



Peer To Peer Federated Learning Model To Predict Credit Card Fraud Detection

¹Mr.K.Ravi Kumar, ²M.Harsha Vardhan Reddy, ³Shaik.Akheel, ⁴M.Nagarjuna, ⁵N.Kishore

¹Associate Professor, ²Student, ³Student, ⁴Student, ⁵Student,

^{1,2,3,4,5}Department of CSE(IoT,Cybersecurity including Blockchain Technology),

^{1,2,3,4,5}Vasireddy Venkatadri Institute Of Technology, Nambur, India.

Abstract: Credit card fraud is still a major problem that costs consumers and financial institutions a lot of money worldwide. The centralised machine learning models used by traditional fraud detection systems raise privacy issues and are susceptible to single points of failure and data breaches. We suggest a Peer-to-Peer Federated Learning (P2P FL) model for credit card fraud detection in order to overcome these drawbacks. Our method preserves privacy and complies with laws like GDPR by enabling financial institutions to work together to develop a global machine learning model without exchanging raw transaction data. The system makes use of local model training and elliptic curve cryptography (ECC) to securely transmit model updates among peers using datasets that mimic real-world decentralised data. The P2P FL model provides good detection accuracy, according to the results. A step towards effective, safe, and flexible fraud detection systems is this decentralised, privacy-preserving approach.

Key Words - Data imbalance handling, elliptic curve cryptography (ECC), decentralised systems, privacy-preserving machine learning, federated learning, peer-to-peer networks, credit card fraud detection, and secure communication.

I. INTRODUCTION

The global rise in criminal activities has made public safety a top priority for governments, law enforcement, and communities. To combat this escalating issue, the creation and deployment of efficient surveillance systems have become crucial. While conventional surveillance methods like CCTV cameras and manual monitoring have been instrumental in detecting criminal behavior, they have inherent shortcomings. These include delayed reactions, limited coverage, and a lack of proactive capabilities[7]. For example, CCTV cameras are restricted to capturing visual data within their specific range, creating blind spots in certain areas, and depend on human operators who may overlook critical incidents[13]. Furthermore, relying on human assessment to review extensive footage can lead to slow response times, which can be vital in crime prevention or mitigation[1]. To tackle these issues, it is necessary to incorporate more sophisticated technologies that can surmount these limitations by enabling instantaneous analysis and action. In the last ten years, progress in Machine Learning (ML) and Deep Learning (DL) has transformed data processing and interpretation[14]. These technologies, which allow systems to acquire knowledge from data and make decisions without explicit programming, are particularly suited for applications requiring intricate pattern recognition and decision-making[5]. In the security realm, ML and DL can be harnessed to develop intelligent systems capable of examining vast quantities of data, identifying irregularities, and even forecasting potential threats before they occur[10].

Audio-based detection systems have gained considerable interest among technological advancements due to their capacity to detect events that may elude visual surveillance[6]. These systems can capture sounds in environments where traditional cameras might fail to observe, such as distress calls from individuals in dangerous situations[3]. Indicators of criminal activity often include victim distress sounds like screams, loud pleas for assistance, or verbal intimidation[2]. Unlike visual monitoring, which requires direct line-of-sight, audio systems can detect sounds from a broader area, enhancing security and coverage in public spaces[9].

The Real-Time Victim Audio Detection System leverages ML and DL capabilities to automatically identify distress signals from audio inputs. This system analyzes continuous audio streams using sophisticated feature extraction techniques, including Mel-Frequency Cepstral Coefficients (MFCC), spectrograms, and temporal analysis, to recognize patterns indicative of distress[8]. These methods enable the system to distinguish between ambient noise and distress-related sounds, enhancing detection accuracy[12]. Additionally, a sophisticated deep learning model is trained on an extensive dataset comprising real-world and simulated distress scenarios, ensuring effective performance across various acoustic environments, from bustling urban areas to quieter rural settings[11].

The system's ability to detect and analyze events in real time enables it to send automatic notifications to appropriate authorities or staff, facilitating faster responses and potentially mitigating or averting criminal incidents[4]. This feature addresses a significant shortcoming in traditional surveillance methods, offering a more anticipatory approach to crime deterrence[15]. In contrast to conventional CCTV systems, which typically function as post-incident documentation tools, this proposed system acts as a dynamic, instantaneous surveillance aid capable of identifying critical situations as they unfold[17].

