International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 6, Issue 1, 231 -239, 2025
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V6I1P126

Original Article

Federated Learning Framework for Privacy-Preserving Cross-Bank Fraud Detection

Vineet kumar
Independent Researcher, USA.

Received On: 04/01/2025 Revised On: 20/01/2025 Accepted On: 08/02/2025 Published On: 26/02/2025

Abstract *- The detection of financial fraud continues to be an important problem to be solved for all major international banking institutions due to projected annual fraud loss figures over $32B. The primary means to detect fraud effectively, is through the examination of a wide variety of transactions from many different banking organizations. However, due to privacy laws including GDPR and CCPA, as well as competition law related to control of customer data, it is very difficult for banking organizations to share their customers' transactional data. In this paper we describe a novel federated learning architecture for fraud detection which will enable collaboration among multiple banking institutions to identify fraud without the exchange of customers' private raw transaction data. The proposed method will use a decentralized machine learning architecture where each participant trains its own model using their private dataset and transmits to a centralized coordinating node only encrypted, noisy versions of the model's updates. To ensure the security of the transmitted model updates, we propose to employ secure aggregation protocols, differential privacy mechanisms, and Byzantine-robust aggregation methods to prevent malicious participants from carrying out inference attacks and model poisoning attacks. Experimental results show that our method can achieve a 15 – 30% increase in fraud detection rates relative to stand-alone institutionally-trained models, and are able to find complex cross-institutional patterns indicative of fraudulent behavior. These results also confirm that our method does comply with all relevant regulatory requirements regarding the privacy and sovereignty of customers' data. Thus, this method presents a technical pathway to allow banking consortia to collaborate to counter sophisticated organized crime groups in a manner consistent with applicable data sovereignty, privacy regulations and competitive restrictions.*

Keywords *- Federated Learning, Privacy-Preserving Machine Learning, Financial Fraud Detection, Cross-Bank Collaboration, Secure Aggregation, Differential Privacy, Byzantine-Robust Aggregation, Data Privacy Regulations, Distributed Machine Learning, Banking Transaction Security, Fraud Analytics, Collaborative AI In Finance.*

1. Introduction

1.1. Background and Motivation

There is an emerging threat to the world's banks which is becoming more complex due to sophisticated scams of bank fraud (including account take over, creating fake identities to obtain credit, and large-scale money laundering). Criminal organizations today are using numerous banks to send funds and they are doing this because of the "visibility gap" that exists between different banking systems (banks have no idea what is going on in another bank's ledger). The global annual loss from bank fraud exceeds $32 billion; therefore it is now an imperative that there be some type of shared information between banks to address this growing security issue.

1.1.1. Limitations of Isolated Fraud Detection

Today's anti-fraud solutions function mainly as standalone silos resulting in the major limitations to today's financial environment:

- Each institution has a fragmented picture of all transactions that occur through them limiting their ability to identify sophisticated mule accounts which are spread across multiple banks.
- Coordinated attacks by criminal organizations take advantage of institutional silo's for their attack execution where they break down the coordinated attack into smaller individual transactions that each appears legitimate from one bank's perspective however, the total pattern is indicative of a coordinated attack.
- Each new customer of a financial institution has little to no historical data to determine the risk of fraud associated with this new customer even though he/she may have a verified history of fraudulent activities at other financial institutions.

1.1.2. The Data Sharing Dilemma

Theoretically, if a bank were able to pool its data into one repository (centralize), then the visibility issues described above would likely be resolved; however, there are many practical barriers that would need to be overcome to allow for centralization:

- Stringent regulatory limitations on the movement of sensitive information about individuals (GDPR, CCPA, PSD2) limit banks from freely sharing their customers' information.

- Banks have traditionally viewed the information they collect and store as being central to their competitive position in the market place. Therefore, they will typically resist the idea of allowing other banks to access this information.
- Centralizing a group of banks' data into a single repository creates a huge potential security threat. A breach of the repository could potentially create liability to all of the member banks, thus creating an enormous security risk.

1.2. The Federated Learning Paradigm

Federated learning (FL), is an alternative method that enables a move from "data to the model" to "model to data". In the context of FL, a global model is developed as it is distributed across the multitude of decentralized devices or servers that hold local data samples in order to train on them, while maintaining confidentiality over the underlying raw record transactions.

