Exploring Responsible Innovation with Privacy Preservation: Federated Learning Policies for Digital Finance Services in Asia

Prohim Tam[1] and Riccardo Corrado[2]

[1]School of Digital Technologies, American University of Phnom Penh, Phnom Penh 12106, Cambodia
[2]CamEd Business School, Phnom Penh 12200, Cambodia

Abstract

Digital finance is being transformed by advancements in Artificial Intelligence (AI) and FinTech, which offer innovation in personalized financial products, fraud detection, accessibility, and risk management. However, managing sensitive customer data poses significant privacy and security challenges. Traditional centralized data collection methods raise concerns, emphasized by several regulations. By leveraging Federated Learning (FL) capability, this paper guides promising solutions and policies that enable institutions to train AI models locally and share only model updates to reduce data-sharing risks between multiple financial institutions. We aim to advance AI-driven innovation in digital finance while ensuring compliance with privacy regulations, focusing on the practical applications of FL in Asia. The key motivations are rooted in enhancing data security through decentralized AI training, improving regulatory compliance, incentivizing participation, and promoting trustworthy collaboration among financial entities. We identify the best-suited FL approaches for Asia's digital finance infrastructure by exploring various applications, including personalized recommendation models, open banking, fraud detection, and automated credit scoring. By assessing literature and analyzing global FL implementations, our study suggests federated settings and actionable policies. Policy implications include adapting privacy regulations to future challenges, setting security standards for distributed AI in finance, and establishing ethical guidelines for FL utilization, which aim to promote responsible innovation while protecting customer privacy, particularly in developing member countries of Asia.

1. Introduction

1.1. Background and Motivation

The successful advancement of digital finance services relies on the effective application of systems, performance indicators, user satisfaction, and privacy preservation for both domestic and international markets. There are several applications to digitalize financial services in Asia, and China's central bank digital currency initiative can be considered as a primary example of how digital finance can revolutionize economies [1]. Digital currencies are emerging as a modern alternative to traditional cash. By transforming physical money into electronic formats, the system promises to streamline financial transactions, reduce costs, and enhance monetary control. Issued and regulated by central banks, these digital currencies can be stored, spent, and tracked electronically, which offers a new paradigm for how we manage money. However, this innovation is not without its challenges. The legal status of the digital currency, the roles and responsibilities of various stakeholders, and concerns about cybersecurity and data privacy all need careful consideration. Therefore, ensuring privacy and security, gaining user trust, addressing technical complexities, and mitigating economic disruption are crucial for the successful implementation of digital currencies, as well as other digital finance applications.

Artificial Intelligence (AI) models in financial services often require vast amounts of user and transaction data to converge the final model with efficient accuracy in validation. Beyond this, many modern applications face penalties for data leakage while developing their AI models, which highlights the critical need for organizations to adopt more adaptive and responsible financial data transfer, storage,

and sharing practices [2,3]. These requirements have caused a shift towards more secure data handling practices. Ensuring that data is shared in a manner that protects user privacy is now a top priority.

Given these concerns, Federated Learning (FL) has emerged as a popular approach among researchers, practitioners, and standards organizations [4]. Introduced by Google in 2016, FL has experienced significant growth and platform development, which results in several deployment options, applicable use cases, standard frameworks, and (hyper)parameter specification guides. FL offers a solution to many of the privacy and data handling challenges faced by data-centralized AI applications. By allowing models to be trained across multiple decentralized devices or edge servers holding local data samples, FL eliminates the risks of data leakage.

1.2. Paper Contributions

With the abovementioned challenges and motivation, our primary contribution to this work can be presented as follows:

- We first investigate the literature review on four primary applications empowered by FL. From the review, we offer a recommendation for an FL platform to fit in with Asia's financial applications and observe the training dataset, environment setup, experiment tools, and implementation results. We explore on the reward mechanism for contributing data or resources with governance use cases to build trust among institutions.
- We contribute to four policy discussions, including 1) **Policymaker** to initiate legislative and regulatory frameworks, 2) **Financial Institutions** to lead the exploration and integration of privacy-aware FL technologies, 3) **Technology Providers** to ensure responsible development of FL solutions in the future while respecting customer privacy rights when developing AI-based automation features, and 4) **International Organizations** to provide with the legal guidance, technical assistance, funding, and knowledge-sharing platforms for supporting the early stage of FL and its implementation.

