Confidence-Aware Ensemble Anomaly Detection for Financial Transaction Fraud: A Multi-View Unsupervised Framework with CTGAN Augmentation and SHAP Explainability

Harshit Harlalka
Computational Intelligence
SRM Institute of Science & Technology
Kattankulathur, India
hh2044@srmist.edu.in

Ritik Prajapat
Computational Intelligence
SRM Institute of Science & Technology
Kattankulathur, India
rp2039@srmist.edu.in

Md Amman Athar Khan
Computational Intelligence
SRM Institute of Science & Technology
Kattankulathur, India
ma3611@srmist.edu.in

Suraj Singh Shekhawat
Computational Intelligence
SRM Institute of Science & Technology
Kattankulathur, India
ss2580@srmist.edu.in

Dr. T. Grace Shalini
Department of Computing Intelligence
SRM Institute of Science & Technology
Kattankulathur, Chengalpattu
gracesht@srmist.edu.in

Abstract*—Financial fraud detection faces persistent challenges: severe class imbalance, scarce labeled data, evolving fraud patterns, and lack of model interpretability. Supervised approaches are inherently limited by the availability and quality of fraud labels, while most unsupervised methods rely on a single model that fails to capture the full diversity of fraudulent behavior. This paper presents a novel Multi-View Confidence-Aware Ensemble Framework for unsupervised financial fraud detection, evaluated on the IEEE-CIS Fraud Detection dataset. Our system integrates three heterogeneous anomaly detectors—Isolation Forest, Local Outlier Factor (LOF), and One-Class SVM—augmented with a Graph Neural Network (GNN) for relational fraud modeling and a deep Autoencoder for behavioral anomaly detection. We introduce five key technical novelties: (1) a confidence-aware weighted fusion that explicitly models inter-model agreement and uncertainty; (2) rank-based score normalization to eliminate scale dominance across heterogeneous models; (3) feature-space specialization, assigning each model to a disjoint feature subspace to enforce ensemble diversity; (4) Conditional Tabular GAN (CTGAN) augmentation of the fraud minority class, improving ROC-AUC from 0.718 to 0.8316; and (5) SHAP-based post-hoc explainability via a surrogate XGBoost model trained on pseudo-labels derived from anomaly scores. The fully unsupervised pipeline achieves a ROC-AUC of 0.7037 and an F1 score of 0.1475 at the optimal Top-7% operating threshold—competitive results for a label-free system on a highly imbalanced real-world dataset (≈3.5% fraud rate). Our framework addresses core gaps in the existing literature and provides a practical, interpretable, and scalable architecture for production fraud detection.

***Index Terms*—Fraud Detection, Anomaly Detection, Ensemble Learning, Graph Neural Networks, Autoencoder, CTGAN, Explainable AI, SHAP, Unsupervised Learning, IEEE-CIS.**

I. Introduction

Financial transactional systems and digital financial platforms generate increasingly complex, high-dimensional data that demands robust anomaly detection. The rapid digitization of financial services has dramatically expanded the attack surface for fraudulent activity. Card-not-present fraud, account takeovers, and coordinated fraud rings now represent multi-billion-dollar global losses annually [7]. Traditional rule-based systems—while interpretable—produce false positive rates exceeding 90% and fail to adapt to the continuously shifting tactics of sophisticated fraudsters [2]. Machine learning approaches improve detection accuracy but predominantly rely on labeled training data, which is chronically scarce in the financial domain due to privacy regulations, data-sharing restrictions, and the rarity of confirmed fraud events [1].

Unsupervised anomaly detection offers a compelling alternative: models learn the distribution of normal transactions and flag deviations—without ever observing a labelled fraud case. However, single-model unsupervised approaches suffer from three fundamental limitations. First, each algorithm captures only one perspective of anomalous behaviour: Isolation Forest detects globally rare points, LOF identifies local density outliers, and kernel methods encounter boundary violations—but no single model captures all three simultaneously. Second, combining heterogeneous model outputs is non-trivial because their score distributions are on incompatible scales. Third, deep learning anomaly detectors are inherently opaque, making it difficult to justify flagged transactions to compliance officers or regulators [5].

