Federated Learning-Based Credit Card Fraud Detection: A Comparative Analysis of Advanced Machine Learning Models

Han Zheng

Questrom School of Business, Boston University, 02215 Boston, United States

Abstract. Because of the privacy concerns about the transaction data, it is essential not to leak it when training prediction models for credit card fraud analysis. Challenges for credit card fraud monitoring include highly imbalanced datasets and the need for advanced models to detect fraud patterns. This paper introduced federated learning and discussed a few federated learning algorithms applied to the problem—these methods include Federated Graph Attention Network with Dilated Convolution Neural Network (FedGAT-DCNN), FedAvg with Convolutional Neural Network (CNN), and Federated Averaging with Distance-based Weighted Aggregation (FedAvg-DWA) with Random Forest (RF). Federated Averaging (FedAvg) aggregates data from local clients and then creates a global model; fedavg-dwa provides dynamic weight averaging, which enhances each client's performance based on their data quality. The FedGAT-DCNN model improves accuracy by integrating GAT with Dilated Convolutions to catch spatial and temporal patterns in transaction data. FedGAT—DCNN performs best on highly imbalanced datasets with a high Area under the Receiver Operating Characteristic Curve (ROC-AUC) score. FedAVG-DWA provides the best performance in different clients' systems. However, system heterogeneity, communication costs, and data imbalance remain critical. Oversampling techniques, model optimization, and reduced communication rounds were used to mitigate the issues. Therefore, federated learning's ability can enhance credit card fraud sensing issues without privacy concerns. The paper highlights both the benefits and challenges of using federated learning in the domain.

1 Introduction

Credit card fraud is a type of crime that harms the wealth of another party by stealing or scamming their digital properties or cash. This type of fraud has caused billions of dollars globally for individuals and banks. Even after numerous notifications from governments and mechanisms trying to detect and catch these frauds, the total amount of money has been lost increased from $27.9 billion in 2018 to $33.5 billion in 2022 [1]. Furthermore, according to the study carried out by Pushpita et al. [2], frauds have at least 18 methods to cause financial loss through credit cards.

Corresponding author: andyzh@bu.edu

Machine learning is one area that has been developing exceptionally fast in the past decades; ANN, decision trees (DT), and RFs are some examples of machine learning algorithms that have been applied to different areas like chemistry, biomedicine, and finance. For instance, in the medical area, Karthik et al. use White Blood Cell Segmentation Network (WBC-Net), developed upon Nested U-Net (UNet++) and Residual Network (ResNet), to attain a better WBC segmentation performance counting and identifying blood smear images' white blood cells, helping the diagnosis of COVID-19 [3]. Xia and Kais have used the Boltzmann machine to calculate accurate molecular potential energy surfaces in quantum chemistry. Using machine learning techniques speeds up the calculation process and the accuracy of the result [4]. Bat-neural Network Multi-agent System (BNNMAS) is proposed by Shahrabi et al. to predict the XETRA DAX Price Index (DAX) [5]. After comparing the result with the Genetic Algorithm Neural Network (GANN) and Generalized Regression Neural Network (GRNN), their model significantly outperforms the former models in accuracy and reliability.

In all different areas, researchers have carried out experiments trying to detect credit card fraud. Sahin et al. have applied a Support Vector Machine (SVM) and DT to credit card fraud monitoring problems [6]. In their paper, they have built seven classifiers based on the two algorithms and demonstrated that Classification and Regression Trees (C&RT), based on a DT, outperforms other models based on an SVM machine [7]. In further research, Sahin et al. developed a cost-sensitive DT that outperformed traditional data mining methods, including DTs, artificial neural network (ANN), and SVM. All of the works are important for preventing future credit card fraud activities. However, these researches have ignored customers' personal data privacy concerns. In recent studies, some researchers have noticed this problem and started introducing federated learning, an algorithm that trains the models on local devices and aggregates to a central model without concerns for personal data leakage, to create new models for fraud detection. The study performed by Salam et al. specifically put data privacy concerns into their model. They use different existing models, including (RF), Logistic Regression, K-Nearest Neighbors (KNN), DT, and Gaussian Naïve Bayes (NB) combined with a federated learning method. Their results show that RF outperforms other classifiers [8].

This paper is going to review the most recent methods for the application of federated learning on financial credit card fraud prevention. The remainder of this paper will be organized as follows: Section 2 will give modern federated learning methods applied in credit card fraud alert system problems. Section 3 will raise some potential challenges and deficiency. Section 4 will summarize the paper and give conclusions from the methods discussed above.

2 Method

2.1 Introduction of federated learning

Federated learning is a decentralized machine learning method in which many and different clients work with each other to train a model without the concern of leaking local data. In order to achieve customers' privacy, the training on the data happened on the local devices, and the model parameters after the update of each training will be sent to the server instead of uploaded to a central server. The server will aggregate all parameters to create a global model and then send it back to the clients for further training. Since the raw data always stays in the client devices, the data leakage concern will be eliminated. In addition, federated learning provides more efficient use of data from multiple sources.