The paper is organized as follows. Section II presents the background of FL and its applicability in digital finance. Section III describes our framework for analyzing the FL use cases and platform recommendations. Section IV showcases the performance evaluation and key indicators leading to policy recommendations, as well as the conclusion.

2. FL and Its Applicability in Digital Finance

FL methodologies vary widely and can be configured to specific needs. The main methodologies include asynchronous and synchronous approaches, Independent and Identically Distributed (IID) and non-IID data handling, and (hyper)parameter for model selection. Additionally, FL can be integrated with other privacy-preserving technologies such as homomorphic encryption and differential privacy to further enhance the security of digital finance systems [5].

Each configuration handles specific FL application targets. In asynchronous FL, local model uploads are made as soon as the training is finished, which leads to faster training times but potentially less consistent models. In contrast, synchronous FL involves waiting for updates from all nodes before averaging to the global model, which ensures consistency but slows down the training process. For data types, financial datasets are often non-IID, which presents a challenge for traditional machine learning models. However, FL is well-suited to handle non-IID data by leveraging local data distributions to improve overall model performance. The selection of the right hyperparameters has a crucial role for the success of FL, a process that involves tuning parameters such as learning rates, batch sizes, and the number of training epochs to optimize model performance, based on the criticality of particular financial services. Given the distributed functions of FL, the outcome is better, but the process can be more

complex than the centralized training. For instance, Imteaj et al. [6] leveraged asynchronous FL to address the prediction of customer financial distress by proposing a novel model to handle non-IID, which shows improved prediction accuracy through distributed data across multiple agents. The authors used different batch sizes ranging from 0, 3, 5, and 10. Optimal participant numbers were determined as 9, 10, and 11 for different batch sizes respectively. The authors used a highly imbalanced dataset, called "Give Me Some Credit" from Kaggle, with 150,000 samples. The proposed model outperformed the local mean model and showed close training accuracy to the best local model in every training round. By using asynchronous FL, the authors obtained an efficient model for handling non-IID financial datasets. However, the drawback remains in terms of potential computational overhead due to asynchronous updates, and dependency on agent resource availability which may prolong convergence if agents have limited computational capabilities.

Figure 1 illustrates the procedure states of FL between multiple financial institutions, compared to traditional learning of uploading raw data to the central coordinator/developer. The main procedure of FL can be described in seven consecutive steps in a single round of communication, as follows:

1. The Coordinator/Developer initiates the process by defining a global model that encapsulates the core AI architecture (e.g., recurrent neural networks). The model is designed with specific objectives for financial services, such as predicting credit scores, detecting fraudulent transactions, or generating personalized financial advice. The global model's parameters, including the learning rate, loss function, and evaluation metrics, are pre-defined to guide the subsequent training phases.
2. Once the global model is configured, the system broadcasts to all local participating entities, including users, banks, and other financial institutions. These entities serve as the nodes in the FL network. The global model is distributed securely, which ensures that the model reaches each participant's device or local server without exposing any sensitive information.
3. Each participant, whether an individual user or a financial institution, trains the global model locally using their own dataset. For example, users might use mobile banking data, banks may utilize transaction histories, and other entities might employ additional financial records. This training occurs in a specified timeslot, which allows participants to adapt the model to each unique data environment. The output of this process is a local model that reflects the optimal model (minimized loss) drawn from the participant's data.
4. After local training, the participant's device or server generates the local model parameters (e.g., weights and biases) and sends them to an edge aggregator. The transmission is designed to be efficient and secure, which ensures that no raw data is exchanged – only the computed model updates.
5. The edge aggregator receives local models from multiple participants and adjusts them according to specific service targets. This step involves aligning feature spaces across different datasets or scaling updates based on participant importance. The aggregator then combines these adjusted models into a single aggregated model, which represents a more generalized understanding based on the collective data relevance.
6. The aggregated model from the edge is then sent back to the core entity (global server). This process involves a secure communication channel to ensure that only authorized updates are transmitted, which maintains the integrity of the model aggregation process.
7. The global server collects all aggregated models from various edge aggregators and performs a model-averaging process. Instead of accessing raw data, the server only requires the parameters of each local model, which are averaged to refine the global model. The updated global model is then redistributed in the next training round, which iterates until the model converges to an optimal solution.

