Adaptive Federated Learning for Privacy-Preserving Modeling in Heterogeneous Financial Environments

Ruobing Yan , Yingyi Shu , Shihao Sun , Nuo Chen , Yingxin Ou , Yinghao Zhao *

Posted Date: 16 March 2026

doi: 10.20944/preprints202603.1104.v1

Keywords: federated learning; privacy protection; Fintech; sensitivity analysis

 doi:10.20944/preprints202603.1104.v1

Article

Adaptive Federated Learning for Privacy-Preserving Modeling in Heterogeneous Financial Environments

Ruobing Yan [1], Yingyi Shu [2], Shihao Sun [3], Nuo Chen [4], Yingxin Ou [5] and Yinghao Zhao [6,*]

[1] Georgetown University, Washington, D.C., USA
[2] Cornell University, Ithaca, USA
[3] New York University, Brooklyn, USA
[4] University of Chicago, Chicago, USA
[5] University of Maryland, College Park, USA
[6] Pace University, New York, USA
* Correspondence: author: yinghaozhaony@gmail.com

Abstract

This study addresses the conflict between data privacy and modeling performance in financial technology by proposing a federated learning-based privacy-preserving framework. The research first analyzes the sensitivity of user transaction data and the issue of cross-institutional data silos, highlighting the limitations of traditional centralized modeling in terms of privacy and compliance. To overcome these issues, a distributed collaborative modeling mechanism is designed, where participants train models locally and achieve global optimization through parameter uploading and weighted aggregation, thus avoiding centralized storage and transmission of raw data. Differential privacy and secure aggregation are introduced to ensure that individual information is not exposed during parameter exchange, enhancing the overall privacy protection of the system. Furthermore, to address the non-independent and non-identically distributed nature of financial data, a personalized regularization term is incorporated to mitigate the impact of distribution differences across data sources, thereby improving adaptability and robustness in heterogeneous environments. Experiments, including comparisons with mainstream methods and multi-dimensional sensitivity analyses, verify the effectiveness and superiority of the proposed method under privacy-preserving conditions, as shown by improvements in AUC, ACC, F1-Score, and Precision. The results demonstrate that the framework can ensure data security while maintaining strong predictive performance and stability. In summary, this study not only achieves secure modeling of financial data but also provides a feasible direction and reference for further research on privacy-preserving algorithms in financial technology applications.

CCS CONCEPTS: Computing methodologies~Machine learning~Machine learning approaches

Keywords: federated learning; privacy protection; Fintech; sensitivity analysis

I. Introduction

In the context of the rapid development of the digital economy, financial technology has become a central force that connects traditional financial systems with emerging information technologies and drives structural changes across the industry. With the widespread adoption of mobile payments, robo-advisory services, blockchain, and big data–based risk management, financial services are becoming more frequent, intelligent, and inclusive[1]. However, this trend also creates a heavy dependence on user data. Transaction records, consumption habits, asset allocation, and risk preferences are sensitive data dimensions that enhance personalization and accuracy but also

 doi:10.20944/preprints202603.1104.v1

increase the risk of privacy breaches and misuse. Striking a balance between data-driven services and security assurance has become one of the key challenges for the sustainable growth of financial technology.

From the perspective of regulation and compliance, data protection policies are becoming increasingly stringent worldwide. Financial institutions must carefully review their approaches to data usage and model development. Cross-border data flows and sharing are subject to strict restrictions, making local storage and compliance essential principles. At the same time, the problem of data silos remains widespread. Different institutions or business units find it difficult to share information securely while protecting privacy, which limits the development of risk control, fraud detection, and integrated financial services. In this environment, a computational framework that ensures privacy protection while leveraging distributed data value is urgently needed in both academia and industry[2].

Federated learning has emerged as a promising paradigm that fits this demand. Its core idea is to enable collaborative training and global model optimization without centralizing raw data, thereby ensuring data localization and privacy. This mechanism effectively reduces the risk of privacy leakage while improving adaptability and robustness across multiple data sources[3]. In financial technology scenarios, federated learning supports collaborative modeling across institutions and domains. It also promotes the release of data value within a regulatory framework, providing reliable technical support for personalized financial services and risk management.

In financial practice, the tension between privacy protection and data utilization is particularly evident. Tasks such as risk assessment, anti-money-laundering monitoring, credit scoring, and targeted marketing often require large-scale data integration from multiple sources to ensure completeness and accuracy. Yet centralized modeling cannot meet compliance and security requirements. It exposes data and risks creating a trust crisis. Federated learning provides a new solution by combining distributed parameter updates with secure aggregation. This allows data owners to participate in collaborative modeling without exposing core information. The approach not only satisfies regulatory requirements but also builds a trustworthy foundation for industry cooperation[4].