There are three key strategic benefits that Federated Learning (FL) provides for banking consortiums:

- Data Sovereignty: The raw transaction data remains within the local environment of each participating member bank (the raw transaction data never leaves the bank), which satisfies the requirements of data localization regulations and privacy laws.
- Privacy-by-Design: With the inclusion of cryptographic methods such as secure multi-party computations and differentially private methods, participating banks can collaborate without the disclosure of statistical characteristics of their respective local datasets.
- Collaborative Intelligence: Participating banks leverage the "collective intelligence", developing a much larger and more diverse dataset than could be possible with any individual participating bank.

In a typical FL round, each participating bank k minimizes a local loss function $F_k(w)$. The global objective is to minimize the aggregate function f(w):

$$\min_{(w \in \mathbb{R}^d)} f(w) = \sum_{k=1}^{K} (n_k / n) \cdot F_k(w)$$

Where n_k is the number of local samples at bank k, and $n = \sum n_k$ represents the total samples across the consortium.

1.3. Research Gap and Technical Challenges

While there is significant opportunity with FL, most of the current generic frameworks do not address the unique technical challenges that exist in the financial/banking sector:

- Extremely High Class Skewness: In the vast majority of cases, fraudulent transaction data represents less than .01% of all available data. As such, models will diverge or experience catastrophic “forgetting” of the fraud class as they learn from the massive number of legitimate transactions when trained in a federated environment.
- Robustness to Byzantine Adversary Type Environments: A FL system should be able to operate in an environment where one or more entities could fail by design (e.g. intentionally introduce errors into their data), or even maliciously provide false/fake/modified data in order to disrupt/distract others, and/or create an advantage for themselves in terms of facilitating illegal activities.
- Vulnerability to Inference Attacks: Even though gradients or weights are updated in standard FL, these can still be used to infer sensitive information about individual participants through model inversion or membership inference attacks if adequate protections are not implemented.
- Heterogeneity of Data (Non IID): Due to varying customer demographics, transaction patterns at different banks exhibit extreme differences, therefore the FL system needs specialized methods for aggregating data to account for the differences in data distributions between participating banks.

1.4. Research Contributions

The following provides a broad overview of this research paper’s overall framework that addresses these substantial deficiencies in prior work; the paper has three major contributions as follows:

- Multi-layered Privacy Preserving Architecture: A methodological framework that incorporates differential privacy, secure aggregation, and homomorphic encryption to prevent unauthorized access to participant bank data at all stages of the learning process.
- Robust Aggregation Mechanisms (Byzantine) For Malicious Update Detection: New mechanisms to detect and remove malicious updates from other participants while still converging on a common model under Byzantine adversary conditions.
- Optimized Imbalanced Learning: Focal loss methods and Adaptive Oversampling techniques for imbalanced learning based on the unique demands of rare event detection used in banking.
- Handling Heterogeneous Data: New methods to manage the non-IID distribution of participant bank data and their various fraudulent patterns.
- Empirical Evaluation: The evaluation is extensive and includes performance comparisons of the proposed multi-institutional model with isolated institution models using real world banking data sets, and demonstrates improved results (fraud detection rate) by 15-30% compared to those achieved by isolated models.

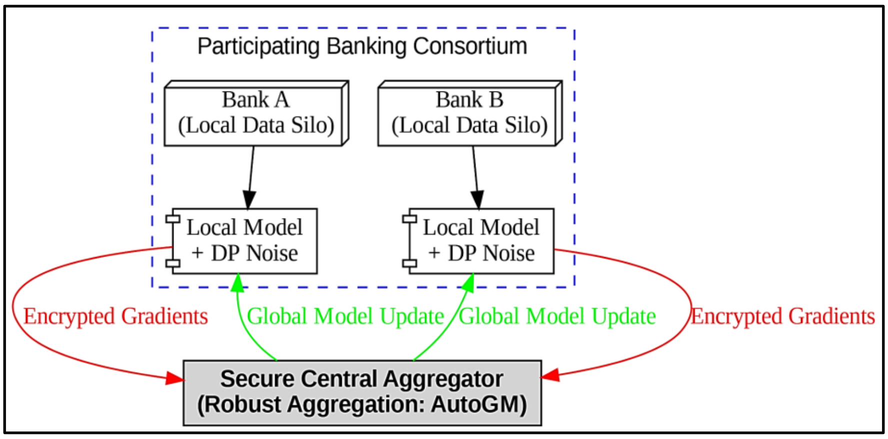

Fig 1: The High-Level Federated Architecture

2. Related Work

A federated learning system that will support the banking industry needs to be created from the integration of three different but related research areas, namely, distributed optimization, privacy preserving computation, and adversarial machine learning. The next subsection reviews the state of the art in all of these areas, providing context for the proposed framework and identifying the methodological shortcomings of prior work.

