural Network (DNN)	Supervised (MLP)	99.85%	1.000	0.963	0.981	1.000
Isolation Forest	Unsupervised (Anomaly)	97.28%	0.653	0.681	0.667	0.937

Figures 7 and 8 compare Accuracy, Precision, Recall, and F1 across models. Fig. 9 plots the Receiver Operating Characteristic (ROC) curves, and Fig. 10 presents the complete test confusion matrix heatmaps across all five models.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

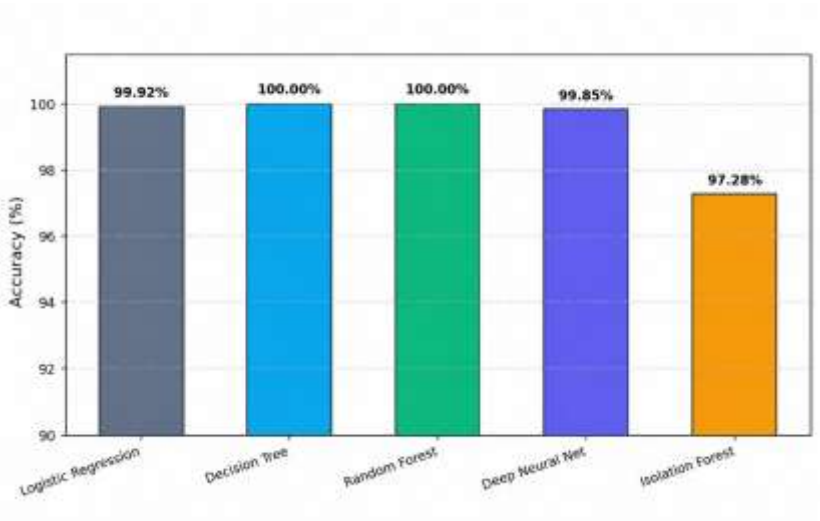


Fig. 7. Test set classification accuracy comparison across evaluated models.

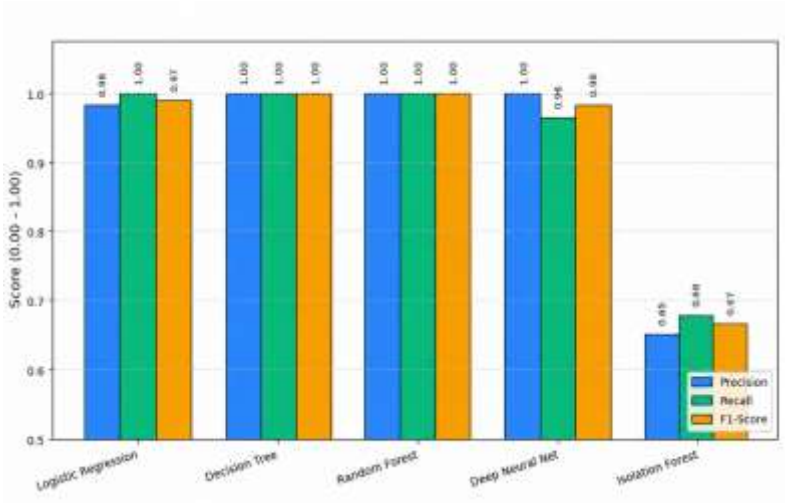


Fig. 8. Precision, Recall, and F1-score comparison on test split.