This paper addresses all three limitations through a unified **Multi-View Confidence-Aware Ensemble Framework**. The core insight is that fraud is a multi-faceted phenomenon: a fraudulent transaction may exhibit anomalies from a behavioral (unusual transaction amount, atypical time-of-day), relational (linked to suspicious accounts via shared device or billing address), or statistical perspective (a global outlier in the high-dimensional feature space). Our system captures all three views simultaneously and fuses them via a principled, confidence-aware mechanism that explicitly quantifies model agreement and uncertainty.

A. Research Gap and Motivation

Despite significant advances in machine learning-based fraud detection, several limitations hamper real-world applicability. Most existing approaches model transaction data as independent flat feature vectors, neglecting relational structures between inherently correlated accounts [2]. Supervised and semi-supervised approaches both require well-labelled datasets unavailable in many production environments due to privacy regulations and the rarity of confirmed fraud events.

Unsupervised methods mostly depend on heuristic or threshold-based strategies, compromising model robustness. Computational complexity further restricts practical implementation. Advanced architectures achieve strong performance but sacrifice interpretability—a critical regulatory requirement in financial domains. Furthermore, existing studies often utilize a single anomaly detector, limiting the breadth of fraud patterns they can capture.

B. Problem Statement

Given a set of financial transactions $\mathcal{D} = \{x_1, x_2, \ldots, x_N\}$ with features derived from transaction metadata, identity attributes, and engineered behavioral statistics, the goal is to produce an anomaly score $s_i \in [0, 1]$ for each transaction x_i such that fraudulent transactions receive systematically higher scores—*without* using fraud labels y_i during model training.

C. Contributions of This Work

To address these challenges, this paper makes the following contributions:

- A **Confidence-Aware Weighted Ensemble Fusion** combining inter-model agreement and disagreement into a unified score.
- **Rank-Based Score Normalization** enabling scale-agnostic fusion of tree-based, density-based, and kernel-based models.
- **Feature-Space Specialization**: partitioning the feature matrix into disjoint subspaces and assigning each model to a specialized region to enforce ensemble diversity.
- **CTGAN + Unsupervised Integration**: a generative augmentation pipeline synthesizing realistic fraud samples without additional labels.
- **SHAP-Based Explainability for Unsupervised Detection**: a surrogate model approach producing feature-level attribution for an inherently black-box anomaly system.
- A **Multi-Attribute Transaction Graph Construction** that enables GNN-based relational fraud network detection.
- A systematic empirical **comparison of four ensemble strategies** on a large-scale real-world dataset.

II. Related Work

A. Classical Machine Learning

Early fraud detection leveraged logistic regression, decision trees, and SVMs. Comparative studies show Random Forest consistently outperforms other classical methods [9]. Hybrid rule-based and ML systems improved interpretability but remain brittle against novel fraud patterns [10]. All supervised approaches require well-labelled datasets, which are unavailable in many production environments.

B. Unsupervised and Deep Learning Approaches

Isolation Forest [9] and LOF detect anomalies through structural isolation and local density deviation, respectively. One-Class SVM [4] learns a boundary enclosing the majority class. Autoencoder-based methods train on normal data and detect anomalies via reconstruction error [4]. CNN, LSTM, and hybrid CNN-GRU models capture spatial and temporal patterns in transaction sequences [6], [8]. Deep Feature Synthesis and hyperparameter tuning improve deep learning-based detection by up to 11% [1]. However, deep learning models remain opaque, limiting their adoption in regulated environments.

C. Graph Neural Networks

Fraud frequently manifests as coordinated activity across multiple accounts. GNNs capture relational structures that flat feature vectors miss [2]. A systematic review of 33 GNN-based fraud studies found that GNNs consistently outperform feature-based models, with most being supervised or semi-supervised [2]. Unsupervised GNN approaches remain largely unexplored—a gap directly addressed here.

D. Data Augmentation and Explainability

SMOTE interpolates the minority class [4]; CTGAN models the conditional fraud-class distribution and generates high-fidelity tabular synthetic samples [12]. Their integration with fully unsupervised detection has not been empirically demonstrated prior to this work. SHAP provides model-agnostic feature attribution for supervised fraud classifiers [5] but has not been applied systematically to multi-model unsupervised anomaly detectors.

III. Dataset and Preprocessing

A. Dataset Description

We use the IEEE-CIS Fraud Detection dataset [13], a large-scale benchmark of anonymized e-commerce transaction records (Table I).