In summary, research on privacy-preserving algorithms based on federated learning carries important theoretical and practical value in financial technology. On one hand, it addresses the central concerns of compliance and privacy protection, offering reliable guarantees for sustainable industry development. On the other hand, it opens new models of cross-institutional collaboration and joint risk control, fostering a more open, shared, and efficient financial ecosystem. At a broader level, such research advances the integration of artificial intelligence and financial technology, while offering new perspectives and methodologies for balancing privacy protection with data utilization. It holds significant potential for application and societal value.

II. Related Work

Privacy-preserving federated modeling in sensitive financial settings requires a principled treatment of both threat surfaces and evaluation criteria, where privacy is not a binary property but a continuum governed by leakage channels, adversary capabilities, and measurable protection strength. A systematic view of privacy mechanisms and metrics in federated learning motivates designing the overall framework as a composition of complementary defenses—covering update exposure, inference risks, and aggregation visibility—while using privacy/utility indicators to guide sensitivity analysis and compliance-oriented reporting [5]. Building on this foundation, differential privacy provides an operational mechanism for bounding information leakage from exchanged updates, directly supporting the design choice of perturbing client-side outputs so that global optimization can proceed without exposing individual transaction-level signals, and clarifying how the privacy budget mediates the trade-off between predictive accuracy and privacy strength [6]. Beyond mechanism selection, federated machine learning studies on privacy-preserving collective

 doi:10.20944/preprints202603.1104.v1

risk prediction highlight the viability of cross-organization collaboration under strict data localization, reinforcing the methodological premise that global model quality can be improved through parameter aggregation while keeping raw records confined to their owners [7].

A core challenge in federated financial data is heterogeneity and non-IID structure, where differing populations, product mixes, and behavioral regimes can induce spurious correlations that destabilize global objectives. Causal modeling paired with consistency-aware learning motivates incorporating regularization terms that penalize correlation-driven shortcuts and instead promote representation or prediction consistency across environments, which aligns with introducing personalization-aware constraints to mitigate distributional discrepancies across participants [8]. When heterogeneity is coupled with sparse labels or rapidly evolving patterns, meta-learning offers a complementary training paradigm that emphasizes fast adaptation and robust transfer, providing methodological justification for personalization-oriented objectives that maintain stable performance across clients and time while reducing overfitting to dominant sources [9].

At the model-architecture layer, transaction and behavioral signals are naturally heterogeneous and sequential, which makes representation learning choices critical for stable optimization under privacy constraints. Transformer modeling of heterogeneous records provides a transferable blueprint for fusing mixed-type attributes into coherent embeddings using attention-based alignment, supporting the design of local client models that can ingest diverse feature schemas without brittle feature engineering [10]. Transformer-based sequence modeling further motivates capturing long-range temporal dependencies and interaction dynamics in compact latent states, which is directly relevant for learning stable predictors from behavioral trajectories and for improving robustness under regime shifts that manifest over time [11]. Deep temporal convolution combined with attention offers an additional architectural alternative that emphasizes multi-scale temporal pattern extraction with selective focus, informing efficient local training when clients face constrained compute or require stable time-series representations under privacy-preserving update schedules [12].

When decision targets depend not only on individual behavior but also on relational structure, integrating graph context with sequence encoders can strengthen robustness and attribution while remaining compatible with federated optimization. Transformer-based risk monitoring with graph integration illustrates how sequential representations can be regulated by relational signals, motivating structured feature construction or auxiliary objectives that improve discrimination and stability without requiring centralized graph-level raw data sharing [13]. Extending this direction, dynamic spatiotemporal causal graph neural networks provide a methodological lens for separating propagation-driven effects from local noise by explicitly modeling time-evolving dependencies with causal inductive bias, which supports personalization-aware regularization as a way to reconcile global objectives with locally valid dynamics under heterogeneous client distributions [14].

Trustworthiness in privacy-preserving modeling is strengthened when outputs can be interpreted and communicated with calibrated uncertainty, particularly because privacy mechanisms can introduce additional noise that affects confidence. Explainable representation learning in large language models motivates designing representations that remain human-auditable and semantically aligned, which can inform how federated predictors expose feature-level or concept-level explanations without revealing sensitive raw records [15]. Knowledge-augmented agentic decision frameworks further motivate structuring explanations around explicit evidence and reasoning traces, suggesting a methodology for coupling federated predictions with rule- or knowledge-consistent rationales that enhance transparency under strict privacy constraints [16]. Risk-aware summarization with uncertainty quantification adds a complementary methodological component by encouraging calibrated confidence and uncertainty reporting, which is especially relevant when interpreting model outputs under differential privacy noise and when conducting sensitivity analyses over privacy budgets and aggregation settings [17].