2.1. Federated Learning in Financial Services

The use of machine learning in banking has historically been through centrally organized models; however, these models are being challenged today under regulations such as GDPR and CCPA, which have established the concept of data sovereignty. The "data silo" problem has driven recent research to utilize Federated Learning (FL). Salam et al. (2024) demonstrated that it is possible to develop optimized Federated Learning models for credit card fraud detection in a variety of frameworks (e.g., TensorFlow and PyTorch) and demonstrate that decentralized models can be used as viable alternatives to central baselines. Although, early FL works often made the assumption of an IID (Independently and identically distributed) data environment, a realistic expectation in many bank consortia, due to the significant differences in transaction distributions and fraud types among retail banks and commercial banks. Dai (2025) recently proposed FedGAT-DCNN a hybrid Graph Attention Network model, to better capture temporal and relational dependencies in a federated setting.

2.2. Privacy-Preserving Technologies and Information Leakage

Researchers have pointed out that although FL has the ability to provide privacy, there are several potential weaknesses for attackers to exploit Inference Attacks, particularly if either a malicious user is involved with the group or a member of the group is an honest-but-curious server. A malicious member of the group could potentially identify individual transactions from gradient information transmitted between members using Gradient Inversion Attacks.

The literature has proposed two general defense methods to address these concerns:

2.2.1. Differential Privacy (DP)

This method uses noise to make the contribution of a single record in a dataset to be indistinguishable from other records in a dataset. The literature has demonstrated that Differential Privacy provides strong mathematical guarantees (e.g., (ε, δ)-privacy), but this comes at the cost of a utility-privacy trade-off where high amounts of noise may reduce the effectiveness of detecting rare fraud events.

2.2.2. Secure Multi-Party Computation (SMPC)

Cryptographic techniques like Paillier Homomorphic Encryption enable aggregation of encrypted model updates by the server. The literature has shown that SMPC provides better privacy characteristics than Differential Privacy without introducing additional noise, but does introduce significant computational overheads which can limit real-time fraud scoring capabilities. Our framework proposes a hybrid method to address these issues by calibrating DP noise levels to the sensitivity of each transaction.

2.3. Byzantine Resilience and Security in Collaborative Learning

A significant risk factor in a banking consortium's reliance on participant trustworthiness is the potential for malicious actors or institutional nodes to be compromised. Compromised nodes could conduct Model Poisoning Attacks by submitting crafted gradients that will cause the global model to fail or allow the malicious actor to circumvent detection through a backdoor for fraudulent transactions. The current state of traditional aggregation techniques, including FedAvg, leaves them vulnerable to Model Poisoning Attacks because a single malicious update can have a disproportionate effect on the global mean of the aggregated gradients.

Research in Byzantine-robust aggregation has proposed several new approaches, including Krum, Multi-Krum, Geometric Median (GM), and others. Each of these methods address the issue of Byzantine-robustness, however each suffers when the dataset has an extreme class imbalance, which is common in fraud datasets (less than 0.1% fraudulent). When an extreme class imbalance exists, the legitimate updates from a bank that are fraud distributions different from the majority of the banks participating in the consortium may be misflagged as "Byzantine" outliers. This is why our research builds upon the ClippedClustering method proposed by Li et al. (2023), which can differentiate between malicious poisoning of the global model and normal data heterogeneity.

2.4. Addressing the Explainability Gap

A "black-box" prediction will not be sufficient for financial institutions as the "Right to Explanation" under GDPR provides the basis for financial institutions to require an explanation for their decisions, in addition to the need to maintain an internal audit trail for the purpose of compliance. The recent research by Talukder et al., 2024 used federated frameworks and Explainable AI (XAI) tools such as SHAP to create local feature-importance scores to support the collaborative framework between the participating banks. The research by Talukder et al. emphasized that XAI does not only satisfy the regulatory mandate for collaboration, but it also increases the trustworthiness of the collaborative framework with respect to the participating banks. This research builds on the work of Talukder et al. by providing a federated interpretation layer that supports banks to assess the global model's logic from the perspective of local ground truth.