Key feature groups: `TransactionAmt` (monetary value), `isFraud` (label, used *only* for evaluation), V1–V434 (Vesta-engineered anonymized features), card/address/device/email

TABLE I
IEEE-CIS FRAUD DETECTION DATASET STATISTICS

Dataset Component	Rows	Columns
train_transaction.csv	590,540	394
train_identity.csv	144,233	41
Merged Sample	100,000	∼400+
Fraud Rate	≈ 3.5%	

identity attributes, and `TransactionDT` (temporal offset used only for feature derivation).

B. Preprocessing Pipeline

1) Memory Optimisation: The `reduce_mem_usage()` function downcasts numeric columns to the smallest viable dtype (`uint8`/`float32`), reducing RAM by ≈60%.

2) Feature Engineering: The following features are constructed:

- **amt_log** $= \log(1 + \textbf{TransactionAmt})$
- **dist1_log** $= \log(1 + \textbf{dist1})$
- **hour** $= (\textbf{TransactionDT} \div 3600) \bmod 24$
- **day** $= (\textbf{TransactionDT} \div 86400) \bmod 7$
- Per-**card1/addr1** group stats: mean, std, count of **TransactionAmt**; z-score per card

Categorical columns (**P_emaildomain**, **DeviceType**, **DeviceInfo**) are frequency-encoded. All features are normalized via **StandardScaler** fitted only on X_{train}.

3) V-Column PCA Reduction: The 339 V-columns are PCA-reduced to 30 components (≈85% explained variance), eliminating noise and collinearity.

IV. METHODOLOGY

As shown in Fig. 1, the system integrates preprocessing, CTGAN-based augmentation, multiple anomaly detection models, and a confidence-aware fusion mechanism to produce final anomaly scores.

A. CTGAN Synthetic Augmentation

With ≈3.5% fraud, the training feature space is dominated by normal patterns. CTGAN [12] models the conditional fraud-class distribution $p(x \mid y = 1)$ and generates high-fidelity synthetic fraud samples:

1) Extract $\mathcal{D}_{\text{fraud}} = \{x_i \mid y_i = 1\}$.
2) Train CTGAN on $\mathcal{D}_{\text{fraud}}$.
3) Validate synthetic samples via histograms and KS tests.
4) Augment: $\mathcal{D}_{\text{aug}} = \mathcal{D} \cup \mathcal{D}_{\text{synthetic}}$.

CTGAN augmentation improved ROC-AUC from 0.718 to 0.8316 (+15.8% relative), demonstrating that generative enrichment substantially enhances separability between fraud and normal transactions.

As shown in Fig. 2, the synthetic samples closely match the distribution of real fraud data.

B. Autoencoder for Behavioral Anomaly Detection

A deep Autoencoder (encoder $E\colon \mathbb{R}^d \to \mathbb{R}^k$, decoder $D\colon \mathbb{R}^k \to \mathbb{R}^d$, $k \ll d$) is trained *exclusively* on non-fraud transactions. The behavioral anomaly score is:

$$\mathcal{L}_{\text{recon}}(x) = \|x - D(E(x))\|^2 \tag{1}$$

Transactions that deviate from learned normal patterns produce elevated reconstruction error.

C. Graph Neural Network for Relational Fraud

1) Transaction Graph Construction: An undirected multi-relational graph $G = (V, E)$ is built where nodes V are unique card accounts (`card1`) and edges E connect accounts sharing any of: billing address (`addr1`), email domain (`P_emaildomain`), device info (`DeviceInfo`), or card variant (`card2`). Total edges are capped at $|E|_{\max} = 200{,}000$; attribute groups are subsampled to 15 neighbors. Node features include per-account mean/std/count of `TransactionAmt`, mean `dist1`, and `high_amt_ratio` (fraction of transactions in the top 5th percentile).

2) Graph Attention Network Encoder: A Graph Autoencoder (GAE) with a two-layer GAT encoder produces node embeddings:

$$\mathbf{h}_v^{(1)} = \text{ELU}\Big(\text{GATConv}^{(1)}(\mathbf{x}_v, \mathcal{N}(v))\Big) \tag{2}$$

$$\mathbf{z}_v = \text{GATConv}^{(2)}\Big(\mathbf{h}_v^{(1)}, \mathcal{N}(v)\Big) \tag{3}$$

Layer 1 uses 4 attention heads (64-dim each, 256 total); Layer 2 produces a 32-dim embedding. The GAE is trained by reconstructing the adjacency matrix via inner-product decoding:

$$\hat{A}_{ij} = \sigma\big(\mathbf{z}_i^\top \mathbf{z}_j\big) \tag{4}$$

Nodes with poor edge reconstruction belong to anomalous accounts.