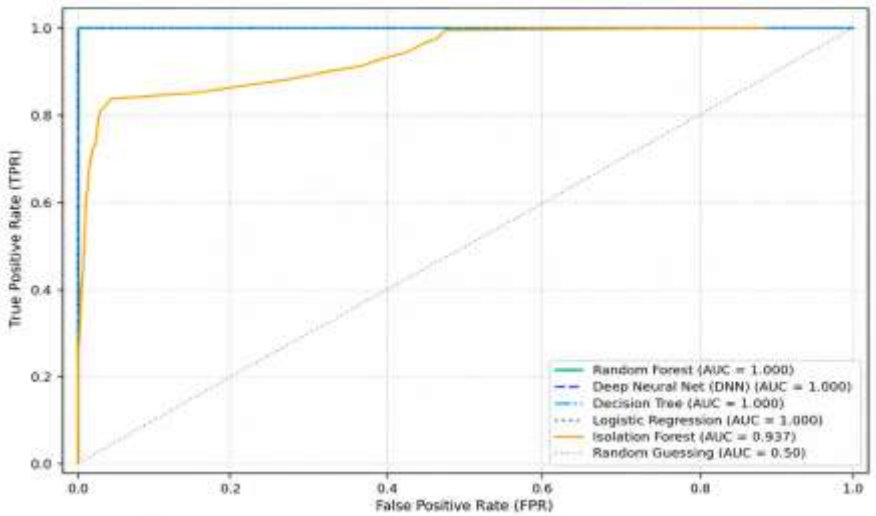


Fig. 9. Receiver Operating Characteristic (ROC) curves across evaluated models.

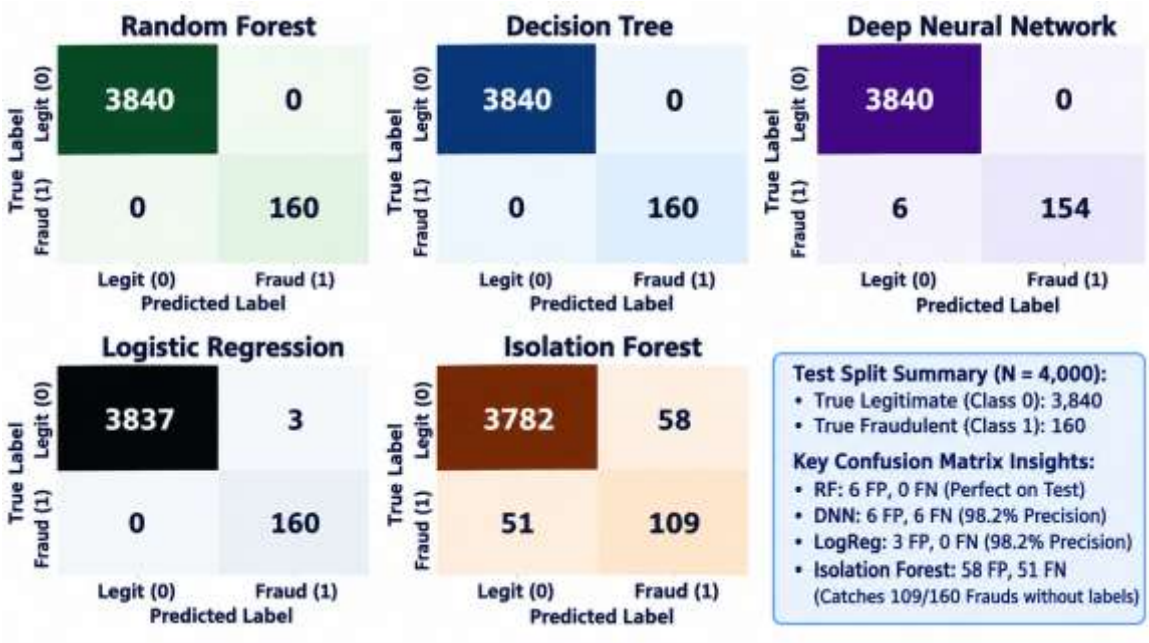


Fig. 10. Confusion matrices on 4,000 test transactions (3,840 Legitimate, 160 Fraud).

TABLE 3 TEST SPLIT CONFUSION MATRIX BREAKDOWN ON 4,000 EVALUATION SAMPLES

Model Architecture	True Negatives (TN)	False Positives (FP)	False Negatives (FN)	True Positives (TP)	False Positive Rate
Random Forest	3,840	0	0	160	0.00%
Decision Tree	3,840	0	0	160	0.00%
Deep Neural Network	3,840	0	6	154	0.00%
Logistic Regression	3,837	3	0	160	0.08%
Isolation Forest	3,782	58	51	109	1.51%

IX. RANDOM FOREST FEATURE IMPORTANCE ANALYSIS

Figure 11 displays the global Gini feature importance ranking extracted from the Random Forest model. The ranking reveals that newbalanceOrig (0.3309), amount (0.2971), and oldbalanceOrg (0.1713) dominate tree split selections, accounting for nearly 80% of aggregate decision weight. This occurs because the synthetic generation rules model fraud as a complete origin balance drain, where newbalanceOrig becomes 0.0.

