

Advancements of Credit Card Fraud Detection Based on Federated Learning

Hongwei Wang^a

Software Engineering, Iowa State University, Ames, Iowa, United State

Keywords: Machine Learning, Federated Learning, Credit Card Fraud Detection

Abstract: Credit card fraud is an increasing concern as digital payment systems and e-commerce continue to expand. The purpose of this research is to investigate the use of federated learning (FL) to identify credit card fraud while maintaining the privacy of data across financial institutions. Unlike traditional centralized models, federated learning can let to train models collaboratively without disclosing raw data. This study reviews critical federated learning approaches, such as vertical and horizontal FL, and highlights their effectiveness in addressing privacy concerns, data heterogeneity, and class imbalance. Integrating advanced techniques like expert systems, explainable Artificial Intelligence (AI) tools and hybrid resampling methods enhances AI models' detection accuracy and transparency. Federated learning offers significant advantages for safeguarding financial transactions by improving real-time fraud detection and adaptive learning. The study also identifies challenges like model scalability, communication overhead, and security threats like inference attacks. Nonetheless, the potential for federated learning to transform credit card fraud detection by combining privacy-preserving, interpretable, and scalable solutions is considerable, positioning it as a critical technology in the prevent financial fraud.


1 INTRODUCTION

Credit card fraud has become increasingly common as digital payment systems and e-commerce platforms gain widespread popularity. Fraudsters exploit vulnerabilities in the financial system through a lot of different means, such as unlawful get credit card information, identity theft, skimming, phishing, and sophisticated malware. The increasing number of online and mobile transactions has led to a rise in card-not-present fraud, where attackers make unauthorized purchases without physically using the card.

Moreover, the large volume and speed of financial transactions, especially in real-time payment systems, make human monitoring insufficient for identifying complex fraud patterns. Artificial intelligence (AI) is essential in tackling these difficulties, allowing for more accurate, real-time fraud detection. Unlike traditional rule-based systems, AI and machine learning models may learn from massive volumes of transactional data, detecting subtle patterns and correlations that could indicate fraudulent behavior. These algorithms constantly evolve, adjusting to new

fraud strategies as they emerge. The AI can increase the accuracy of finding the fraud plan. Artificial intelligence can analyze significant data sets to decline the number of false alarms and missed alarms. Also, the AI can do self-update according to the dataset updating, suited to respond to rapidly changing fraud techniques.

Many fields have started to consider the application of AI, for example, in the study carried out Banerjee et al. (Banerjee et al., 2018). Crop management, pest control, disease management, soil and irrigation management, vegetation management, and yield prediction are all applications of AI techniques. The essay emphasizes how AI helps improve decision-making processes in agriculture by offering solutions to problems like pest infestation, improper soil treatment, and disease detection. Artificial intelligence is also used in business analytics. In the study that Schmitt et al. (Schmitt et al., 2023) carried out, they use Automated Machine Learning (AutoML) to improve business analytics and decision-making processes. It investigates how AutoML frameworks can enhance machine learning (ML) adoption across industries, particularly for non-

^a <https://orcid.org/0009-0002-5089-5976>

expert users, by automating the selection of models, hyper-parameter tuning, and evaluation processes. A large part of the business analysis is credit card fraud detection. Hasan et al. emphasizes transparency and interpretability in AI models, particularly in regulated industries like finance. The study assesses four machine learning models (ANN, SVM, Random Forest, and Logistic Regression) using a dataset that is extremely skewed. When it came to identifying fraudulent transactions, ANN had the highest precision and SVM the highest recall. In the study of Alenzi and Aljehane (Alenzi et al., 2020), the research also emphasizes the financial advantages of employing explainable AI models for adaptive learning and real-time fraud detection. It is creating a logistic regression-based artificial intelligence system to identify credit card fraud. It deals with the rising problem of credit card fraud brought on by an increase in internet sales. It draws attention to the difficulties in handling erratic, unbalanced data and dataset outliers. Additionally, there are numerous connections between federated learning and credit card fraud detection. The paper that Abdul Salam et al (Abdul Salam et al., 2024) proposes a FL approach for prevent credit card fraud, addressing critical issues such as data privacy and class imbalance in financial organizations. Traditional centralized models have data sharing limits, therefore federated learning is a feasible alternative since it allows institutions to exercise a shared model without disclosing sensitive data. Addressing class imbalance, the research employs a variety of resampling strategies, including oversampling (Smote, AdaSyn) and undersampling (RUS), and analyses their efficacy across multiple machine learning algorithms. The Random Forest (RF) classifier has the best accuracy (99.99%), beating approaches such as Logistic Regression and K-Nearest Neighbors. The study demonstrates how hybrid resampling methods combined with federated learning significantly improve fraud detection while maintaining data privacy.