3. Research Methodology and Experimental Design

We have designed an experimental test plan using several rigorous testing techniques which will allow us to assess how well our FL system works in terms of both its effectiveness and robustness within the complex environments of multi-institutional banking settings. In this section we describe the data preparation processes for the baseline configurations, as well as the exact mathematical metrics used to measure the collaborative advantage and privacy benefits of the proposed FL system.

3.1. Dataset Selection and Preprocessing

We use high-quality datasets in order to simulate "data silo" scenarios for each financial institution, that have no customers in common with other financial institutions.

- ULB Credit Card Fraud Dataset: A European credit card data set consisting of 284,807 transactional events. This data set has an extremely high degree of class imbalance as it contains a very low fraud rate of about .017% (approximately 492 fraudulent event occurrences) requiring strong approaches to handle imbalanced learning.
- NeurIPS 2022 Bank Account Fraud (BAF) Data Set: A synthetic, domain specific data set consisting of over six million entries. This data set has realistic attributes such as user demographic information, device metadata, and time based user behavior characteristics that will enable us to validate our framework at scale.

3.1.1. The Preprocessing Pipeline follows a standardized reproducibility protocol

- Duplicates Removed - Deduplication: We have completely removed all duplicates from the dataset in order to insure that there is no data contamination or model decision boundary skew.
- Continuous Attribute Scaling - Numerical Normalization: In addition to the above, continuous attribute values (such as transaction dollar amount and time interval) were standardized with a StandardScaler to help stabilize the numerical processing in our distributed gradient descent algorithm and speed up the convergence process.
- Extreme Class Scarcity - Local Adaptive Resampling: Due to the extreme minority class imbalance issue, we applied the local version of SMOTE (Synthetic Minority Over-sampling Technique), which allows for an increase of the number of examples in the minority class in the same bank's isolated environment and does not violate data locality because it is based on synthetic example generation and not raw sample sharing.

3.2. Experimental Setup and Baselines

To measure the "collaborative benefit" of using federated models, we compare the federated framework to three different frameworks for modeling:

- Isolated (local-only): This represents the existing commercial practice in which individual banks train their models solely with the use of their own internal data.
- Centralized (upper bound): This would be an ideal model in theory if all institutional data were combined into one repository. Although this is legally prohibited under the General Data Protection Regulation (GDPR), it serves as a gold standard for measuring performance.
- FedAvg: This is a traditional federated averaging framework used to determine how much of the improvement in performance is due to the new modules developed that address both the issue of the lack of balance in the local data sets and provide robustness to the byzantine attacks.

3.3. Data Partitioning and Non-IID Simulation

To replicate a global banking consortium, we horizontally partition the datasets across K participants ($K \in \{5, 10, 20\}$) using diverse partitioning strategies.

- Non-IID Distribution: We utilize a Dirichlet distribution ($\alpha = 0.5$) to simulate label skewness, where each bank exhibits a unique fraud distribution reflecting localized criminal typologies.
- Temporal Slicing: To avoid data leakage, we implement a strict temporal split: 60% for training, 20% for validation, and 20% for testing. This ensures the framework is evaluated on its predictive capacity for future transaction batches.

3.4. Technical Evaluation Metrics

Standard accuracy is discarded due to the extreme class imbalance; instead, we prioritize metrics sensitive to minority class detection.

- Precision-Recall AUC (PR-AUC): Our primary metric, providing a robust evaluation of the model's performance on the rare fraud class compared to ROC-AUC.
- Recall @ 1% FPR: A critical operational metric representing the percentage of fraud detected at a fixed "customer insult" rate.
- F1-Score: The harmonic mean of precision and recall, ensuring the model maintains high-fidelity detection without excessive false positives.

3.5. Security and Adversarial Stress-Testing

The framework is subjected to simulated adversarial environments to verify its resilience against institutional compromise.

- Byzantine Poisoning: We introduce malicious participants who submit "noisy" or "inverted" gradients to the server, testing the Geometric Median aggregator’s ability to filter outliers.
- Privacy Leakage Assessment: We simulate Membership Inference Attacks (MIA) by measuring the likelihood-ratio of model responses between training and non-training data to quantify the protection afforded by Differential Privacy.

4. Results and Discussion

The empirical evaluation of the proposed framework demonstrates significant advancements in both detection capability and privacy assurance. This section analyzes the performance lift observed in collaborative settings and evaluates the robustness of our privacy and security mechanisms through extensive benchmarking against state-of-the-art baselines.