The processing flow ensures that sensitive financial data remains within the local environment of each participant, thereby safeguarding privacy while allowing for the AI development of a robust and accurate global model that can be applied across various financial services.

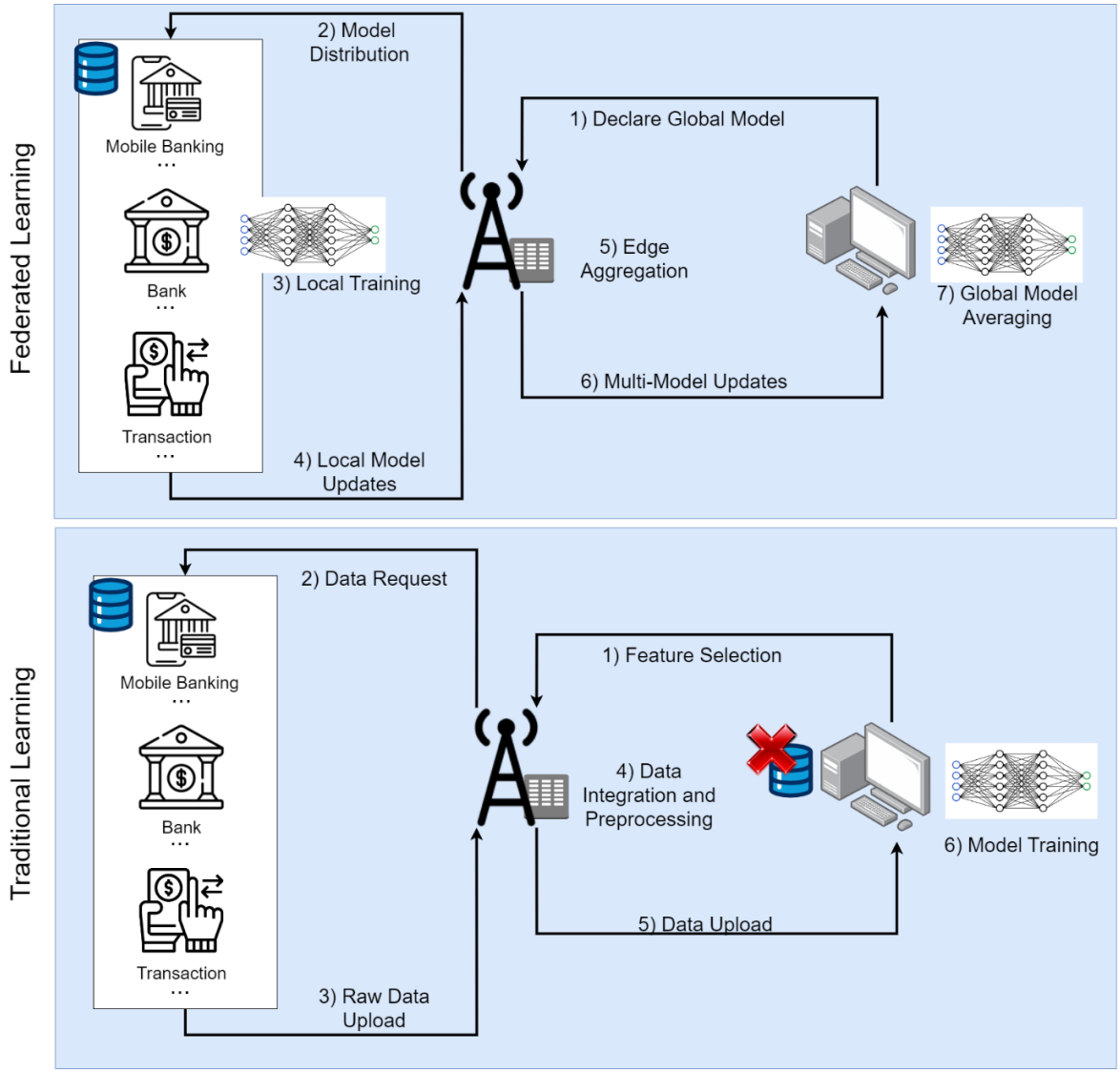

Figure 1. FL Procedure Flows in Digital Finance Learning.

Table 1 presents the most well-known FL platforms used in research and development. From our FL background's observation, we provide recommended use cases in financial scenarios based on the organization's existing IT infrastructure and future goals.

Table 1. FL Platforms and Suitable Use Cases in Digital Finance.

Ref.	Platform	Recommended Use Cases	Developer	Available
[7]	TFF	For financial firms already using TensorFlow that offers ease of integration and deployment across various environments through high-level interfaces	Google	https://www.tensorflow.org/federated
[8]	FATE	For financial services requiring security and compliance, particularly in regulated environments such as banking or industrial-grade platform with secure computation protocols	WeBank AI	https://fate.fedai.org/
[9]	PySyft	For highly privacy-sensitive financial applications requiring advanced security	Open-Minded	https://openmined.github.io/PySyft/

		features such as differential privacy, multiple deployment environments, and secure learning		
[10]	IBM FL	For large financial enterprises focused on the deployment across computing environments and custom fusion algorithms	IBM	https://ibmfl.res.ibm.com/
[11]	NVIDIA FLARE	For financial institutions looking for scalable and lightweight solutions with strong analytics and orchestration capabilities, including training and validation workflows, federated analytics, and management dashboard (Open-source SDK)	NVIDIA	https://github.com/NVIDIA/NVFlare
[12]	FedML	For scalable and secure financial applications across diverse data silos and large models, particularly for firms focusing on MLOps integration	FedML Inc.	https://fedml.ai/home