For the rest part of paper, first, how the other researcher uses Federated Learning in credit fraud or different ways that the researcher will commonly use AI to prevent or protect against credit card fraud in Section 2, compare these methods to each other, and emphasize the challenge facing in this area in Section 3. In Section 4, summarize the paper and get the conclusion according to the paper that this paper discussed here.

2 METHOD

2.1 Preliminaries of Federated Learning

Federated Learning (FL) is a decentralized machine learning technique that accept numerous devices or institutions to coach a model without giving raw data. Unlike traditional approaches, where data is centralized, FL allows for model training at the data source, ensuring privacy and security. After local training, clients send only model modifications (weights, gradients, etc.) to a central server, which aggregates these updates to enhance the model as a whole. The clients receive the updated model back, and the process keeps on until the model is optimized. FL is especially effective in sensitive industries such as business, hospitality, and mobile applications, where data privacy is a top priority. The pipeline includes model initialization, local training, aggregation, and global model updates. FL complies with GDPR laws by guaranteeing that data remains with the owner while benefiting from collective learning across many sources.

2.2 Vertical Federated Learning

The Boliang Lv et al.'s (Lv et al., 2021) paper explores the application of Vertical Federated Learning (VFL) for identifying e-banking fraud accounts by combining financial and social data while preserving privacy. Vertical federated learning allows financial institutions and social media companies to collaborate without sharing sensitive data, using encrypted sample alignment and model training techniques. The system was tested using 60,000 data samples, demonstrating that federated learning improves accuracy, precision, and recall compared to local models. This method is beneficial in detecting fraudulent accounts before transactions occur, safeguarding users, and enhancing fraud prevention. The study carried out by Fan et al. (Fan et al., 2024) highlights federated learning's effectiveness in cross-industry data collaboration. The study introduces a unique defence method termed Federated Learning Similar Gradients (FLSG), intended to prevent inference assaults in VFL. When various institutions want to collaborate on model training without sharing raw data yet have different feature sets for the same users, they employ VFL. In this strategy, the study proposes replacing actual gradients with identical gradients estimated using cosine distance, making label inference attacks more difficult. This strategy improves VFL security by protecting privacy while

retaining model accuracy, which is especially useful in industries such as banking and healthcare.

2.3 Centralized Federated Learning

A paper written by Taoufik El Hallal et al. (Hallal et al., 2024) explores the application of centralized federated learning for credit card fraud detection, where a central server aggregates model updates from multiple decentralized sources to build a global model while preserving data privacy. This method addresses the challenge of data privacy in the financial sector by allowing institutions to collaborate without sharing raw data. The study highlights how federated learning improves fraud detection accuracy compared to traditional models, mainly when dealing with imbalanced datasets. By leveraging advanced techniques such as Convolutional Neural Networks (CNN) and sampling methods, the federated learning model enhances the detection of fraudulent transactions across institutions. The study performed by Myalil et al. (Myalil et al., 2021) suggests using Centralized Federated Learning to detect fraudulent transactions in financial institutions. In this method, a central server manages the learning process by combining the models trained locally by individual banks on their private data. The centralized method protects privacy while enabling robust fraud detection across several banks without exchanging sensitive transaction data. The study addresses issues such as non-IID data distribution (where banks have diverse transaction patterns) and hostile actors attempting to disrupt the learning process. Epsilon Cluster Selection, an upgraded aggregation technique, is proposed to address these issues, making the fraud detection procedure more reliable. Baabdullah et al.'s (Baabdullah et al., 2024) article employs Centralized Federated Learning, in which a global model is trained on a central server, and starting parameters are given to local models located on fog nodes (banks). Each bank trains its model with local data and returns updated parameters to the global model, maintaining privacy and prohibiting data sharing. The use of blockchain improves data security and immutability. The method is used to improve classification performance and accuracy in detecting fraudulent credit card transactions while protecting privacy.