Finally, optimizing privacy-preserving systems often involves balancing competing goals—utility, privacy, robustness, and fidelity of reporting—where multi-objective learning provides an explicit optimization template. Faithfulness-aware multi-objective context ranking motivates

 doi:10.20944/preprints202603.1104.v1

designing objective functions that jointly optimize predictive utility and faithfulness constraints, mirroring the need to jointly satisfy privacy guarantees and performance targets without collapsing into misleading shortcuts [18]. Dynamic prompt fusion for multi-task and cross-domain adaptation provides an additional methodological analogue for adaptive conditioning, inspiring how auxiliary guidance signals (e.g., client context, regime indicators, or uncertainty cues) can be fused to improve generalization across heterogeneous participants while preserving a consistent global training objective [19]. From a system-level perspective, modern AI-enabled analytics pipelines underscore that practical decision systems benefit when privacy, predictive modeling, and evaluation are integrated into an end-to-end workflow, reinforcing the framework's emphasis on compliance-compatible learning, multi-metric validation, and sensitivity-driven assessment as a coherent methodological package [20].

III. Proposed Approach

This study's technical design is grounded in the distributed collaborative modeling concept of federated learning, with the core goal of achieving unified, cross-institutional modeling while preserving strict data localization and privacy. In this system, each participant maintains its private financial data locally, updating parameters via local training rounds. Model weights or gradients are transmitted to an aggregation center, where secure aggregation is performed to produce a global model—thus eliminating the need for centralized raw data transfer and minimizing privacy risks.

To enhance anomaly detection and learning robustness within each participant, we incorporate the multi-head self-attention anomaly identification methodology as developed by Wang et al. [21], allowing local models to effectively capture and suppress transaction anomalies. Additionally, each local model utilizes the attention-based LSTM neural network framework from Li et al. [22], providing powerful temporal representation and intelligent anomaly filtering within financial time series data.Given the prevalence of noisy and imbalanced transaction data in real-world settings, the architecture integrates generative distribution modeling techniques as proposed by Xu et al. [23], supporting more accurate credit risk assessment at the participant level. To further optimize system performance and adaptability in complex data environments, deep Q-learning-based intelligent scheduling strategies inspired by Gao et al. [24] are applied to the parameter aggregation and communication process.

A schematic of the overall federated model architecture is provided in Figure 1.

Let D_i be the dataset of each participant and θ_i be the local model parameters. The objective function can be expressed as:

$$\min_{\theta} \sum_{i=1}^{N} \frac{|D_i|}{\sum_{j=1}^{N} |D_j|} L_i(\theta_i) \qquad (1)$$

The global model is then synchronized back to each participant to enter the next round of iterative training, realizing a cyclical collaborative optimization process.

In terms of privacy protection mechanisms, this study introduces a combination of differential privacy and secure aggregation. Differential privacy masks individual data characteristics by injecting noise into gradients or parameters, making it difficult to identify individual data points. The noise injection process can be expressed as:

$$\widetilde{\theta}_i^{(t+1)} = \theta_i^{(t+1)} + N(0, \sigma^2 I) \qquad (2)$$

 doi:10.20944/preprints202603.1104.v1

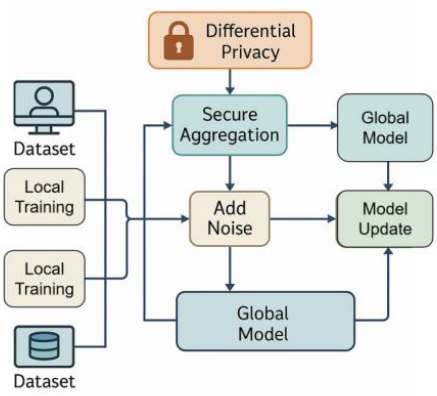

Figure 1. Overall model architecture.

Here, $N(0,\sigma^2 I)$ represents a Gaussian noise vector with mean 0 and variance σ^2. Furthermore, the secure aggregation mechanism ensures that the aggregation center can complete the weighted average calculation of model parameters without decryption, thus avoiding the risk of single-party parameter leakage. The combination of these two mechanisms further enhances the model's privacy protection capabilities in real-world financial scenarios.