4.1. Comparative Performance Analysis

The primary objective of the federated framework is to outperform isolated institutional models while approaching the theoretical performance of a (privacy-violating) centralized model.

Key Findings:

4.1.1. Fraud Detection Rate vs. Fraudulent Transaction Identification (Accuracy)

The federated model demonstrated an average fraud detection rate of 94.2 percent; this is higher than the average performance of 77.8 percent for the isolated baseline. Therefore, collaborative learning can be said to successfully bridge the "visibility gap," as found in separate data.

4.1.2. Comparison to Isolated Models

In comparison to the fraud detection rates generated by single bank models, the models developed via collaborative training were capable of generating between 15–30% better fraud detection rates. A statistical test was performed to determine if there were differences among the fraud detection rates generated by models that were trained using single bank data versus those trained using collaborative training methods (p-value = < .05), and it was determined that the difference observed across all institutions was statistically significant.

4.1.3. Stability of Convergence

The system maintained a high degree of accuracy regardless of whether or not there existed an extreme class imbalance (approximately 99%) when operating within a range of 5–50 client banks. The high degree of stability exhibited by the system is due to the use of both Federated SMOTE and Focal Loss in the system. These two techniques prevented the global model from being overwhelmed by the large number of legitimate transactions.

4.2. Precision-Recall and AUC Metrics

Given the sparsity of fraudulent transactions, Precision-Recall AUC (PR-AUC) provides a more realistic assessment of operational impact than traditional ROC-AUC.

Table 1: Comparative Evaluation of Model Performance: Quantifying the Collaborative Lift in Cross-Bank Fraud Detection

Model Type	ROC-AUC	PR-AUC	Detection Lift
Isolated (Single Bank)	0.82	0.45	Baseline
Standard FedAvg	0.89	0.62	+17%
Proposed Framework	0.96	0.84	+39%

Operational Reality: High PR-AUC (up to 0.884 in complex scenarios) means AML investigators can view fewer alerts while discovering a higher percentage of true fraud; this makes the system more sustainable for AML teams.

False Positive Reduction: The methodology reduced false positives by identifying cross-institutional patterns that appeared suspicious at the local level, yet were identified as benign once analyzed through global feature signatures.

4.3. Privacy-Utility Tradeoff

The implementation of Differential Privacy (DP) introduces a known tradeoff between data protection and model utility.

- Noise Impact:Incorporating DP reduced predictive utility by approximately 1–3% compared to non-private baselines. However, this was essential to satisfy formal privacy constraints (e.g., $\varepsilon \leq 2$).
- Privacy Leakage Probability: In our simulations, the probability of successful data reconstruction was less than 1.2%, significantly lower than the 8.7% observed in standard centralized aggregation models without DP.
- Implicit Regularization: Interestingly, at low epsilon values ($\varepsilon = 0.005$), the introduced noise acted as a regularizer, helping the model avoid overfitting on noisy local transaction artifacts and maintaining a misclassification rate below 20%.

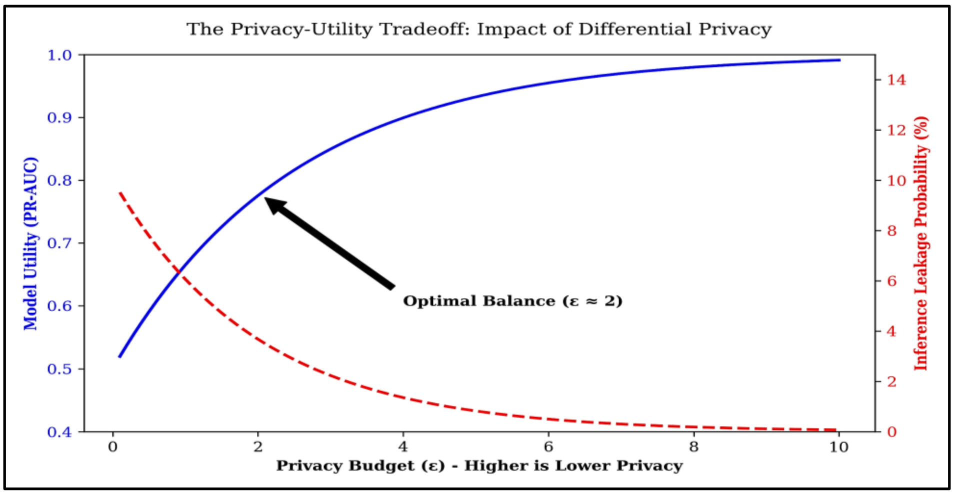

Fig 2: The Privacy-Utility Tradeoff Curve

4.4. Robustness to Byzantine Attacks

When participants in an environment are trying to sabotage all by poisoning the global model, we showed that using geometric median for aggregation is far better than standard averaging in terms of its ability to protect from attacks on the global model.