2.4 Horizontal Learning

Zheng et al. (Zheng et al., 2020) present a Horizontal Federated Learning framework paired with meta-learning for detecting counterfeit credit cards. In this strategy, many institutions work together to build a

fraud detection model without revealing sensitive client information. Each bank trains a local model with its private dataset and updates the global model via a centralized server. The model features a novel K-tuplet loss method to increase feature extraction and discover previously unseen fraud events. Experimental results from various datasets show that this federated meta-learning technique outperforms traditional methods regarding accuracy and recall while maintaining privacy. Sha (Sha, 2023) describes a Horizontal Federated Learning technique for detecting financial fraud utilizing big data technology. This strategy involves multiple financial institutions working together to train machine learning models using their local datasets, which have similar properties but are stored separately across geographies. The institutions only share model updates, not raw data, to ensure privacy and compliance with legislation. The research emphasizes the advantages of federated learning in overcoming data silos and enhancing model accuracy for identifying fraud across financial systems. Experiments show that the proposed solution is more accurate and preserves privacy than typical centralized models. Li et al. (Li, 2023) investigates the use of Horizontal Federated Learning for credit risk management in several institutions. In this technique, banks work together to train local models on their respective datasets, which have the same feature space but are scattered across institutions. The FedAvg algorithm combines locally trained models into a global model while protecting sensitive client data. The study addresses Non-independent and Identically Distributed (non-IID) data by employing techniques like the Chi-square test to increase data homogeneity. The experimental results demonstrate a considerable improvement in recall and F1 scores, indicating the efficacy of federated learning for managing credit risk while protecting data privacy.

3 LIMITATIONS AND CHALLENGES

While FL offers many advantages, some challenges and limitations must be considered to ensure its successful implementation. These limitations include data interpretability, scalability and applicability as shown below.

3.1 Applicability

In the three challenges, first challenges facing the federal financial sector (especially in economic areas

such as banking) is the existence of non-independent Identically Distributed (non-IID) data. In practice, different organizations (e.g., banks, financial service providers, or payment gateways) handle various transactions, customer data, and fraud patterns. For example, a large multinational bank may process many transactions across countries and currencies, while a smaller regional bank may have a smaller, more homogeneous dataset. This imbalance in data distribution can severely impact global models trained in a federated learning environment. In standard machine learning, models are typically trained on datasets that follow a uniform distribution (IID), simplifying convergence and improving model accuracy. In federated learning, data differences between organizations can slow model convergence, reduce accuracy, and create bias in the global model. This bias can lead to inaccurate fraud detection in organizations with smaller or less diverse datasets, which can cause the international model to favor data from larger organizations.

3.2 Scalability

With more and more organizations join Federated Learning networks, systems must efficiently handle an increasing number of clients. Scalability becomes an important concern in large networks with hundreds or thousands of users. The aggregation process must handle model updates from many clients without becoming a bottleneck, and the communication overhead must remain manageable even as the network grows. Hierarchical federated learning and layered aggregation approaches are potential solutions to improve scalability. In these approaches, local models are aggregated at intermediate nodes before being passed up to the central server, decreasing the burden on the main server and increasing communication efficiency.

3.3 Interpretability

Although FL enhances data privacy by decentralized raw dataset, it is not entirely resistant to security and privacy threats. One of the most important issues is the risk of attacks, where an attacker can attempt to infer sensitive information about an individual customer or dataset from shared model updates. In the context of credit card fraud detection, such attacks could compromise customer data, expose transaction history, or even reveal personal financial information. Another security risk is the presence of malicious participants in the FL process. In some cases, participants may intentionally submit corrupted model updates to affect the performance of the global

model, which may partiality the model toward inaccurate fraud predictions. These Byzantine attacks may undermine the reliability of the worldwide model, thereby reducing its effectiveness in detecting fraudulent transactions.