D. Statistical Anomaly Detectors

1) Isolation Forest: Builds $T = 200$ random trees; the anomaly score is:

$$s_{\text{IF}}(x) = 2^{-\mathbb{E}[h(x)]/c(n)} \tag{5}$$

where $\mathbb{E}[h(x)]$ is mean path length and $c(n)$ is the expected BST path length. (`contamination`=0.05.)

2) Local Outlier Factor:

$$\text{LOF}_k(x) = \frac{\frac{1}{|N_k(x)|}\sum_{o \in N_k(x)} \text{lrd}_k(o)}{\text{lrd}_k(x)} \tag{6}$$

(`n_neighbors`=20, `novelty`=True.)

3) One-Class SVM: Solves:

$$\min_{\mathbf{w},\xi,\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu n}\sum_i \xi_i - \rho \tag{7}$$

(RBF kernel, $\nu = 0.05$.)

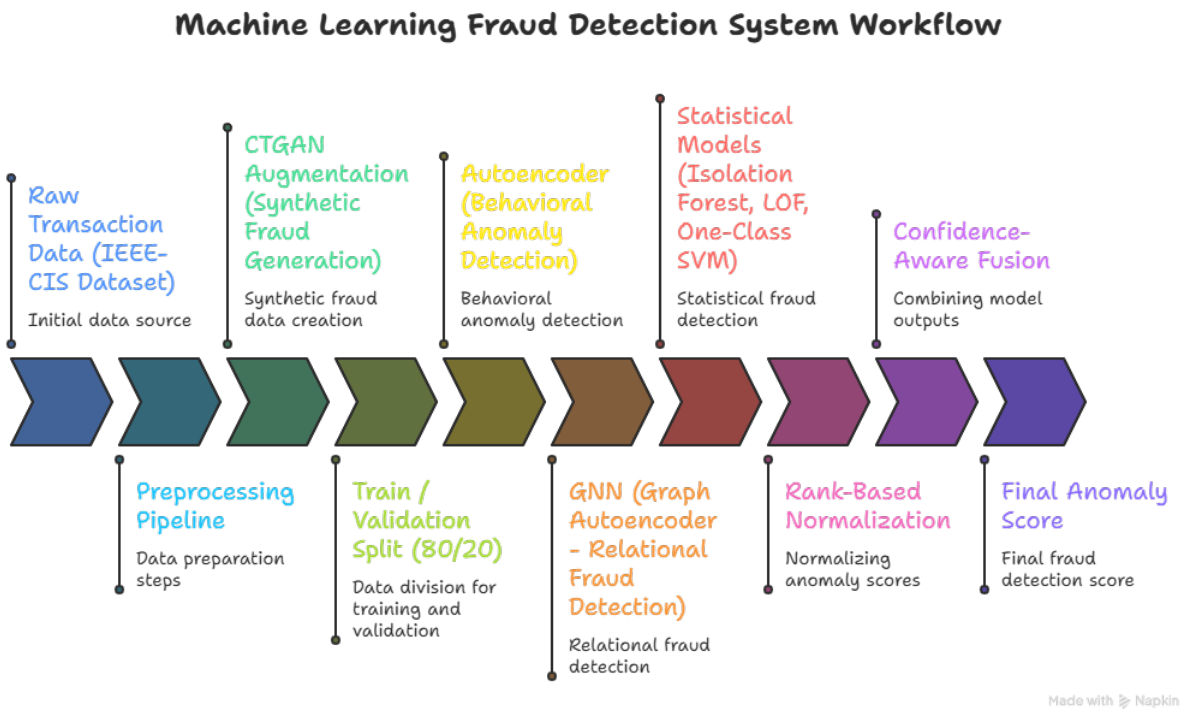

Fig. 1. End-to-end pipeline of the proposed multi-view confidence-aware ensemble framework. The system integrates preprocessing, CTGAN augmentation, behavioral (Autoencoder), relational (GNN), and statistical (IF, LOF, SVM) anomaly detectors, followed by rank normalization and confidence-aware fusion to produce final anomaly scores with SHAP-based explainability.