[13]	Flower	For complex financial institutions requiring large-scale experiments and integration across multiple languages and ML frameworks	Adap ML	https://flower.ai/

As a summary of the FL platforms and their applicability in digital finance, the primary questions to answer are: 1) How to determine FL types and specifications for different use cases of digital finance applications? 2) Which FL settings are better in terms of target objectives, whether to optimize the restricted privacy requirements or accuracy? and 3) What policies do financial institutions consider before training and deploying the final FL model?

3. Framework for Analysis and Practical Guidelines

In this section, we organize by starting to observe the use cases of FL in modern financial services, then we conduct the practical instructions and guidelines for different categories based on simulation platforms. Additionally, we will highlight the importance of incentive mechanisms for encouraging data contribution and computational resource sharing, which showcases how these elements contribute to building trust among institutions.

3.1. Use Case Observation

Four primary applications are investigated before conducting the practical FL framework and policy implications, including 1) Recommendations for consumer services, 2) Open-banking, 3) Credit card fraud, and 4) Automated credit scoring.

3.1.1. Recommendations for Consumer Services

There are several recommendations in FL-empowered financial services, including personalized products, stock prediction, risk assessments, loan projection, and wealth management. In our study, we reviewed two primary use cases – based on the relevance and depth of analysis for assisting the implementation in Asia – as given in Table 2, where it is possible to observe a summary of each study's objectives, platforms, and achieved outcomes using FL integration.

Table 2. Existing FL in Digital Finance's Recommendation Services: Two Selected Cases.