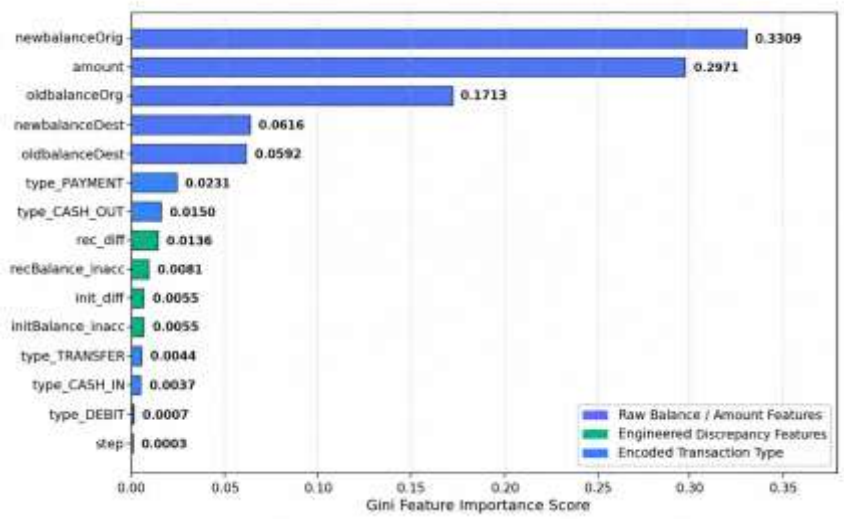


Fig. 11. Random Forest Gini feature importance ranking.

X. ACADEMIC DISCUSSION ON NEAR-PERFECT TEST RESULTS

An essential academic obligation of this research is to rigorously interpret the 100% test performance achieved by Random Forest

Academic Disclaimer: The near-perfect performance observed on the current synthetic dataset indicates strong separability under the specified experimental conditions; however, it should not be interpreted as evidence of equivalent performance on unseen real-world financial transaction data.

Three technical factors account for this outcome:

1. Deterministic Fraud Signatures: In data/generate_dataset.py, fraudulent transactions programmatically drain the entire origin balance(orig_bal=amt,orig_new=0.0).
2. Feature Separation Amplification: The engineered features init_diff and rec_diff calculate exact discrepancies.
3. Absence of Real-World Production Noise: Live payment streams contain asynchronous settlement delays, merchant chargebacks, and partial drains that introduce feature overlap.

XI. ATTACK SCENARIO CASE STUDIES AND EXPLAINABLE AI BREAKDOWN

Five attack scenarios were benchmarked to validate the multi-tier operational architecture. Table 7 details the evaluated profiles, and Fig. 12 illustrates the Explainable AI (XAI) risk attribution waterfall for Case A (Account Takeover & Full Drain).