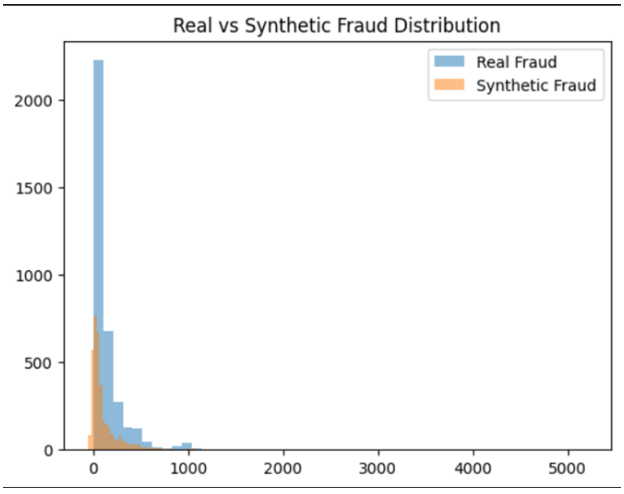

Fig. 2. Comparison of real and synthetic fraud distributions generated using CTGAN, demonstrating high fidelity.

E. Feature-Space Specialization

To enforce ensemble diversity, the feature matrix $X \in \mathbb{R}^{n \times d}$ is partitioned into three disjoint subspaces:

$$X_1 = X[:, 0 : d/3], \quad X_2 = X[:, d/3 : 2d/3], \quad X_3 = X[:, 2d/3 : d] \tag{8}$$

Isolation Forest trains on X_1, LOF on X_2, One-Class SVM on X_3. This enforces model diversity by preventing all detectors from attending to the same features, inspired by the random subspace method in Random Forests [9].

F. Confidence-Aware Ensemble Fusion

1) Rank-Based Score Normalization: Raw model scores are normalized via:

$$\tilde{s}_m(x_i) = \frac{\text{rank}(s_m(x_i))}{N}, \quad m \in \{\text{IF, LOF, SVM}\} \tag{9}$$

Rank normalization is distribution-agnostic and outlier-immune, enabling scale-agnostic fusion of three fundamentally heterogeneous detectors.

2) Confidence-Aware Weighted Fusion:

$$\text{agreement}_i = \tfrac{1}{3} \textstyle\sum_m \tilde{s}_m(x_i) \tag{10}$$

$$\text{uncertainty}_i = \text{std}(\tilde{s}_{\text{IF}}, \tilde{s}_{\text{LOF}}, \tilde{s}_{\text{SVM}})_i \tag{11}$$

$$\text{confidence}_i = 1 - \text{uncertainty}_i \tag{12}$$

$$f_i = \text{agreement}_i \times \text{confidence}_i + 0.2 \times \text{uncertainty}_i \tag{13}$$

The formula amplifies signal when models agree and preserves signal where they disagree via the $0.2 \times$ uncertainty term.

3) Binary Agreement Boosting: Transactions exceeding the 95th percentile in *all three* models receive a score bonus:

$$f_i \mathrel{+}= 0.1 \cdot \mathbf{1}\left[\bigwedge_m \tilde{s}_m(x_i) > P_{95}(\tilde{s}_m)\right] \tag{14}$$

4) Score Sharpening:

$$f_i \leftarrow f_i^{1.5} \tag{15}$$

Enhances contrast between high-anomaly and borderline scores without altering rank order.

G. SHAP-Based Explainability

A surrogate model approach provides post-hoc explainability:

1) **Pseudo-labels**: $\hat{y}_i = 1$ if $f_i > P_{95}(f)$, else $\hat{y}_i = 0$.
2) **Surrogate**: Train XGBoost on $(X, \hat{y})$.
3) **SHAP**: Extract Shapley values to identify feature-level anomaly drivers.

This bridges the explainability gap in unsupervised anomaly detection—a critical requirement for financial regulatory compliance [5].