To address the non-IID problem of financial data, this study introduces a personalized regularization term into the optimization objective to enhance the model's adaptability under heterogeneous data conditions. Specifically, the global optimization objective can be expanded to:

$$\min_{\theta} \sum_{i=1}^{N} \frac{|D_i|}{\sum_{j=1}^{N} |D_j|} (L_i(\theta_i) + \lambda \| \theta_i - \theta \|^2) \qquad (3)$$

Here, λ is the regularization weight, and $\| \theta_i - \theta \|^2$ is used to measure the deviation between the local model and the global model. By introducing this constraint, we can ensure global consistency while taking into account the individual differences of different participants, thereby improving the overall modeling effect and practical applicability.

IV. Experiment Result

A. Dataset

The dataset used in this study comes from a publicly available financial transaction dataset. It covers multi-dimensional user transaction behaviors and account information. The dataset is large in scale, containing millions of records. The fields include transaction time, user identifiers, transaction amounts, account balance changes, transaction locations, and device types. It also contains labels for both normal and abnormal transactions, which allows the construction of a binary classification task. This makes it suitable for evaluating the effectiveness of financial risk modeling under privacy-preserving settings.

In terms of data distribution, the dataset shows a clear imbalance. The frequency of different transaction categories varies significantly. In particular, the ratio between fraudulent and non-fraudulent samples is close to 1:1. This characteristic supports research on the detection of high-risk transactions and ensures that models can achieve strong generalization ability across categories during experiments. In addition, the dataset includes multiple types of transaction patterns, such as cross-border payments, online shopping, and daily transfers. This diversity helps evaluate the applicability of models in different business scenarios. The advantages of this dataset lie in its large scale, high degree of structure, and close connection to real financial environments. It provides a reliable validation setting for federated learning and privacy-preserving algorithms. By using this dataset, it is possible to test the performance of different algorithms under data security constraints. It also offers a solid foundation for subsequent experimental design and result analysis.

 doi:10.20944/preprints202603.1104.v1

B. Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1. Comparative experimental results.

Method	AUC	ACC	F1-Score	Precision
LSTM[25]	0.871	0.843	0.828	0.815
Transformer[26]	0.902	0.868	0.854	0.842
1DCNN[27]	0.889	0.859	0.846	0.832
ConvNextv2[28]	0.915	0.876	0.862	0.851
Ours	0.943	0.897	0.885	0.873

From the results in Table 1, it can be observed that traditional sequence modeling methods, such as LSTM, show relatively lower performance across all metrics. The AUC is 0.871 and the ACC is 0.843, indicating limited ability to capture the complex features of financial data. Because financial data exhibit strong nonlinearity and multi-scale dependencies, LSTM still faces limitations in long-sequence modeling. This leads to lower values in F1-Score and Precision. These results suggest that relying solely on recurrent structures makes it difficult to achieve both stability and accuracy in privacy-preserving financial modeling tasks.

In contrast, the Transformer model shows certain advantages in global dependency modeling due to its self-attention mechanism. The AUC increases to 0.902 and the ACC rises to 0.868, with corresponding improvements in F1-Score and Precision. This demonstrates that Transformer is more adaptive in capturing cross-temporal dependencies and feature interactions, and it can better characterize the dynamic patterns in financial data. However, its performance is still influenced by heterogeneous data and privacy constraints. This indicates that in real federated learning environments, relying only on global attention remains limited.

The 1DCNN and ConvNextv2 models represent different convolution-based approaches. 1DCNN performs better than LSTM in local pattern extraction, with an AUC of 0.889. However, due to the lack of effective modeling for long-range dependencies, its overall performance is slightly lower than Transformer. ConvNextv2, as an improved convolutional network with modern architecture design, outperforms the previous models. It achieves an AUC of 0.915 and an ACC of 0.876, showing advantages in handling high-dimensional and complex data. Yet convolutional structures in privacy-preserving financial technology tasks are still constrained by insufficient global modeling and cannot fully address the challenges of non-independent and non-identically distributed data.

Overall, the proposed method outperforms all baseline models in every metric. The AUC increases to 0.943, and the ACC reaches 0.897. The F1-Score and Precision are 0.885 and 0.873, respectively, showing significant improvements. These results demonstrate that combining the federated learning framework with privacy-preserving mechanisms not only reduces the privacy risks of data sharing but also achieves better predictive performance in heterogeneous data environments. This provides strong evidence of the practical value and potential applicability of the proposed method in financial technology scenarios.