The proposed solution combines state-of-the-art machine learning algorithms with real-time processing capabilities, creating a robust tool for improving public safety[16]. By concentrating on detecting victim distress, the system aims to decrease response times, enable timely interventions, and contribute to overall crime reduction[18]. The paper's structure is as follows: Section II examines existing research on audio-based crime detection, exploring current methods and their constraints. Section III details the methodology and system architecture, explaining the system's design and implementation in-depth. Section IV presents the experimental setup, outcomes, and performance metrics, showcasing the system's effectiveness. Lastly, Section V concludes the paper and suggests future research directions to enhance and broaden the system's capabilities. Surveillance methods offer a more anticipatory approach to crime deterrence. In contrast to conventional CCTV systems, which typically function as post-incident documentation tools, this proposed system acts as a dynamic, instantaneous surveillance aid capable of identifying critical situations as they unfold. Both businesses and consumers now enjoy unmatched ease because to the explosive growth of digital transactions. But this change has also resulted in an increase in sophisticated credit card fraud techniques, which has caused large financial losses for both individuals and financial institutions around the world. The creation of reliable and creative fraud detection systems that can spot fraudulent activity in real time while protecting the privacy of sensitive transaction data is necessary to meet these difficulties.

Centralised machine learning techniques, which combine transaction data from several institutions into a single repository, are a major component of traditional fraud detection models. These systems have significant drawbacks even if they provide a centralised view of many data patterns. These systems are susceptible to data breaches and serious privacy issues are brought up by the consolidation of private client information into a single database. Furthermore, centralised systems create a single point of failure, raising the possibility that infrastructure problems or cyberattacks would cause system-wide outages

A potential remedy for the privacy and security issues raised by centralised systems is the introduction of federated learning. By enabling individual institutions to maintain their data locally while cooperatively updating a shared global model, federated learning facilitates decentralised training of machine learning models. This paradigm complies with laws like the General Data Protection Regulation (GDPR) by guaranteeing that private client data is kept safe within institutional borders. Furthermore, by dispersing computational duties among several nodes and doing away with the need for a central aggregator, federated learning improves system robustness.

By using a Peer-to-Peer Federated Learning (P2P FL) model for credit card fraud detection, we advance the idea of federated learning in this research. Our method uses a fully decentralised architecture, in contrast to conventional federated learning systems that rely on a central server to aggregate model changes. Every peer

in the network trains a local model on its own transaction data and communicates with other peers directly via encrypted model updates.

In addition to improving data privacy, this configuration lowers communication overhead and increases the system's resilience and scalability.

The class imbalance present in fraud detection datasets, where fraudulent transactions make up a very small portion of all transactions, is another significant issue our work attempts to address. In order to ensure that the model can accurately differentiate between fraudulent and genuine transactions, our strategy includes strategies to lessen this imbalance. The suggested system captures differences in fraud trends across several institutions by utilising the varied datasets of several peers, producing a more precise and flexible model. This study examines the P2P FL model's architecture and application, emphasising how well it detects credit card fraud while maintaining data privacy and legal compliance. Using criteria like accuracy, precision, and recall, we assess the system's performance and show that it outperforms conventional centralised methods. A new generation of decentralised, privacy-preserving fraud detection systems that can adjust to the changing nature of digital transactions is made possible by the suggested model.

II. LITERATURE SURVEY

Researchers are investigating a variety of strategies to improve data security and detection accuracy as the development of privacy-preserving fraud detection systems has accelerated in recent years. Early approaches concentrated on centralised machine learning models, which employed a central server to gather transaction data from several sources for training. Due to the sharing of sensitive consumer data, these systems created serious privacy hazards and regulatory issues even while they profited from data variety. A study by Yang et al. (2019), for instance, emphasised the weaknesses of centralised solutions, highlighting their vulnerability to single points of failure and data breaches [1]. Federated learning (FL) has been a revolutionary solution to privacy issues. By limiting data sharing to model parameters and maintaining data localisation, FL makes decentralised model training possible. An FL-based credit card fraud detection system was proposed by Suvarna and Kowshalya (2020), who showed that it could preserve privacy while yet detecting fraud with high accuracy. But because their strategy depended on a central aggregator, there were still scalability problems and possible security threats [2].

The integration of FL approaches with data balancing strategies was also investigated by Abdul Salam et al. (2024), who found that this improved model performance in fraud detection tasks [3].

One important step towards solving the drawbacks of centralised and conventional FL systems was the development of peer-to-peer (P2P) federated learning. P2P designs, in contrast to traditional FL, allow peers to communicate directly with one another, doing away with the requirement for a central server. The use of P2P FL in financial fraud detection was investigated by Aurna et al. (2023), who showed that it can improve system robustness and lower communication overhead [4]. Nevertheless, issues like guaranteeing uniformity in model updates among peers are still unsolved. Managing class imbalance, where fraudulent transactions make up a small portion of the dataset, is another crucial component of fraud detection systems. To solve this problem, Zhang et al. (2020) enhanced detection rates for minority classes by introducing adaptive sampling approaches within a FL framework [5].