2.2. FedGAT-DCNN

FedGAT-DCNN is a sophisticated model proposed by Li and Walsh integrating GATs and DCNNs into a federated learning framework to enhance credit card fraud identification [9]. GATs assign varying levels of attention to nodes within a graph-based transactional dataset, enabling the model to focus on critical transaction patterns. The ability of GATs is crucial in detecting fraud networks like users and merchants, which the relationship between entities is intricate and dynamic. GATs allow the model to selectively prioritize suspicious nodes by using attention mechanisms, improving fraud detection accuracy.

DCNN complements GATs by expanding the receptive field of the model and not increasing computational costs. This enables the model to capture long-range dependencies in transactional sequences, such as irregular spending behaviors that may signal fraudulent activity over time. Combined, GATs and DCNNs allow the model to detect spatial and temporal patterns, effectively improving the detection of fraudulent behaviors that conventional models often miss.

Li and Walsh's work demonstrates that their method is effective. They tested their model on two datasets: 2018CN, an imbalanced fraud dataset, and 2023EU, a more balanced one. The model achieves a high ROC-AUC score of 0.9712 on the 2018CN dataset and an exceptional ROC-AUC of 0.9992 on the 2023EU dataset. The result shows its adaptability to different data distributions and robustness in fraud detection scenarios. The model's high accuracy and ability to minimize false positives highlight its potential for practical deployment in financial institutions, offering enhanced security without having concerns about data privacy through federated learning.

2.3 FedAvg

FedAvg, as a critical algorithm in federated learning [10], collect and adds up all the updates in multiple iteration by finding their average to create a new global model and then give back to all devices. As one algorithm in federated learning, FedAvg ensures high privacy and efficiency. Additionally, it is robust in environments with Non-identically Distributed (non-IID) data, allowing devices with different data distributions to contribute to a shared model. This algorithm has been widely adopted in applications where data privacy is essential, such as healthcare and finance.

2.3.1 FedAvg with CNN

Yang et al. combined FedAvg with CNN to detect credit card fraud across multiple financial companies [11]. To handle the imbalanced nature of the credit card anomaly detection dataset, Yang's team employed the Synthetic Minority Over-sampling Technique (SMOTE) to rebalance the local datasets. Their results show that their model outperformed traditional centralized methods by achieving a higher AUC score of 0.969 and an F1 score of 0.9534. After adjusting the communication rounds and other parameters, they maintain accuracy and computational efficiency simultaneously.

2.3.2 FedAvg-DWA with random forest

FedAvg-DWA is a modified FedAvg algorithm which assigns sample weights to handle class imbalance [12]. The server aggregates multiple updates based on the distance between each client's local model and the global model, ensuring better model generalization like for fraudulent transactions as minority class. Their method combining FedAvg-DWA with RF

outperforms other standard federated learning algorithm such as FedAvg and federated proximal (FedProx) in terms of accuracy and F1 score.

3 Discussion

3.1 Class imbalance

In federated learning, class imbalance [13], especially in credit card risk prevention, exists significant challenges. Fraud detection datasets are often highly imbalanced, with only a tiny percentage of the total data are real fraudulent transactions. This imbalance leads to difficulties in training models. Many algorithms in federated learning tends to favor majority classes (legitimate transactions) over the minority class (fraudulent transactions), resulting in decreased accuracy in detecting frauds. In federated learning, this issue is further complicated by the non-identical and non-independent distribution of data across different clients (e.g., banks). Each client's data are different from another client, the converge ability for global model will be highly affected due to the imbalance problem in each client.

To address these issues, existing techniques including resampling (oversampling fraud cases using methods like Synthetic Minority Over-sampling Technique (SMOTE) or cost-sensitive learning are often applied. However, the privacy concerns and limited visibility into local datasets make these methods harder to implement in federated learning. Specialized loss functions, such as Focal Loss, or the introduction of dynamic weighting mechanisms, have been invented to enhance the model's sensitivity to genuine fraud transactions as minority classes in federated learning. As a result, it is crucial to mention class imbalance in federated learning thus improving fraud detection rates without compromising client data privacy.

3.2 System heterogeneity

System heterogeneity [14], as another major challenge in federated learning, has been waiting to be resolved. System heterogeneity means the variability in computational resources, communication speeds, and storage capacities among participating devices or clients in a network of federated learning. The variability among heterogeneous devices can highly affect the speed and performance of federated learning algorithms. "Stragglers", as some devices, will decrease the efficiency of the overall training process because the limitation of their computational resources.