For future prospects, such as Shapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). Federated learning will continue to facilitate collaboration between financial institutions while ensuring data privacy and sharing model improvements without exposing sensitive information. By combining this method with expert systems, rule-based decision-making may be achieved, which improves the dependability and interpretability of machine learning models. To increase trust and compliance, tools such as SHAP and LIME are critical to delivering interpretable AI. These tools can explain model decisions and allow auditors and stakeholders to understand why certain transactions are flagged as fraudulent. Making machine learning models more transparent can enhance user trust and ensure regulatory compliance. Together, these technologies will shape the future of fraud detection, making it safer, more adaptable, and easier to explain.

4 CONCLUSIONS

This paper provides the three different learning of the federated learning, which is VFL, CFL and HL. Federated learning is promising to improve credit card fraud prevention and detection by facilitating collaboration between financial institutions while safeguarding data privacy. However, several challenges need to be overcome to implement federated learning successfully. These challenges include addressing data heterogeneity, as financial institutions have diverse datasets with varying structures, leading to issues with model convergence. Communication overhead is also a concern, especially when multiple institutions are involved, as frequent model updates need to be communicated. Furthermore, federated learning is vulnerable to security threats, such as inference attacks or malicious participants attempting to compromise the global model. To reduce these concerns, methods like secure multi-party computation and differential privacy are essential.

REFERENCES

- Abdul Salam, M., Fouad, K. M., Elbably, D. L. et al. 2024. Federated learning model for credit card fraud detection with data balancing techniques. *Neural Comput & Applic*, 36, 6231–6256.
- Alenzi, H. Z., & Aljehane, N. O. 2020. Fraud detection in credit cards using logistic regression. *International Journal of Advanced Computer Science and Applications*, 11(12).
- Banerjee, G., Sarkar, U., Das, S., & Ghosh, I. 2018. Artificial intelligence in agriculture: A literature survey. *International Journal of Scientific Research in Computer Science Applications and Management Studies*, 7(3).
- Baabdullah, T., Alzahrani, A., Rawat, D. B., & Liu, C. 2024. Efficiency of federated learning and blockchain in preserving privacy and enhancing the performance of credit card fraud detection (CCFD) systems. *Future Internet*, 16(6), 196.
- Fan, K., Hong, J., Li, W., Zhao, X., Li, H., & Yang, Y. 15 Jan. 2024. FLSG: A novel defense strategy against inference attacks in vertical federated learning. *IEEE Internet of Things Journal*, 11(2), 1816-1826.
- Li, Y., & Wen, G. 2023. Research and practice of financial credit risk management based on federated learning. *Engineering Letters*, 31(1), 28–46.
- Lv, B., Cheng, P., Zhang, C., Ye, H., Meng, X., & Wang, X. 2021. Research on modeling of e-banking fraud account identification based on federated learning. In 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech), AB, Canada (pp. 611-618).
- Myalil, D., Rajan, M. A., Apte, M., & Lodha, S. 2021. Robust collaborative fraudulent transaction detection using federated learning. In 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), Pasadena, CA, USA (pp. 373-378).
- Schmitt, M. 2023. Automated machine learning: AI-driven decision making in business analytics. *Intelligent Systems with Applications*, 18, 200188.
- Sha, X. 2023. Research on financial fraud algorithm based on federal learning and big data technology. Columbia University. <https://arxiv.org/abs/2405.03992>
- El Hallal, T., & El Mourabit, Y. 2024. Federated learning for credit card fraud detection: Key fundamentals and emerging trends. 2024 International Conference on Circuit, Systems and Communication (ICCSC), Fes, Morocco, pp. 1-6.
- Zheng, W., Yan, L., Gou, C., & Wang, F. 2020. Federated meta-learning for fraudulent credit card detection. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20) Special Track on AI in FinTech*, 4654-4660.