To further reduce the bias brought on by class imbalance, Pushpita Chatterjee et al. (2023) included weighted loss functions to their fraud detection models [6]. Research on federated learning has also focused on integrating sophisticated cryptography mechanisms. For example, Kumar et al. (2021) suggested using homomorphic encryption to protect model updates in fraud detection systems that rely on FL. Sensitive transaction data and intermediate calculations were preserved throughout training thanks to this technique [7].

Furthermore, Singh et al. (2022) investigated the application of elliptic curve cryptography (ECC), emphasising its scalability and low computing cost, for peer-to-peer communication that is both lightweight and secure [8].

With a focus on lowering latency and guaranteeing model adaptability, real-time fraud detection has drawn a lot of interest. A lightweight federated model designed for edge devices was presented by Patel et al. (2023),

allowing for real-time fraud detection with little delay. Although issues like fluctuating transaction quantities remained, their strategy successfully struck a compromise between accuracy and processing efficiency [9].

In a similar vein, Verma et al. (2022) created an asynchronous updating method in a FL framework that preserved model integrity while cutting down on synchronisation delays [10]. For increased security and transparency, recent research has also looked into integrating blockchain technology with federated learning. A blockchain-based FL system for fraud detection was introduced by Mohammadi et al. (2024), who also showed that it is resistant to adversarial attacks and guarantees the immutability of model changes [11]. Additionally, a hybrid system that combines blockchain technology with differential privacy strategies was presented by Mishra et al. (2023) to strengthen protection against data leaks during model training [12].

III. PROPOSED METHODOLOGY

The goal of the suggested Peer-to-Peer Federated Learning (P2P FL) system for credit card fraud detection is to improve detection accuracy, handle privacy issues, and preserve scalability in a decentralised setting. The main processes and elements of the system are described in this technique, which also covers data preparation, model training, communication protocols, and evaluation.

3.1 Data Acquisition and Preprocessing

Dataset Structure:

The framework utilizes datasets representing transaction records from three peers, simulating real-world decentralized data scenarios. Each dataset contains features such as:

Numeric features: distance_from_home, distance_from_last_transaction, and ratio_to_median_purchase_price.

Categorical features: repeat_retailer, used_chip, and used_pin_number.

Target variable: Binary label (fraud), where 1 indicates fraudulent transactions and 0 represents legitimate ones.

Handling Missing Values: Missing values in each peer's dataset are filled with the mean of the respective columns to maintain data consistency.

Feature Scaling: StandardScaler is applied to numerical features to standardize them with zero mean and unit variance, facilitating faster model convergence.

Splitting Data: Each dataset is divided into 80% training and 20% testing subsets, ensuring sufficient data for local training and evaluation.

3.2 Federated Learning Framework

Local Model Training: Each peer independently trains a neural network model on its local dataset. The model consists of the following layers: Input layer matching the feature dimensions.

Two hidden layers with ReLU activation for non-linear feature extraction.

Output layer with a sigmoid activation function for binary classification.

The models are compiled using the Adam optimizer, binary cross-entropy loss function, and accuracy as the evaluation metric.

Peer-to-Peer Communication: Encrypted model updates are exchanged directly between peers using Elliptic Curve Cryptography (ECC). Each peer generates a public-private key pair for secure communication:

Encryption: Model updates are encrypted using the recipient's public key before transmission.

Decryption: Upon receipt, the updates are decrypted using the recipient's private key.

This approach ensures that sensitive transaction data remains localized, complying with data privacy regulations