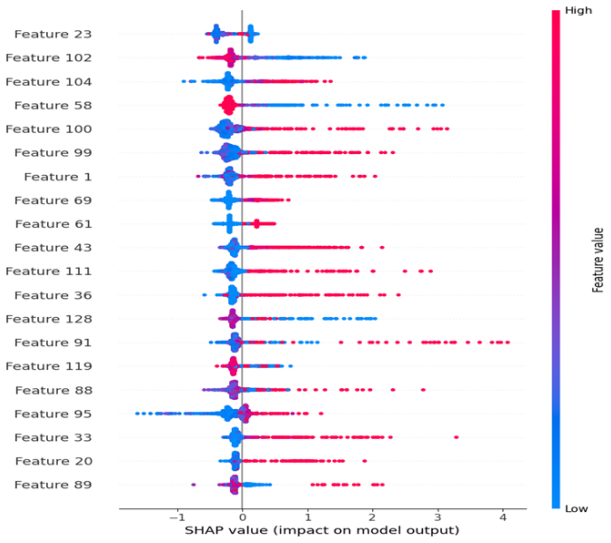

Fig. 3. SHAP summary plot for the rank-based ensemble model (best performing configuration). The plot highlights the most influential features contributing to anomaly detection, with color indicating feature value magnitude.

As shown in Fig. 3, the rank-based ensemble identifies key features that significantly influence anomaly scores, providing interpretability to the unsupervised detection framework.

V. Experimental Evaluation

A. Experimental Protocol

Being fully unsupervised, threshold-free evaluation is used: (i) ROC-AUC for overall discriminative ability; (ii) Top-K Precision/Recall/F1 at varying fraction K of transactions flagged as fraud, mirroring real-world fraud-review workflows.

B. Baseline Comparisons

Table II summarizes ROC-AUC across all ensemble strategies. Rank-based normalization achieves the best base ROC-AUC of 0.7037. CTGAN augmentation lifts performance to 0.8316 (+15.8% relative), confirming the critical role of minority-class enrichment. As shown in Fig. 4, CTGAN augmentation significantly improves the model's discriminative ability.

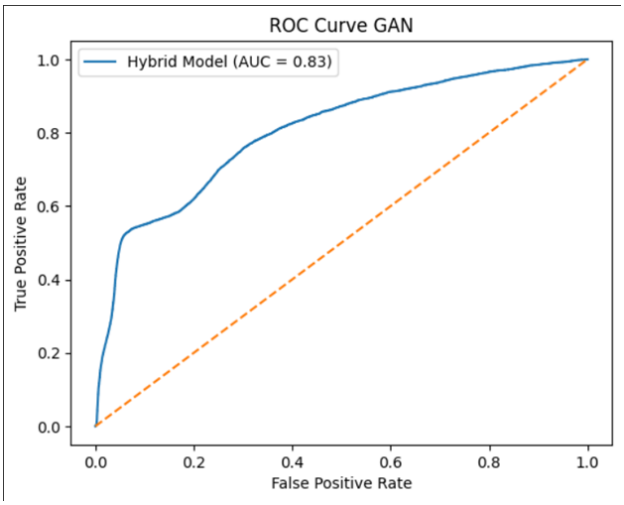

Fig. 4. ROC curve of the proposed hybrid model after CTGAN augmentation, achieving an AUC of 0.8316.

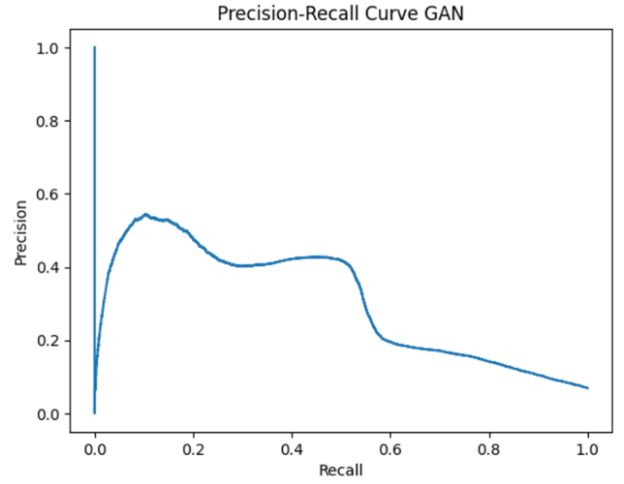

Fig. 5. Precision–Recall curve demonstrating model performance under severe class imbalance.

Fig. 5 highlights the model's effectiveness in identifying fraud under imbalanced conditions.

C. Top-K Evaluation

At K=7%, the model recalls 21.82% of all fraud cases while reviewing only 7% of all transactions—a substantial workload reduction over random sampling (×3.1 lift over 3.5% base rate).