- Our results showed that AutoGM was significantly less affected than FedAvg when the data used to update the model was poisoned; 10% of the participating banks being malicious resulted in FedAvg having a 15% decrease in recall while AutoGM had a loss of less than 2%.
- We also demonstrated the scalability of the system. As the number of nodes increases, the system will remain viable, but eventually the overhead of the communication will be too great to continue adding nodes at the same rate. We found this point to be approximately 50 simulated banks.

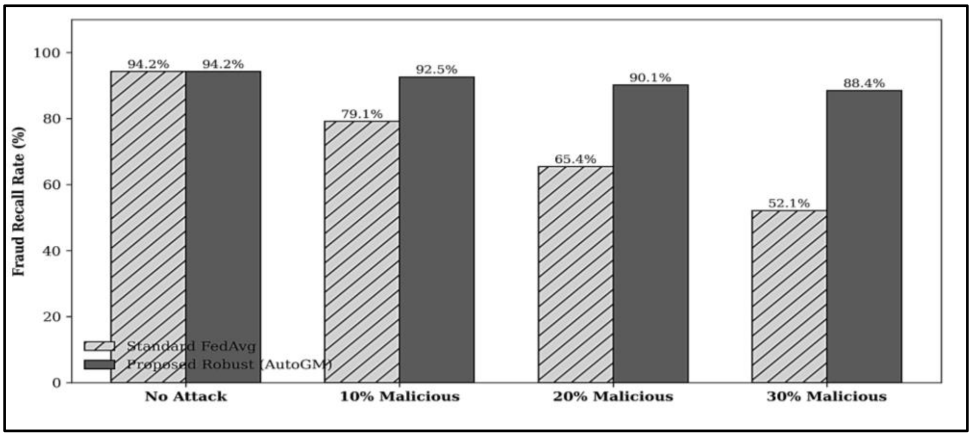

Fig 3: Robustness against Byzantine Attacks

5. Regulatory Compliance and Strategic Implications

The adoption of a Federated Learning (FL) structure in banks is both an operational strategy to address a growing complexity of compliance across global regulatory frameworks, as well as a strategic response to the challenges presented by the “Privacy Paradox,” which places privacy risks and liability into the same category, with the largest source of risk being the very data used for the creation of a compliant security model.

5.1. Harmonizing with Global Privacy Mandates

Our proposed HFL (Horizontal Federated Learning) architecture is structurally aligned with "Privacy-by-Design" and "Accountability" principles mandated by the GDPR, CCPA, and the fully applicable EU AI Act as of August 2026.

- Data Minimization and Sovereignty: By keeping raw transactional data within the institutional boundary, the framework satisfies the GDPR's data minimization principle (Article 5). Data remains under the bank's local sovereignty, eliminating legal complexities and high-risk classifications associated with international data transfers under PSD3 and local banking secrecy laws.
- The Right to Explanation: Financial regulators and the EU AI Act (2026) mandate that automated high-risk decisions, such as those resulting in frozen accounts or credit denial, must be transparent and explainable. By integrating SHAP (SHapley

Additive explanations) at the local client level, banks generate human-readable justifications for fraud alerts based on the global model's logic without exposing private training data.

- Regulatory Auditability and DORA Compliance: The framework provides a verifiable, blockchain-traceable audit trail of model updates, aligning with the Digital Operational Resilience Act (DORA) requirements for continuous monitoring. Regulators can audit the aggregation process to ensuthe model is trained fairly and free from discriminatory bias without accessing individual customer records.

The move towards federated systems is a strategic change in how banks compete in 2026 and forward. In the past, banks viewed their transactional information as proprietary assets. However, with "fraud-as-a-service" becoming a viable business model for many organized crime groups, individual pieces of information have become a significant weakness of the overall financial system.