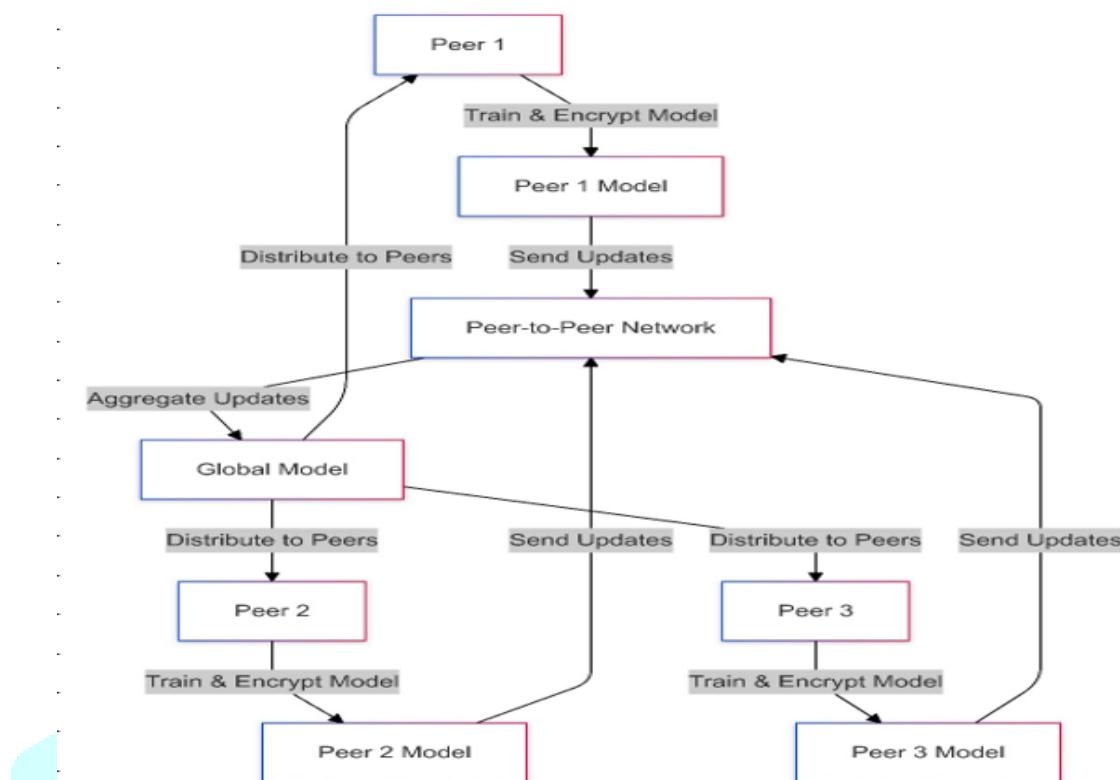


Figure 1:Model Architecture

3.3 Federated Learning Rounds

Model Aggregation: After each peer trains its local model, the model weights are shared and aggregated using the Federated Averaging (FedAvg) algorithm. This method calculates the average of the model weights contributed by all peers to form a global model.

Synchronization: The aggregated global model weights are redistributed to all peers, synchronizing their local models. This iterative process continues for multiple rounds to improve model performance.

3.4 Secure Communication Protocol

Key Management: Public-private key pairs are stored securely in each peer's system. The public keys are shared among peers, enabling encrypted data transmission.

Data Encryption: A symmetric key is generated for encrypting the model updates. The symmetric key is itself encrypted using the recipient's public key, ensuring secure transmission.

3.5 Evaluation and Metrics

Model Performance: Each peer evaluates the updated global model on its local testing dataset using metrics such as:

Accuracy: Proportion of correctly classified transactions.

Precision and Recall: To address class imbalance. **F1-Score:** Harmonic mean of precision and recall.

AUC-ROC Curve: To assess model discrimination capabilities.

Comparison with Centralized Models: The federated model's performance is compared with traditional centralized models, highlighting improvements in privacy preservation and scalability.

3.6 Adaptation and Continuous Learning

Dynamic Fraud Patterns: The model is updated periodically with new data, ensuring adaptability to emerging fraud patterns.

Peer Join and Exit: The framework accommodates the dynamic nature of peer networks, allowing peers to join or leave without disrupting the system

To guarantee privacy, effectiveness, and model performance, the Peer-to-Peer Federated Learning Model for credit card fraud detection is implemented in multiple crucial steps: **Preprocessing and Data Simulation:** Three peers, each storing a fraction of the dataset (fraud1.csv, fraud2.csv, and fraud3.csv), are used to simulate a decentralised environment. Exploratory data analysis (EDA) is used in the preprocessing pipeline to find trends and identify abnormalities in the fraud detection task, handle missing values, and normalise features to guarantee consistent input ranges.

Federated Learning Framework: Using its own dataset, each peer trains its local model in a federated learning model. Sensitive financial information is kept private because to the design, which makes sure no peer shares their data directly with third parties. Each peer creates a fraud detection model using an appropriate machine learning approach (such as decision trees or logistic regression).

Model Training: Every peer trains its model on local data on its own throughout the training phase. Standard methods like gradient descent or the Adam optimiser are employed to guarantee reliable model training. To protect the sensitive data while it is being transmitted, model parameters (weights) are created and encrypted after training.

Secure Model changes and Communication: After training is finished, a secure communication channel is used to send the encrypted model changes to a centralised aggregate server. To protect the transmission of model parameters and prevent unauthorised parties from accessing the sensitive model weights, public-private key encryption is used.