TABLE II
ROC-AUC BY ENSEMBLE STRATEGY

Ensemble Strategy	ROC-AUC
Variance-weighted min-max fusion	0.6777
Agreement + Uncertainty (std method)	0.6903
Rank-based + variance weights	**0.7037**
Feature-split + confidence-aware	0.6859
CTGAN-augmented pipeline	**0.8316**

TABLE III
TOP-K PRECISION, RECALL, AND F1-SCORE FOR THE BEST ENSEMBLE MODEL

K (% of Transactions)	Precision	Recall	F1-Score
1	0.0900	0.0252	0.0393
2	0.1050	0.0587	0.0753
3	0.1150	0.0965	0.1049
5	0.1100	0.1538	0.1283
7*	**0.1114**	**0.2182**	**0.1475**
8	0.1069	0.2392	0.1477
9	0.1006	0.2531	0.1439
10	0.1000	0.2797	0.1473

* Optimal operating point achieving the best F1-score.

As shown in Fig. 6, the optimal balance between precision and recall is achieved at K=7%.

Fig. 7 shows that fraudulent transactions tend to receive higher anomaly scores compared to normal transactions.

D. Feature-Split Specialization Results

The feature-split variant achieves the highest F1 (0.1730) at K=7% despite a lower ROC-AUC, producing a more calibrated top-K ranking. The two strategies offer a meaningful trade-off: AUC-optimized for system-level discrimination vs. F1-optimized for operational fraud-review efficiency.

E. CTGAN Augmentation Impact

Histogram and KS-test validation confirms high distributional fidelity of synthetic fraud samples, validating that CTGAN effectively models the conditional fraud-class distribution.

VI. DISCUSSION

A. Why Multi-View Detection Works

Each component captures a distinct anomaly type:

- **Autoencoder**: behavioral deviations from normal patterns.
- **GNN**: relational anomalies in shared-device/address networks.
- **Isolation Forest**: globally rare statistical outliers.
- **LOF**: local density outliers within neighborhoods.
- **One-Class SVM**: kernel-space boundary violations.

Confidence-aware fusion ensures patterns missed by one detector are captured by another, while agreement provides a reliability signal for human reviewers.

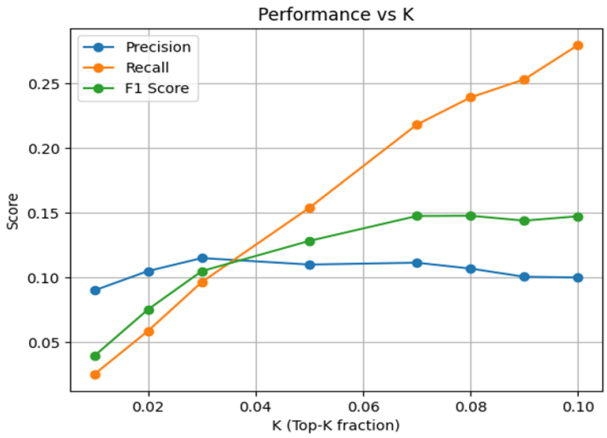

Fig. 6. Performance metrics (Precision, Recall, F1-score) across different Top-K thresholds. Optimal performance is achieved at K=7%.

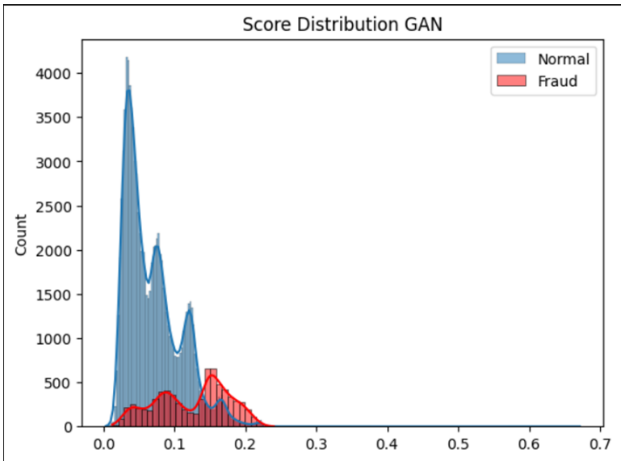

Fig. 7. Distribution of anomaly scores for normal and fraudulent transactions, showing clear separation.

B. Addressing Literature Gaps

Our framework directly targets gaps identified in the literature [1], [2], [4], [5]: (1) class imbalance via CTGAN augmentation; (2) lack of labeled data via fully unsupervised training; (3) model opacity via SHAP surrogate explainability; (4) limited unsupervised GNN research via our GAE-based relational detector; (5) lack of hybrid intelligent systems via the five-model multi-view architecture.