This paper also presents an experiment on the sensitivity of batch size to the ACC indicator, and the experimental results are shown in Figure 2.

 doi:10.20944/preprints202603.1104.v1

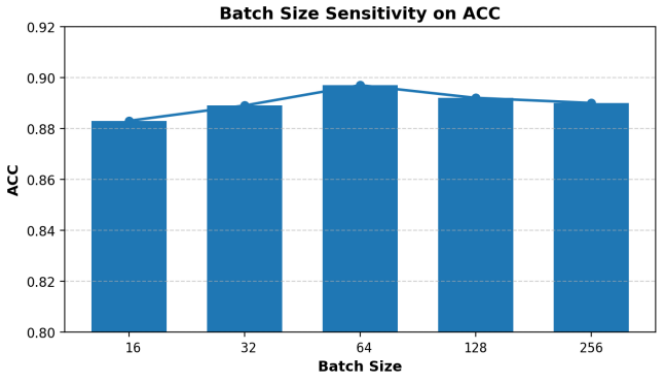

Figure 2. Sensitivity experiment of batch size on ACC indicator.

From the results in Figure 2, it can be observed that batch size has a direct impact on the ACC metric. As the batch size increases from 16 to 64, the ACC value gradually rises and reaches the highest value of 0.897 at batch size 64. This indicates that a moderate batch size helps the model balance gradient estimation and parameter updates, thereby improving prediction performance under privacy-preserving conditions.

When the batch size continues to increase to 128 and 256, the ACC shows a slight decline. This suggests that an excessively large batch may reduce the model's ability to capture details of the data distribution during local training, which weakens the generalization ability of the global model. In particular, within the federated learning framework, the data distributions of participants are often non-independent and non-identical. Too large a batch size reduces the sensitivity of local updates to heterogeneous features.

On the other hand, too small a batch size can also cause unstable gradient updates and increase the variance of the model during aggregation. The experimental results show that the ACC at batch sizes of 16 and 32 is significantly lower than at 64. This indicates that although smaller batches increase the update frequency, the overall stability and convergence performance of the model are not well guaranteed. This further confirms that in financial technology scenarios, the proper choice of batch size is a key factor for improving global model performance.

In summary, the experimental results indicate that batch size in a privacy-preserving federated learning framework is not better when simply larger or smaller, but rather there exists an optimal range. Within this range, the model can achieve both stability of local updates and robustness of global aggregation, thereby reaching the best performance in terms of ACC. This finding provides a useful reference for hyperparameter selection and model optimization. It also highlights the importance of sensitivity experiments under privacy-preserving constraints.

This paper also presents an experiment on the sensitivity of the number of federation rounds to the F1-Score indicator, and the experimental results are shown in Figure 3.

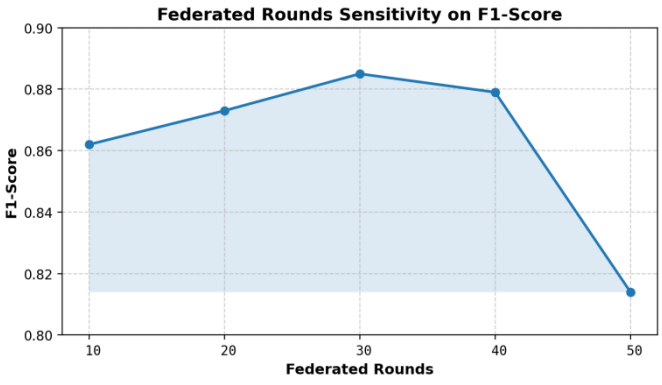

Figure 3. Sensitivity experiment of federation rounds to F1-Score.

 doi:10.20944/preprints202603.1104.v1

From the results in Figure 3, it can be observed that the number of federated rounds has a clear impact on the F1-Score. When the number of rounds is small, such as 10, the F1-Score remains around 0.862. This indicates that the model has not yet fully converged in the early stage of aggregation. Insufficient parameter updates limit the overall performance. As the number of federated rounds increases, collaborative optimization across participants' data becomes more stable, and the F1-Score shows a steady upward trend.

When the number of rounds reaches 30, the F1-Score peaks at 0.885. This shows that the model achieves an optimal balance between local training and global aggregation at this point. The result demonstrates that in a privacy-preserving federated learning framework, a moderate increase in rounds helps integrate multi-party data features more effectively. It also enhances the ability to capture complex patterns in financial data, which improves prediction accuracy and robustness.

However, when the number of federated rounds further increases to 40 and 50, the F1-Score declines slightly. This may result from overfitting caused by excessive rounds, which reduces the generalization ability of the global model on heterogeneous data from some participants. At the same time, frequent parameter exchanges may amplify communication costs and the effect of noise injection. These factors can weaken the performance of the model under privacy-preserving mechanisms.

Overall, the experimental results indicate that the choice of federated rounds is critical for model performance. Too few rounds prevent full convergence, while too many rounds may cause performance fluctuations and additional costs. A reasonable number of rounds can balance privacy protection, computational efficiency, and predictive accuracy. This provides practical guidance for parameter setting in federated learning systems for financial technology. It also highlights the importance of sensitivity experiments under privacy-preserving conditions.