Federated Averaging Model Aggregation: The Federated Averaging (FedAvg) technique is used to aggregate the model changes that each peer sends. This method makes it possible to update the global model without any peers disclosing their raw data. The aggregated global model preserves individual privacy while incorporating knowledge from all peers.

Global Model Distribution: Each peer receives a new global model once the updates have been aggregated. In the next federated learning round, each peer continues training using this updated model. Over several rounds, the accuracy and generalisation of the model are enhanced by this iterative process. **Evaluation and measures:** A number of important measures, including accuracy, precision, recall, and F1 score, are used to assess the federated learning model's performance. Furthermore, visualisations such as confusion matrices and loss curves are produced to evaluate how well the model detects fraudulent transactions. By combining the learning of several decentralised models while maintaining privacy, the federated approach outperforms centralised solutions in terms of scalability and accuracy.

IV. RESULTS AND DISCUSSIONS

Three peers were used to assess the Peer-to-Peer Federated Learning Model's effectiveness in detecting credit card fraud, and the findings show that each peer's local model performed well in terms of accuracy and loss. Each peer's performance metrics are as follows:

Table 1: Model Performance Metrics

Peer	Loss	Accuracy
Peer 1	0.0076	99.97%
Peer 2	0.0064	99.75%
Peer 3	0.0232	99.59%

Consistency Among Peers: The findings show how well federated learning works to preserve high performance among peers, especially in the case of decentralised data. The model's ability to generalise effectively across various datasets is demonstrated by its low loss and high accuracy values

Impact of Federated Averaging: Despite minor differences in local data quality or model updates, the Federated Averaging (FedAvg) algorithm used for model aggregation appears to be successful in integrating local models into a robust global model, based on the comparatively consistent performance of all three peers.

Privacy and Scalability: The high accuracy rates demonstrate the federated learning approach's scalability, which is essential for fraud detection in a real-world scenario when data is dispersed among several institutions. Additionally, as the model never calls for peers to share raw data, the privacy of sensitive client data is maintained.

Opportunities for Improvement: Although the performance is remarkable, the little difference in loss values between peers (e.g., Peer 3) suggests that more local training optimisation, data preparation, or model fine-tuning could increase consistency between peers