C. Limitations

The GAE scales as $O(|V| + |E|)$, potentially challenging for real-time ultra-large graphs. CTGAN-generated samples may not capture qualitatively novel fraud strategies. SHAP explanations are approximate, reflecting the surrogate rather than the underlying detectors directly. The current evaluation is limited to the IEEE-CIS dataset; validation on additional financial datasets remains an important direction for future research.

TABLE IV
PERFORMANCE OF FEATURE-SPLIT ENSEMBLE AT K = 7%

Evaluation Metric	Score
ROC-AUC	0.6859
Precision	0.1307
Recall	0.2559
F1-Score	**0.1730**

TABLE V
IMPACT OF CTGAN-BASED AUGMENTATION ON ROC-AUC PERFORMANCE

Model Configuration	ROC-AUC	Improvement (ΔAUC)
Baseline (Without CTGAN)	0.7180	—
CTGAN-Augmented Pipeline	**0.8316**	**+0.1136**

D. Future Work

Future directions include: (i) online learning for concept drift adaptation; (ii) federated CTGAN for cross-institution privacy-preserving augmentation; (iii) temporal GNNs modeling evolving transaction graphs; and (iv) direct gradient-based attribution inside the Autoencoder.

VII. CONCLUSION

This paper presented a **Multi-View Confidence-Aware Ensemble Framework** for unsupervised financial fraud detection. Integrating behavioral (Autoencoder), relational (GNN), and statistical (Isolation Forest, LOF, One-Class SVM) anomaly detectors through confidence-aware rank-normalized fusion, with CTGAN augmentation and SHAP explainability, the system achieves ROC-AUC of 0.7037 without labels and 0.8316 with CTGAN augmentation, alongside a Top-7% F1 of 0.1475 on the IEEE-CIS dataset. The five novel contributions—confidence-aware fusion, rank normalization, feature-space specialization, CTGAN integration, and surrogate SHAP explainability—collectively address the core open challenges in unsupervised financial anomaly detection and provide a modular, interpretable foundation for production deployment.

ACKNOWLEDGMENT

The authors acknowledge the IEEE-CIS Fraud Detection dataset provided via Kaggle [13].

REFERENCES

[1] M. N. Uddin, "Fraudulent Transaction Detection and Prevention Using Deep Neural Networks," Ph.D. Thesis, Central Queensland University, 2024.
[2] S. Motie and B. Raahemi, "Financial Fraud Detection Using Graph Neural Networks: A Systematic Review," *Expert Systems with Applications*, 2024.
[3] S. Naseer et al., "Enhanced Network Anomaly Detection Based on Deep Neural Networks," *IEEE Access*, 2018.
[4] S. Jiang et al., "Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network," *Systems*, 2023.
[5] O. A. Bello et al., "Enhancing Cyber Financial Fraud Detection Using Deep Learning Techniques," *Int. J. Network and Communication Research*, 2022.
[6] Q. Yu et al., "Deep Learning for Cross-Border Transaction Anomaly Detection in Anti-Money Laundering Systems," 2025.
[7] A. Tanikonda et al., "Deep Learning for Anomaly Detection in E-commerce and Financial Transactions," *J. Information Systems Engineering and Management*, 2025.
[8] M. T. R. Mazumder et al., "Anomaly Detection in Financial Transactions Using Convolutional Neural Networks," *J. Economics, Finance and Accounting Studies*, 2025.
[9] D. Tikoo et al., "Anomaly Detection in Transactions using Machine Learning: A Comparative Study," in *Proc. ICIMMI*, 2024.
[10] M. B. Maneela et al., "Anomaly Detection in Transactions Using Machine Learning," *Edu-Tech Enterprise*, 2025.
[11] W. Wang and Y. Li, "Multi-layer Deep Learning-based Anomaly Detection for Digital Currency Transactions," in *Proc. AIVRID*, 2025.
[12] [Author(s)], "SEAI Research Project: CTGAN-Based Synthetic Fraud Augmentation for Unsupervised Anomaly Detection," Internal Project Report, 2026.
[13] IEEE Computational Intelligence Society, "IEEE-CIS Fraud Detection Dataset," Kaggle, 2019. [Online]. Available: https://www.kaggle.com/c/ieee-fraud-detection