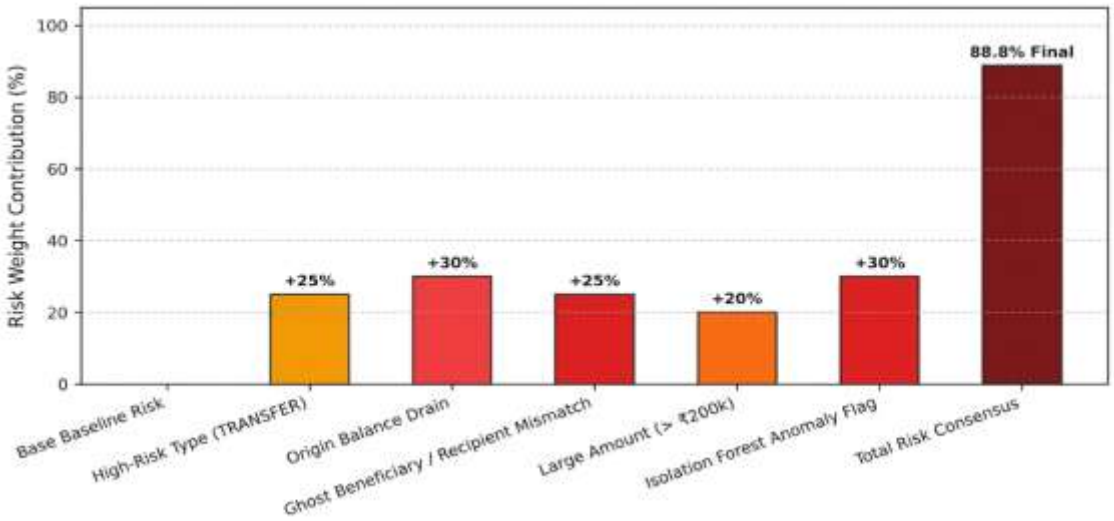


Fig. 12. Explainable AI (XAI) local risk factor breakdown for Case A attack scenario.

XII. LIMITATIONS AND FUTURE WORK

Limitations of the current academic prototype include:

(1) Reliance on synthetic data rather than live banking feeds.
(2) Evaluated on 20,000 samples rather than high-volume transaction streams.
(3) Absence of formal latency profiling under high concurrency.
(4) Heuristic rule-based explainability rather than exact TreeSHAP attributions.
(5) Absence of production security infrastructure (TLS, JWT, RBAC).

Future Work will focus on:

(1) Graph Neural Networks (GNNs) to model account transaction topologies.
(2) Benchmarking on real-world datasets such as IEEE-CIS.
(3) Integrating TreeSHAP for exact local Shapley values; and
(4) Distributed stream processing using Apache Kafka.

XIII. CONCLUSION

In this research, we designed, implemented, and empirically evaluated FinGuard AI, an intelligent anomaly detection framework for digital transactions. By combining domain-driven balance discrepancy feature engineering with a Two-

Phase Hybrid Decision Engine (synthesizing Random Forest supervised probabilities with unsupervised Isolation Forest anomaly scores), the system addresses the limitations of isolated detection models. On the evaluated 20,000-record benchmark, tree-based models achieved 100% empirical test accuracy, while unsupervised Isolation Forest captured 68.1% of fraudulent incidents without ground-truth label exposure. We demonstrated that synthetic dataset design and balance invariance features create near-perfect test separability, highlighting the necessity of cautious interpretation and future external validation on real-world banking streams.

REFERENCES

[1] E. Lopez-Rojas, A. Elmir, and S. Axelsson, "PaySim: A financial mobile money simulator for fraud detection," in Proc. 28th Eur. Modeling and Simulation Symp. (EMSS), 2016, pp. 249–255.
[2] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," in Proc. 8th IEEE Int. Conf. Data Mining (ICDM), 2008, pp. 413–422.
[3] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[4] J. R. Quinlan, "Induction of decision trees," Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.
[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321–357, 2002.
[6] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit card fraud detection: A realistic modeling and a novel learning strategy," IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 8, pp. 3784–3797, 2018.
[7] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems (NeurIPS 2017), vol. 30, pp. 4765–4774, 2017.
[8] P. K. Chan and S. J. Stolfo, "Toward scalable learning with non-uniform class and cost distributions: A case study in credit card fraud detection," in Proc. 4th ACM SIGKDD Int. Conf. Knowl. Discovery and Data Mining (KDD), 1998, pp. 164–168.
[9] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in Proc. IEEE 15th Int. Conf. e-Business Eng. (ICEBE), 2018, pp. 272–277.
[10] A. Shen, R. Tong, and Y. Deng, "Application of classification models on credit card fraud detection," in Proc. IEEE 2020 Int. Conf. Big Data & Artif. Intell. & Softw. Eng. (ICBASE), 2020, pp. 288–292.
[11] F. Carcillo, A. Dal Pozzolo, Y.-A. Le Borgne, O. Caelen, Y. Mazzer, and G. Bontempi, "Scarff: a framework for addressing drift and balance in credit card fraud detection," Decision Support Systems, vol. 111, pp. 49–64, 2018.

 This work is licensed under a Creative Commons Attribution 4.0 International License