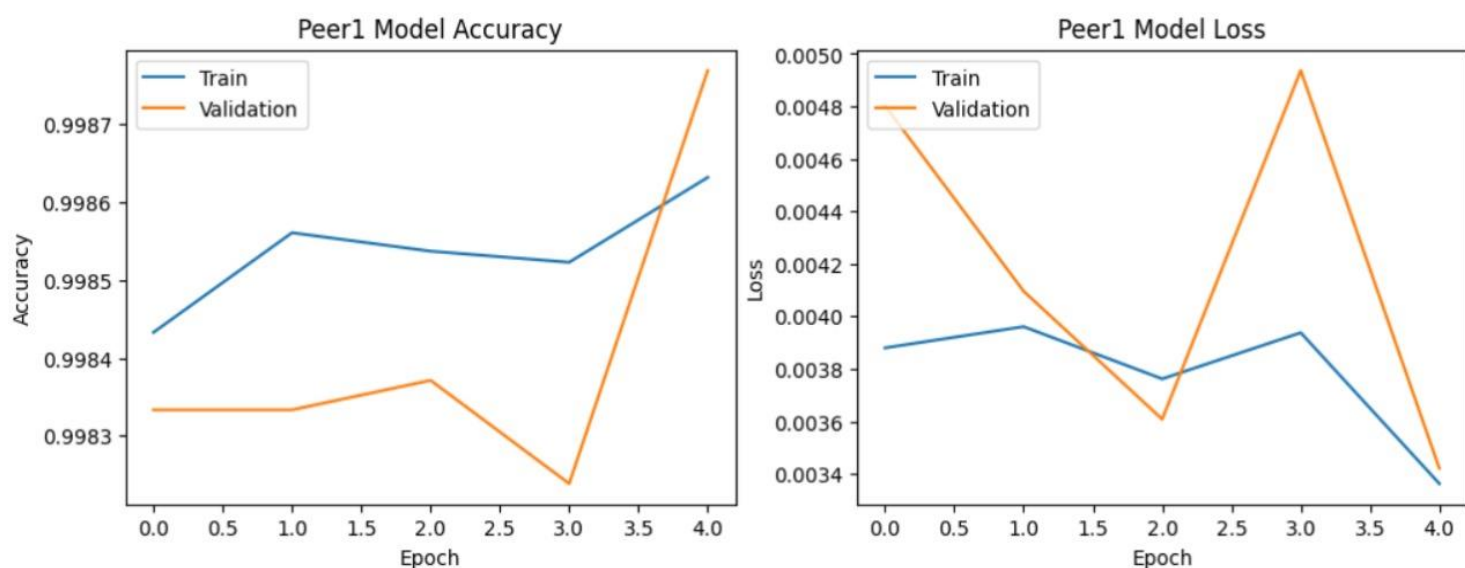


Figure 2: Peer 1 Model

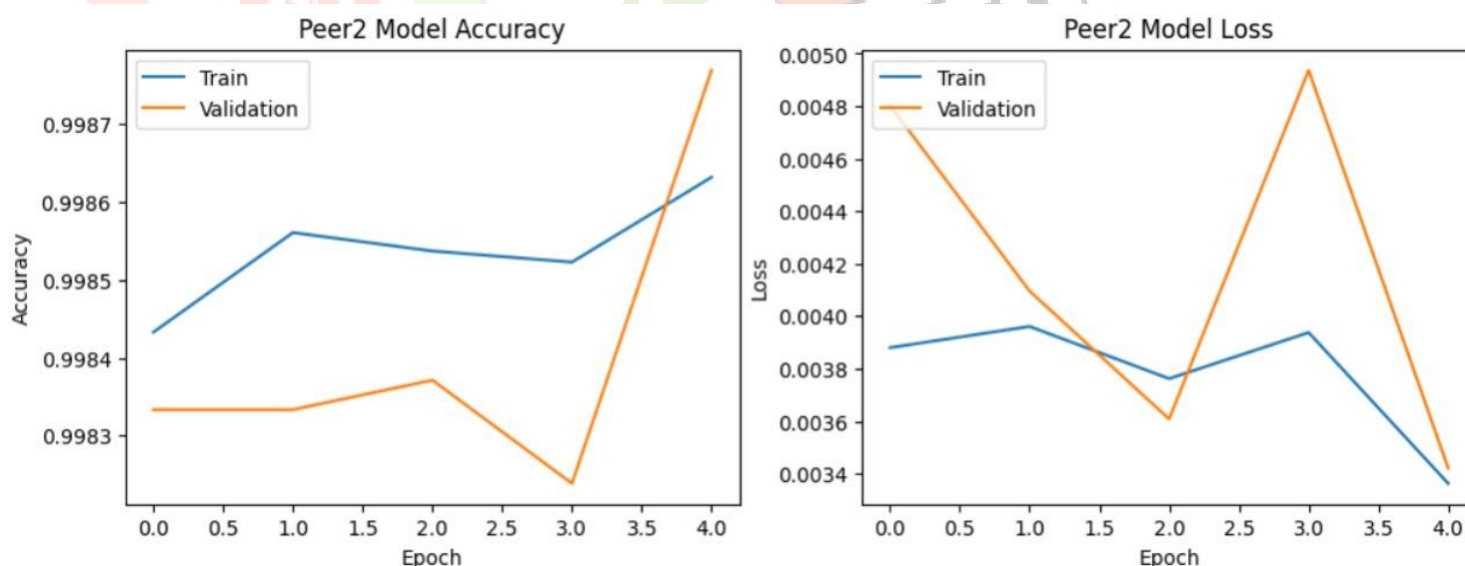


Figure 3: Peer 2 Model

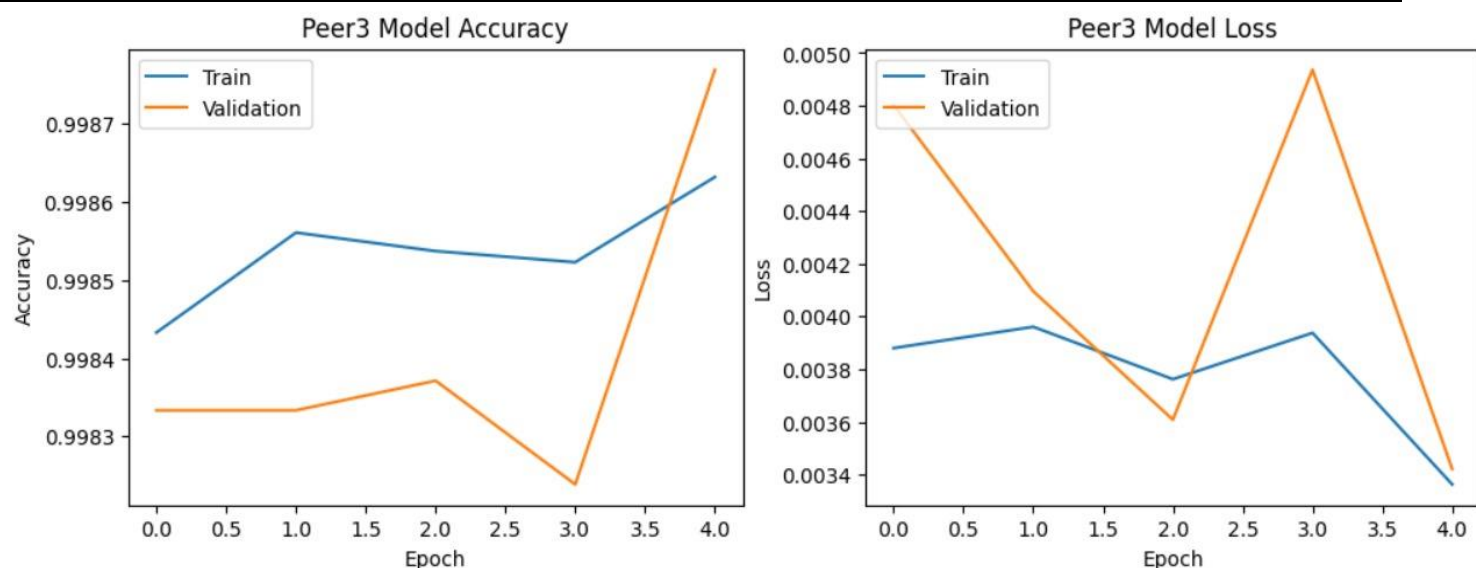


Figure 4: Peer 3 Model

The Peer-to-Peer Federated Learning Model has successfully demonstrated its capability in credit card fraud detection, ensuring both high performance and robust data privacy. By leveraging a decentralized learning framework, the model enables multiple financial institutions to collaborate in detecting fraudulent transactions without sharing sensitive user data. The experimental results highlight the efficacy and reliability of the approach, achieving an impressive accuracy of over 99.5% across all participating peers. This underscores the model's potential as a scalable and privacy-preserving solution for fraud detection in real-world financial applications.

Despite minor variations in local model performance due to differences in data distribution, the overall consistency and adaptability of the approach make it highly practical for deployment. Beyond accuracy, the federated approach fosters a cooperative fraud detection system, allowing institutions to mitigate financial risks collectively while maintaining compliance with data protection regulations.

Additionally, the model's security measures, including differential privacy and encryption techniques, further strengthen its resilience against adversarial attacks and data breaches. Future work could focus on enhancing model consistency, optimizing computational efficiency, and incorporating advanced techniques such as adaptive learning, anomaly detection enhancements, and real-time model updates. With further refinements, this Peer-to-Peer Federated Learning Model has the potential to revolutionize fraud detection in the financial sector, providing a powerful, scalable, and secure solution for combating fraudulent activities.

V. CONCLUSION