In order to identify credit card fraud, especially when different clients (banks) train a common global fraud detection model without sharing sensitive transaction data with each other, system heterogeneity becomes a huge barricade. Since each bank can have their own infrastructure setups, which results in inconsistencies in time as how quickly local models will update and return to the central server. The discrepancies among the time will cause delays in model convergence, since the global model must wait for each client to finish completing the updates.

Some existing techniques in addressing system heterogeneity in federated learning including adaptive client selection, where faster clients will participate more in the model updates. Furthermore, FedAvg and FedProx, as two modified algorithms, would allow partial updates from slower clients. Thus, the slower clients could participate in the updating process without lagging the overall process. These approaches help in scalability and efficiency of federated learning, especially in tasks like credit card fraud prevention and control which time and accuracy are both highly considered. Therefore, more research would be needed to resolve system heterogeneity among different clients.

3.3 Communication costs

One more challenge in federated learning would be communication costs [15]. Since federated learning requires training machines to learn on decentralized data, each client must frequently send updates to the central server after the training of local model. This will ensure the data privacy but causes high communication costs, since model parameters would need to be sent to the central server. In credit card malicious transaction identification, imbalance in dataset enlarges the problem. The high-volume imbalanced data requires more model updates to increase the accuracy of the model and not having overfitting at the same time. Some techniques such as model compression, quantization and limiting the number of communication rounds can alleviate the costs. For example, model updates could ask less frequency and maintain high performance by selecting intelligent client or aggregation methods helps in strike in balance between communication efficiency and fraud detection accuracy. With the existing methods, more research would need to be done in optimizing communication while maintaining robust fraud detection accuracy.

4 Conclusion

The paper investigated some existing methods trying to build a federated learning model on credit card fraud monitoring. FedGAT-DCNN combines GATs with DCNN has given their highest ROC-AUC of 0.9992 on the 2023EU dataset. FedAvg with CNN also shows a high accuracy of AUC score at 0.969. FedAvg-DWA with RF also outperforms traditional FedAvg and FedProx in determining credit card fraud monitoring with high accuracy. However, challenges such as class imbalance, system heterogeneity and communication costs also bring barriers to the federated learning methods. In considering the benefits of federated learning such as a relatively higher privacy comparing to traditional models and its potential in generalization and scalability when system heterogeneity can be efficiently resolved, federated learning will bring more value and accuracy in resolving real life problems. Overall, federated learning has shown its special ability on credit card fraud detection problems.

References

1. Techopedia, Credit Card Fraud Statistics, https://www.techopedia.com/credit-card-fraud-statistics (2024), Accessed time: September 4, 2024.
2. P. Chatterjee, D. Das, D. Rawat, Securing financial transactions: Exploring the role of federated learning and blockchain in credit card fraud detection. Authorea Preprints (2023).
3. Y. Lu, et al., WBC-Net: A white blood cell segmentation network based on UNet++ and ResNet. Applied Soft Computing 101: 107006 (2021).
4. R. Xia, S. Kais, Quantum machine learning for electronic structure calculations. Nature Communications 9.1, 4195 (2018).
5. R. Hafezi, J. Shahrabi, E. Hadavandi, A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price. Applied Soft Computing 29, 196-210 (2015).
6. Y. Sahin, E. Duman, Detecting credit card fraud by decision trees and support vector machines. Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol. 1 (2011).
7. Y. Sahin, S. Bulkan, E. Duman, A cost-sensitive decision tree approach for fraud detection. Expert Systems with Applications 40.15, 5916-5923 (2013).

8. M. A. Salam, et al., Federated learning model for credit card fraud detection with data balancing techniques. Neural Computing and Applications 36.11, 6231-6256 (2024).
9. M. Li, J. Walsh, FedGAT-DCNN: Advanced Credit Card Fraud Detection Using Federated Learning, Graph Attention Networks, and Dilated Convolutions. Electronics 13.16: 3169 (2024).
10. B. McMahan, et al., Communication-efficient learning of deep networks from decentralized data. Artificial Intelligence and Statistics. PMLR (2017).
11. W. Yang, et al., FFD: A federated learning-based method for credit card fraud detection. Big Data–BigData 2019: 8th International Congress, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019. Springer International Publishing (2019).
12. K. Bian, H. Zheng, FedAvg-DWA: A novel algorithm for enhanced fraud detection in federated learning environment. 2023 4th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). IEEE (2023).
13. L. Wang, et al., Addressing class imbalance in federated learning. Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, No. 11 (2021).
14. A. Reisizadeh, et al., Straggler-resilient federated learning: Leveraging the interplay between statistical accuracy and system heterogeneity. IEEE Journal on Selected Areas in Information Theory 3.2: 197-205 (2022).
15. O. Shahid, et al., Communication efficiency in federated learning: Achievements and challenges. arXiv preprint arXiv:2107.10996 (2021).