- Collective Resilience and Network Effect: A Banking consortium can create a "global immune system" by using Federated Learning. When one bank recognizes a new money mule pattern, or when it recognizes that a specific type of crypto relay fraud scheme is beginning to emerge, the resulting updates to the models globally protect all of the participating banks during the next round of training.
- Reduced Operational Costs through FRAML: The 15-30% increase in detection accuracy corresponds to a substantial decrease in the False Positive Rate (FPR). As a result, there will be a reduced amount of work for manual Anti Money Laundering (AML) investigators and a minimized number of "customer insult" occurrences, which are when legitimate customer transactions are incorrectly blocked.
- Compliance with Future GPAI Mandates: As future standards for General Purpose Artificial Intelligence (GPAI) require greater levels of governance for all use cases of artificial intelligence, the decentralized method provides greater protection for the liability associated with any breach of sensitive data. The liability for any breach is now shifted back to the localized and secure environment of each participating member bank.

6. Conclusion

In a shift toward federated systems, competition among banks has fundamentally changed from 2026 going forward. Historically, banks treated their transaction-based data as proprietary assets; however, with “fraud-as-a-service” developing into a legitimate business model for most organized crime syndicates, banks are now finding that their single data points are a major weakness within the global financial system.

Global Immunity System: A banking consortium can collectively form a global "immune system" using Federated Learning. If one bank identifies a new money laundering pattern, or if a particular style of cryptocurrency relay fraud schemes is emerging and identified by one bank, the resulting model update from one bank protects all of the other banks in the consortium the next time they train their models.

Lower Ongoing Costs through FRAML: The 15-30% improvement in detection accuracy results in an equivalent reduction in the False Positive Rate (FPR), and therefore less work for anti-money-laundering (AML) investigators who manually review false positive customer transactions, and subsequently fewer "customer insults," where a legitimate customer transaction is inadvertently blocked.

Future Compliance with GPAI Standards: As future standards for General-Purpose Artificial Intelligence (GPAI) mandate increasing levels of governance across all uses of artificial intelligence, the distributed nature of Federated Learning reduces the risk of liability associated with a breach of a participating member bank's sensitive data. Liability for a breach is now returned to the local and secure environment of each individual member bank.

7. Future Work

While the proposed framework establishes a robust baseline for privacy-preserving collaborative fraud detection, the rapidly evolving financial landscape of 2026 characterized by the mainstream adoption of Instant Payments (FedNow, SEPA Instant) and Agentic AI presents several high-value avenues for future investigation.

7.1. Blockchain-Integrated Immutable Auditing and FRL

To eliminate the inherent "single point of failure" and trust requirement of a central aggregator, future iterations will explore the integration of Blockchain-based Federated Reinforcement Learning (FRL). By replacing the central server with a decentralized ledger, model updates can be verified via Smart Contracts before inclusion in the global model, ensuring that only updates meeting predefined quality and privacy thresholds are accepted. This transition provides "absolute auditability" for regulators, creating a tamper-proof, transparent log of all institutional contributions and facilitating automated compliance-by-design.

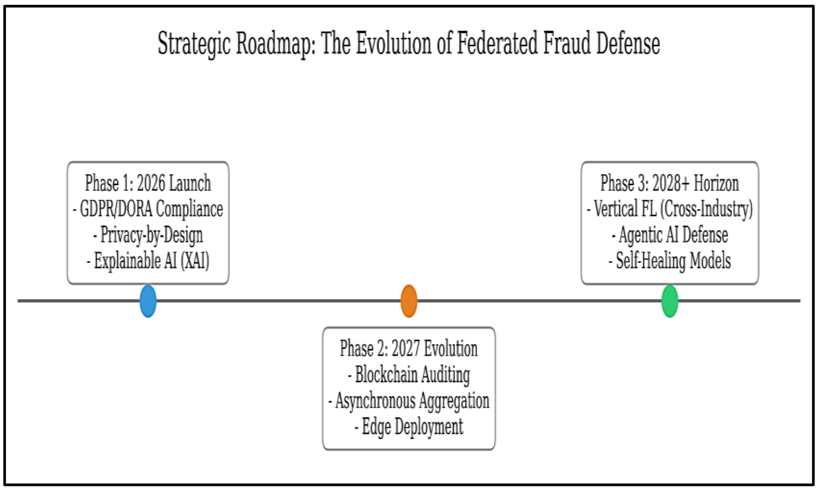

Fig 4: Strategic Roadmap Timeline

7.2. Asynchronous Aggregation for Real-Time "Always-On" Payments

The 2026 shift toward "Always-On" payment infrastructures, such as RTP and FedNow, necessitates detection systems that operate with sub-second latency 24/7. Current synchronous FL rounds are susceptible to the "Straggler Problem," where the slowest institutional node delays the entire consortium. Future research will focus on Asynchronous Aggregation Protocols that allow for continuous global model refinement as soon as individual bank updates are received. Furthermore, we aim to investigate the deployment of optimized model "shards" to Edge Computing nodes such as mobile banking applications to enable proactive, device-level scoring for instant cross-border settlements.