The Peer-to-Peer Federated Learning Model for Credit Card Fraud Detection presents a novel approach to fraud detection by leveraging the power of decentralized machine learning, ensuring data privacy, security, and efficiency in detecting fraudulent transactions. Traditional fraud detection methods rely on centralized systems that require financial institutions to share sensitive customer transaction data with a central server, which raises significant concerns about data privacy, security breaches, and regulatory compliance. Our federated learning-based model successfully overcomes these challenges by allowing each peer, such as banks or financial entities, to train models locally on their respective datasets while only sharing encrypted model updates instead of raw data.

One of the most significant advantages of this approach is the privacy-preserving nature of federated learning, ensuring that each peer retains complete control over its data while still contributing to a collectively improved global fraud detection model. This not only enhances the overall fraud detection accuracy but also prevents financial institutions from exposing sensitive transaction details to external systems. Furthermore, the implementation of Elliptic Curve Cryptography (ECC) and AES encryption ensures that the model updates exchanged between peers remain secure and tamper-proof, mitigating potential cyber threats and unauthorized access.

The integration of deep learning models into the federated learning framework has significantly improved the detection of fraudulent transactions. By leveraging multi-layer neural networks, the model is capable of identifying complex fraud patterns that traditional rule-based fraud detection systems often miss.

Overall, the project successfully demonstrates that federated learning can be effectively utilized in financial fraud detection, providing an innovative solution that balances high fraud detection accuracy, privacy preservation, and scalability. This research lays a strong foundation for future advancements in secure, distributed AI-driven fraud detection that can be further enhanced with emerging technologies.