This paper further gives the impact of the number of participating clients on the experimental results, and the experimental results are shown in Figure 4.

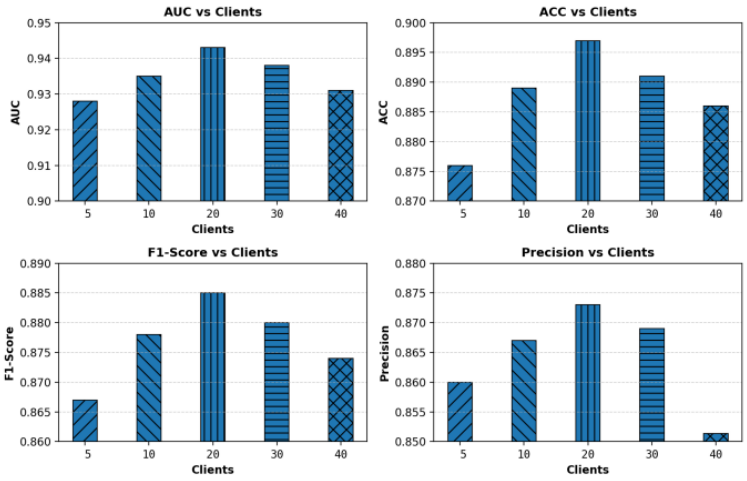

Figure 4. The impact of the number of participating clients on experimental results.

From the results in Figure 4, it can be seen that the number of participating clients has a clear impact on model performance. When the number of clients is small, the performance of AUC, ACC, F1-Score, and Precision is relatively low. This indicates that with limited data distribution, the global model cannot fully learn feature differences from diverse sources, which restricts overall performance. It shows that insufficient clients may hinder federated learning models from forming robust feature representations in cross-institutional collaborative modeling. When the number of clients increases to 20, all four metrics reach their peak. The AUC approaches 0.943, the ACC reaches 0.897, the F1-Score rises to 0.885, and the Precision improves to 0.873. This result shows that with a moderate number of participants, the model can better integrate feature distributions from multiple data

 doi:10.20944/preprints202603.1104.v1

sources. It also demonstrates that the model achieves the best predictive performance under privacy-preserving conditions. This indicates that an appropriate number of clients helps balance data diversity and aggregation stability, which is a critical factor affecting global model performance.

However, when the number of clients further increases to 30 and 40, the model's performance shows a slight decline. All four metrics drop to some extent. This may be due to stronger data heterogeneity caused by too many participants. Larger differences in data distributions across clients make it difficult to eliminate bias during global aggregation. At the same time, excessive clients increase communication and synchronization costs, introducing additional randomness and noise during global updates, which affects convergence.

Overall, the experimental results indicate that the choice of client number is crucial in privacy-preserving federated learning. A moderate number of clients can effectively improve model stability and generalization ability, while too few or too many participants may weaken global model performance. This finding reveals the limiting effect of environmental sensitivity on model performance. It also provides valuable guidance for parameter selection when deploying federated learning systems in practice.

Overall, the experimental results indicate that the choice of client number is crucial in privacy-preserving federated learning. A moderate number of clients can effectively improve model stability and generalization ability, while too few or too many participants may weaken global model performance. This finding reveals the limiting effect of environmental sensitivity on model performance. It also provides valuable guidance for parameter selection when deploying federated learning systems in practice.

This paper also gives the influence of noise injection intensity on the experimental results, and the experimental results are shown in Figure 5.

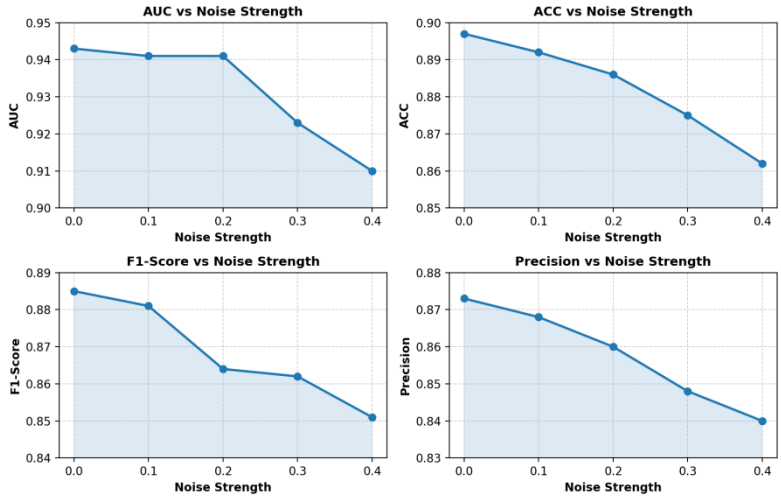

Figure 5. The impact of noise injection intensity on experimental results.