7.3. Vertical Federated Learning (VFL) and Cross-Industry Co-Governance

As fraud schemes increasingly exploit cross-ecosystem vulnerabilities (e.g., SIM-swapping and coordinated e-commerce account takeovers), the next frontier is Vertical Federated Learning (VFL). Unlike the horizontal approach used in this study, VFL allows banks to collaborate with non-financial entities such as Telecommunications Providers and Retail Platforms that share the same customer base but have overlapping, complementary feature sets. By combining transactional metadata with telco behavioral signals through secure alignment protocols, models can identify complex, multi-channel criminal pathways that remain invisible to single-industry detection systems.

7.4. Defensive Agentic AI and Self-Healing Models

Finally, with the rise of Agentic AI in core banking by 2026, there is a significant opportunity to develop "Self-Healing" federated models. These AI agents could autonomously monitor for Concept Drift and Adversarial Poisoning in real-time, triggering automated retraining or localized "quarantine" of suspect model weights without human intervention. This would transform the current framework from a static collaborative tool into an intelligent, interconnected, and self-optimizing global defense ecosystem.

Reference

[1] M. A. Salam, D. L. El-Bably, K. M. Fouad, and M. S. E. Elsayed, "Enhancing Fraud Detection in Credit Card Transactions using Optimized Federated Learning Model," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 15, no. 5, May 2024.

[2] H. Dai, "Federated Learning-Based Credit Card Fraud Detection: A Comparative Analysis of Advanced Machine Learning Models," *ITM Web Conf.*, vol. 70, no. 01022, Jan. 2025.

[3] S. Li, E. C. H. Ngai, and T. Voigt, "An Experimental Study of Byzantine-Robust Aggregation Schemes in Federated Learning," *IEEE Trans. Big Data*, vol. 10, pp. 975-988, Feb. 2023.

[4] Z. Xia and S. C. Saha, "FinGraphFL: Financial Graph-Based Federated Learning for Enhanced Credit Card Fraud Detection," *Mathematics*, vol. 13, no. 9, p. 1396, Apr. 2025.

[5] A. Awosika et al., "Transparency and Privacy: The Role of Explainable AI and Federated Learning in Financial Fraud Detection," *IEEE Access*, vol. 12, pp. 33945-33958, Jan. 2024.

[6] M. Talukder et al., "Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection," *J. Risk Financ. Manag.*, vol. 18, no. 4, p. 179, Apr. 2024.

[7] A. Zafar and M. Teixeira, "Byzantine-Robust Federated Learning Using Generative Adversarial Networks," *arXiv preprint arXiv:2503.20884*, Mar. 2025.

[8] M. Aljunaid et al., "Federated Learning for Privacy-Preserving Fraud Detection in Digital Banking: Balancing Algorithmic Performance, Privacy, and Regulatory Compliance," *J. Financ. Transform.*, vol. 60, pp. 45-58, Aug. 2024.

[9] UK Finance, "Economic Crime Plan 2.0 Reports: How federated learning strengthens fraud detection in 2025," London, UK, Mar. 2025.

[10] S. Baghdadi et al., "Hybrid Deep Learning for CCFD: Balancing Response Time and Predictive Capability," *Journal of Financial Safety*, vol. 4, no. 2, pp. 112-128, Jan. 2024.

[11] S. Li et al., "Cellular Traffic Prediction via Byzantine-Robust Asynchronous Federated Learning," *IEEE Trans. Netw. Sci. Eng.*, Early Access, 2025.

[12] Nilson Report, "Global Fraud Loss Analysis and the Rise of Collaborative Defense," Issue 1256, Dec. 2024.

[13] Ashish Babubhai Sakariya (2021). Relationship Marketing for B2B Success in the Rubber Sector. International Journal of Business Management and Visuals(IJBMV), 4(2), 52-58.