Ref.	Use Case	Objective	Platform	Achieved Outcomes	Year

[14]	FedStock	Investigate the effectiveness of FL for stock market trend prediction compared to traditional centralized and decentralized learning methods (Learning models: Linear Regression, Random Forest, and Support Vector Machine)	Flower with Scikit-Learn on a public dataset containing stock market data from nine key provinces in China	FL outperformed (de)centralized and decentralized learning when using Random Forest or Support Vector Machine models	2023
[15]	FedRisk	Address the inefficiencies in evaluating risks associated with Supply Chain Financing (SCF) and propose a FL framework for order-level risk prediction.	PyTorch for training model on an aerospace supply chain dataset (orders from 2014 to 2022)	FL can enhance prediction accuracy compared to localized models, which is beneficial for SMEs with limited data	2024

In [14], the authors leverage FL for stock prediction on a public dataset from nine provinces between the strongest stock regions in China, namely Hubei, Fujian, Sichuan, Shandong, Beijing, Zhejiang, Jiangsu, Guangdong, and Shanghai. The data cleaning and preprocessing methods were applied to remove missing values, normalize features, and identify correlations between variables. Three machine learning models (Linear Regression, Random Forest, and Support Vector Machine) were trained on the preprocessed dataset using both centralized, decentralized, and FL approaches. A 5-fold cross-validation was employed to evaluate model performance. The key findings can be outlined as follows:

- FL outperformed centralized and decentralized learning when using Random Forest or Support Vector Machine models.
- LR models demonstrated better performance with both centralized and decentralized learning compared to FL.
- SVM underperformed compared to LR and RF due to convergence issues.

This FL use case demonstrated the potential of FL for stock market trend prediction, identified potential limitations of FL in certain scenarios (e.g., Linear Regression model), and provided guidelines for data preprocessing techniques for stock market data. Overall, the study contributes to the understanding of FL's applicability in financial domains and highlights the importance of careful model selection and data preparation for accurate stock market predictions. However, further exploration of FL performance with true parallelism on multi-computer platforms is needed.

Moreover, in [15], the primary goal is to address the inefficiencies in evaluating risks associated with SCF. Traditional methods rely on firm-level data, which is misaligned with the order-level conditions of SCF. To bridge this gap and overcome data limitations, especially for Small and Medium Enterprises (SMEs), the study proposes a FL framework for order-level risk prediction. The study used an aerospace supply chain dataset, provided by a system integrator, consisting of payment records for suppliers on purchase orders from 2014 to 2022. The dataset was filtered to include suppliers with sufficient data (more than 200 order records) for model development. To simulate real-world data accumulation, the data was further partitioned based on order timeline. The key findings can be concluded as follows:

- By leveraging collective data from multiple suppliers without compromising privacy, FL can enhance prediction accuracy compared to localized models, particularly beneficial for SMEs with limited data.
- The FL approach can accurately predict buyers' late payment risk, which is a crucial factor in SCF risk evaluation.

- FL can improve SCF accessibility for SMEs and mitigate the risks associated with SCF for financial institutions.

The benefits of FL are notable for both use cases; however, from our evaluation, the implementation of a token-based reward system to incentivize data contribution should be added on for allowing institutions to earn tokens for the volume and quality of their shared data. A well-balanced management framework must be applied for building trust, which ensures data anonymization and maintains transparency by regularly sharing performance metrics with participating institutions. The coordinator can develop clear collaboration agreements detailing roles and responsibilities and create feedback mechanisms for ongoing communication to strengthen partnerships and enhance model performance.

3.1.2. Open Banking

Open banking, a subset of open innovation, empowers consumers by enabling them to share their financial data through Application Programming Interfaces (APIs) with third-party providers. Open banking supports the inclusion of smaller and medium-sized entities that promote innovative ideas and targeted services for diverse customer segments. This data-sharing revolution has the potential to spark innovation, improve financial services, and enhance competition. However, realizing these benefits requires addressing significant challenges, particularly in data privacy and security.

FL is emerging as a promising solution to these challenges. In the context of open banking, FL can facilitate the development of sophisticated fraud detection models, personalized financial advice, and innovative credit scoring systems. Table 3 presents the different use cases of open banking in the UK, Australia, and China [16].