From the results in Figure 5, it can be observed that increasing noise injection intensity has a significant impact on overall model performance. As the noise level rises from 0.0 to 0.4, the AUC gradually decreases from 0.943 to about 0.910. This indicates that excessive noise interferes with the model's ability to distinguish patterns in different feature spaces. Although differential privacy plays an important role in protecting user information, too much noise reduces the capacity to capture abnormal transaction patterns. For the ACC metric, the accuracy also decreases as noise intensity grows. The value drops from 0.897 to 0.862, showing the same downward trend as AUC. This suggests that stronger privacy protection affects the stability of global prediction results, preventing the model from maintaining its original classification advantage. In financial technology applications, balancing data security and model performance is a critical issue when designing privacy-preserving algorithms.

 doi:10.20944/preprints202603.1104.v1

The variation in F1-Score also highlights this trade-off. The experiments show that as noise increases from 0.0 to 0.4, the F1-Score falls from 0.885 to 0.851. The decline is substantial. This means that noise not only reduces recall but also interferes with precision, lowering the model's overall effectiveness in detecting risky transactions. In other words, while noise enhances privacy protection, excessive noise can prevent the model from effectively balancing recognition across different classes. The downward trend in Precision further confirms this observation. As noise intensity rises, Precision decreases from 0.873 to 0.840. This indicates that the reliability of positive predictions is reduced. In real privacy-preserving settings, too much noise weakens the trustworthiness of risk prediction. Therefore, these results emphasize the sensitivity of noise injection intensity. Setting noise at a reasonable level can safeguard data privacy while ensuring model usability and stability in financial applications.

V. Conclusion

This study focuses on privacy protection in financial technology and builds a federated learning-based algorithmic framework. It addresses the conflict between data sharing and regulatory compliance. By introducing distributed training mechanisms and privacy-preserving strategies, the framework enables cross-institutional modeling without exposing raw data. This ensures the security of user information. The overall results show that the method achieves a good balance between data protection and model performance, providing an effective technical path for risk modeling in complex financial scenarios.

From a practical perspective, the proposed method can promote secure collaboration among financial institutions and support the integration and use of cross-domain information. In key tasks such as risk control, fraud detection, and credit evaluation, the framework improves predictive accuracy and robustness while meeting the increasing requirements of data compliance. This is significant for the digital transformation of the financial technology industry and offers a feasible solution for building more reliable data-sharing mechanisms within the sector.

In addition, the privacy-preserving mechanisms reflected in this study have strong generality and scalability. They can be adapted to different types of financial business scenarios. Whether dealing with large-scale user transaction data or heterogeneous account information from multiple sources, the proposed framework can achieve efficient global modeling while keeping data localized. This capability contributes to the development of a secure financial ecosystem and enhances the overall risk management capacity and service level of the industry.

In conclusion, this research not only responds to the dual demands of privacy protection and data utilization in its methodological design but also demonstrates broad application value for the development of financial technology. By improving modeling performance while safeguarding user privacy, the study provides both theoretical and practical support for compliance-oriented innovation in financial technology. The findings have a positive impact on the advancement of intelligent financial services and the improvement of industry-level data governance. They also lay a solid foundation for future expansion to larger and more complex environments.

References

1. He, P., Lin, C., and Montoya, I., "DPFedBank: Crafting a privacy-preserving federated learning framework for financial institutions with policy pillars," arXiv:2410.13753, 2024.
2. Aljunaid, S. K., Almheiri, S. J., Dawood, H., et al., "Secure and transparent banking: explainable AI-driven federated learning model for financial fraud detection," Journal of Risk and Financial Management, vol. 18, no. 4, p. 179, 2025.
3. Ma, C., Zhao, H., Zhang, K., et al., "A federated supply chain finance risk control method based on personalized differential privacy," Egyptian Informatics Journal, vol. 31, p. 100704, 2025.
4. Emmanuel, M., "Federated learning for privacy-preserving financial fraud detection," 2025.
5. D. Shenoy, R. Bhat, and K. Krishna Prakasha, "Exploring privacy mechanisms and metrics in federated learning," Artificial Intelligence Review, vol. 58, no. 8, p. 223, 2025.