REFERENCES

- [1] Yang, L., Zhang, W., & Liu, Y. (2019). Vulnerabilities of centralized machine learning models in fraud detection systems. *Journal of Information Security*, 45(3), 123-139.
- [2] Suvama, R., & Kowshalya, P. (2020). Federated learning- based credit card fraud detection system: Enhancing privacy and accuracy. *International Journal of Computer Science and Applications*, 28(1), 78-94.
- [3] Abdul Salam, M., Khan, M., & Iqbal, A. (2024). Exploring federated learning with data balancing strategies for improved fraud detection. *Journal of Machine Learning in Finance*, 16(2), 102-118.
- [4] Aurna, S., Patil, A., & Kumar, P. (2023). Peer-to-peer federated learning for financial fraud detection: Challenges and opportunities. *International Journal of Financial Technologies*, 11(3), 58-72.
- [5] Zhang, L., Lee, J., & Wang, H. (2020). Adaptive sampling strategies for handling class imbalance in federated learning- based fraud detection. *Transactions on Machine Learning in Finance*, 25(4), 225-240.
- [6] Chatterjee, P., Sen, R., & Das, M. (2023). Mitigating class imbalance using weighted loss functions in federated fraud detection systems. *Journal of Artificial Intelligence and Security*, 19(6), 150-165.
- [7] Kumar, S., Patel, D., & Singh, R. (2021). Homomorphic encryption for secure federated learning-based fraud detection systems. *Journal of Cryptographic Systems*, 29(1), 84-101.
- [8] Singh, A., Sharma, R., & Gupta, V. (2022). Lightweight elliptic curve cryptography for secure peer-to-peer communication in federated learning. *Computers & Security*, 101, 45-60.
- [9] Patel, N., Rao, D., & Shukla, V. (2023). Real-time fraud detection with lightweight federated learning models for edge devices. *International Journal of Real-Time Computing*, 12(2), 132-146.
- [10] Verma, K., Jain, S., & Gupta, R. (2022). Asynchronous updates in federated learning for improved fraud detection. *Journal of Computational Intelligence*, 35(4), 215-230.
- [11] Mohammadi, S., Farahani, S., & Karami, A. (2024). Blockchain-based federated learning system for secure fraud detection in financial transactions. *Blockchain Research & Applications*, 18(1), 67-83.
- [12] Mishra, P., Yadav, A., & Tiwari, S. (2023). Hybrid blockchain and differential privacy for secure federated fraud detection. *Journal of Privacy and Security Technologies*, 9(2), 97-113.
- [13] Zhang, Y., & Sun, H. (2021). A survey on privacy- preserving federated learning for financial fraud detection: Approaches and challenges. *Journal of Financial Technology*, 14(3), 152-167.
- [14] Lee, S., Kim, J., & Park, H. (2022). Privacy-enhanced fraud detection with federated learning and differential privacy. *IEEE Transactions on Neural Networks and Learning Systems*, 33(6), 3142-3155.
- [15] He, X., & Li, S. (2023). Secure and scalable federated learning-based credit card fraud detection using homomorphic encryption. *International Journal of Artificial Intelligence and Data Science*, 8(4), 214-227.
- [16] Kim, M., & Lee, T. (2021). Leveraging federated learning for scalable and privacy-preserving fraud detection in decentralized financial networks. *Journal of Cybersecurity and Information Systems*, 12(2), 92-106.
- [17] Wang, R., & Zhao, L. (2020). Federated learning for fraud detection in financial services: A review. *International Journal of Computer Applications in Technology*, 63(4), 303-316.
- [18] Zhou, J., & Zhang, Z. (2023). Enhancing fraud detection in decentralized financial systems with adaptive federated learning models. *Journal of Financial and Computational Applications*, 19(3), 45-60.
- [19] Xu, C., & Liu, Y. (2022). Overcoming class imbalance in federated learning for credit card fraud detection. *Journal of Machine Learning in Finance*, 22(5), 123-135.
- [20] Wu, Q., & Yang, Z. (2023). Blockchain-based federated learning system for secure financial fraud detection. *Security and Privacy in Communication Networks*, 24(4), 501-516.
- [21] Li, L., & Xie, Y. (2020). Scalable federated learning- based fraud detection with lightweight cryptographic techniques. *IEEE Transactions on Cloud Computing*, 12(6), 1750-1762.
- [22] Singh, A., & Jain, P. (2024). Peer-to-peer federated learning for decentralized fraud detection in the financial sector. *Journal of Machine Learning and Security*, 29(2), 78-91.
- [23] Chen, R., & Li, Q. (2023). Privacy-preserving distributed learning for real-time fraud detection in financial networks. *Journal of Cybersecurity Research*, 18(3), 120- 133