Literature Review about FL applications in credit card fraud (CCF)

Federated Learning (FL) has become an innovative approach in the domain of Credit Card Fraud (CCF) detection, offering a decentralized solution to the traditional centralized machine learning models. This literature review synthesizes the current state of research on the application of FL in CCF detection.

Introduction

The integration of FL in CCF detection systems is motivated by the need to enhance data privacy and security while improving the detection accuracy of fraud models. FL allows multiple financial institutions to collaboratively train a shared model without the need to exchange or centralize their sensitive transaction data.

Key Research Contributions

1. Data Privacy and Security: FL preserves data privacy by keeping the raw data localized and only sharing model updates, as proposed in the study by Mustafa Abdul Salam et al. (2024).

2. Optimization Techniques: The application of metaheuristic optimization algorithms, such as the improved Gorilla Troops Optimizer (AGTO) and Coati Optimization Algorithm (COA), has been shown to improve the initial global model, reducing communication costs and accelerating convergence in FL, as documented in the paper by Mustafa Abdul Salam et al.

3. Handling Class Imbalance: The use of resampling strategies like Synthetic Minority Oversampling Technique (SMOTE) has been highlighted as a crucial preprocessing step to address the highly skewed nature of credit card datasets, improving the performance of unbalanced or skewed data in fraud detection models.

Methodological Frameworks

- FL Model Development: The research by Baabdullah et al. (2024) introduces a CCFD system that integrates FL with blockchain technology, ensuring privacy preservation and data protection while enhancing classification performance and prediction accuracy.

- Algorithmic Approaches: The study employs machine learning and deep neural network learning algorithms, including Random Forest (RF), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM), alongside deep optimization techniques such as ADAM, SGD, and MSGD.

- Data Preprocessing and Feature Extraction: The Europe Credit Card (ECC) dataset, which is highly imbalanced, is used in the experiments. The dataset undergoes PCA transformation and normalization to prepare it for model training.

Performance Evaluation

The performance of FL models in CCF detection is evaluated using metrics such as accuracy, precision, recall, F1-score, computation time, and average loss. The results from the studies indicate that FL models, when optimized with the right algorithms, can achieve high accuracy and efficiency in detecting credit card fraud.

Challenges and Limitations

Despite the promising results, several challenges and limitations have been identified:

- Computational Overhead: The integration of FL with blockchain may impose significant computational demands, particularly when handling large-scale data.
- Data Synchronization: Maintaining data consistency and synchronization across multiple nodes within a blockchain network adds complexity.
- Scalability and Security Concerns: Ensuring scalability while addressing security and privacy concerns in the FL and blockchain systems is critical.

Future Research Directions

- Advanced Optimization Techniques: Further research is needed to develop more efficient optimization algorithms for FL models in CCF detection.
- Real-time Detection Systems: There is a need for real-time CCF detection systems that can adapt to the evolving nature of fraud tactics.
- Regulatory Compliance: Addressing the intersection of FL and blockchain with regulatory compliance will be crucial for the adoption of these technologies in the financial sector.

Conclusion

FL presents a significant advancement in the field of CCF detection, offering a robust framework that enhances data privacy, reduces the risk of data breaches, and improves the overall performance of fraud detection systems. The integration of FL with other technologies like blockchain further strengthens the system's capabilities. However, ongoing research is necessary to overcome the computational, scalability, and regulatory challenges to fully realize the potential of FL in CCF detection.