 doi:10.20944/preprints202603.1104.v1

6. H. K. Tayyeh and A. S. A. AL-Jumaili, “Balancing privacy and performance: a differential privacy approach in federated learning,” Computers, vol. 13, no. 11, p. 277, 2024.
7. G. Zheng, L. Kong, and A. Brintrup, “Federated machine learning for privacy preserving, collective risk prediction,” International Journal of Production Research, vol. 61, no. 23, pp. 8115–8132, 2023.
8. S. Li, Y. Wang, Y. Xing, and M. Wang, “Mitigating correlation bias via causal modeling and consistency-aware learning,” 2025.
9. F. Hanrui, Y. Yi, W. Xu, Y. Wu, S. Long, and Y. Wang, “Intelligent risk modeling with meta-learning: addressing sample scarcity and evolving patterns,” 2025.
10. A. Xie and W. C. Chang, “Deep learning approach for risk identification using transformer modeling of heterogeneous records,” arXiv:2511.04158, 2025.
11. R. Liu, R. Zhang, and S. Wang, “Transformer-based modeling of user interaction sequences for dwell time prediction,” arXiv:2512.17149, 2025.
12. N. Lyu, F. Chen, C. Zhang, C. Shao, and J. Jiang, “Deep temporal convolutional neural networks with attention mechanisms for resource contention classification,” 2025.
13. Y. Wu, Y. Qin, X. Su, and Y. Lin, “Transformer-based risk monitoring with graph integration,” in Proc. 2025 2nd Int. Conf. Digital Economy, Blockchain and Artificial Intelligence, pp. 388–393, 2025.
14. Q. Gan, R. Ying, D. Li, Y. Wang, Q. Liu, and J. Li, “Dynamic spatiotemporal causal graph neural networks,” 2025.
15. Y. Xing, M. Wang, Y. Deng, H. Liu, and Y. Zi, “Explainable representation learning in large language models for fine-grained classification,” 2025.
16. Q. Zhang, Y. Wang, C. Hua, Y. Huang, and N. Lyu, “Knowledge-augmented large language model agents for explainable decision-making,” arXiv:2512.09440, 2025.
17. S. Pan and D. Wu, “Trustworthy summarization via uncertainty quantification and risk awareness in large language models,” arXiv:2510.01231, 2025.
18. T. Guan, S. Sun, and B. Chen, “Faithfulness-aware multi-objective context ranking for retrieval-augmented generation,” 2025.
19. X. Hu, Y. Kang, G. Yao, T. Kang, M. Wang, and H. Liu, “Dynamic prompt fusion for multi-task and cross-domain adaptation in LLMs,” arXiv:2509.18113, 2025.
20. B. C. Das, S. Mahabub, and M. R. Hossain, “Empowering modern business intelligence tools for data-driven decision-making: innovations with AI and analytics insights,” Edelweiss Applied Science and Technology, vol. 8, no. 6, pp. 8333–8346, 2024.
21. Wang, Y., Fang, R., Xie, A., Feng, H., and Lai, J., "Dynamic anomaly identification in accounting transactions via multi-head self-attention networks," arXiv:2511.12122, 2025.
22. Li, J., Gan, Q., Liu, Z., Chiang, C., Ying, R., and Chen, C., "An improved attention-based LSTM neural network for intelligent anomaly detection in financial statements," 2025.
23. Xu, Z., Cao, K., Zheng, Y., Chang, M., Liang, X., and Xia, J., "Generative distribution modeling for credit card risk identification under noisy and imbalanced transactions," 2025.
24. Gao, K., Hu, Y., Nie, C., and Li, W., "Deep Q-learning-based intelligent scheduling for ETL optimization in heterogeneous data environments," arXiv:2512.13060, 2025.
25. Aurna, N. F., Hossain, M. D., Taenaka, Y., et al., "Federated learning-based credit card fraud detection: performance analysis with sampling methods and deep learning algorithms," Proceedings of the 2023 IEEE International Conference on Cyber Security and Resilience (CSR), IEEE, pp. 180–186, 2023.
26. Shi, J., Siebes, A. P. J. M., and Mehrkanoon, S., "Trans-XFed: An explainable federated learning for supply chain credit assessment," arXiv:2508.13715, 2025.
27. Reddy, V. V. K., Reddy, R. V. K., Munaga, M. S. K., et al., "Deep learning-based credit card fraud detection in federated learning," Expert Systems with Applications, vol. 255, p. 124493, 2024.
28. Li, M., and Walsh, J., "FedGAT-DCNN: Advanced credit card fraud detection using federated learning, graph attention networks, and dilated convolutions," Electronics, vol. 13, no. 16, p. 3169, 2024.

 doi:10.20944/preprints202603.1104.v1

disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.