Table 3. Initiatives of Open Banking: Three Selected Cases

County	Key Approaches	Impact
UK	Mandated data sharing by major banks, creation of regulatory sandbox	Increased competition, emergence of fintech startups, development of new financial products
Australia	Consumer Data Right (CDR) enabling data portability, focuses on consumer empowerment	Enhanced financial literacy, improved access to financial services, potential for personalized products
China	Gradual approach with pilot programs, emphasis on technology collaboration (including Tencent Cloud and WeBank's Fintech Lab)	Potential for significant market disruption, development of innovative fintech solutions

Implementing FL in the open banking ecosystem can be highly complex in practice. Data heterogeneity, derived from various data formats and structures across different financial institutions, poses a substantial barrier. Additionally, the imbalanced conditions of financial data, with a predominance of non-fraudulent transactions, require specialized techniques to build effective and reliable models. Overcoming these challenges necessitates careful consideration of incentive structures, data harmonization strategies, robust communication protocols, and platform selection.

By addressing these complexities, FL can propel open banking to its full potential in the future of Asia's digital finance, offering benefits to consumers and financial institutions. From the analysis of the selected cases, we can draw some key aspects related to FL in open banking:

- Open banking serves as a mechanism for innovation and competition within the financial sector. By enabling third-party developers to access and utilize customer data, open banking fosters the development of new products and services.
- Robust regulatory frameworks that prioritize consumer empowerment and data privacy are essential for shaping a successful open banking landscape.

- FL deployment offers a promising solution to integrate with the regulatory framework; however, specific guidelines to handle data heterogeneity and imbalance are highly significant for FL-assisted open banking. With enabled FL, open banking can utilize robust encryption and secure communication channels for data transmission and model updates, which instill confidence among participating institutions.
- Open banking with feedback mechanisms can create structured loops where institutions can share their experiences, challenges, and suggestions regarding the FL implementation. Regular communication motivates collaboration and a sense of community.
- Collaboration among financial institutions, FinTech companies, and regulators is vital for the successful implementation of FL in open banking.

3.1.3. Credit Card Fraud Detection

Credit card fraud, characterized by unauthorized transactions on credit cards, has become a pervasive issue for financial institutions, which results in substantial financial losses and damage to consumer trust. Given the sensitive nature of financial data and the distributed conditions of credit card transactions across multiple institutions, traditional centralized fraud detection methods face significant challenges. To address the limitations of centralized approaches and protect customer privacy, FL emerges as a promising paradigm for collaborative fraud detection. Table 4 summarizes two primary use cases, namely Starlit and FFD.

Table 4. Existing FL in Digital Finance's Credit Card Fraud Detection.

Ref.	Use Case	Objective	Platform	Achieved Outcomes	Year
[17]	Starlit	Address the limitations of existing FL solutions for financial fraud detection, which include issues with security, scalability, and practicality	Flower on AWS ECS (using Python) on transaction data and customer information	With similar training time and network disk volume usage, the proposed Starlit achieved better peak training memory usage and network file volume usage consumption	2024
[18]	FFD	Accurately identify fraudulent transactions while protecting sensitive customer data	Applying FL on European credit card transactions from September 2013, provided by the ULB ML Group	FFD framework achieved a 10% increase in Area Under the ROC Curve (AUC) compared to traditional FDS methods, which indicates improved fraud detection accuracy	2019

In [17], the authors developed a scalable and privacy-preserving FL mechanism, named Starlit, to enhance financial fraud detection. The study utilizes the Flower framework, while integrating with SecureBoost, Private Set Intersection, and Differential Privacy. The system is deployed on AWS ECS, and the Synthetic dataset is conducted by the global transaction organization. The dataset includes transaction data and customer information. In summary, Starlit successfully addresses the limitations of existing FL solutions in fraud detection as follows:

- Achieving linear scalability with the number of participants, avoiding computationally expensive operations.
- Incorporating all phases of the FL process, including identity alignment, into the implementation and evaluation.
- Designing Starlit to be resilient against client dropouts.

- Enabling secure identification of discrepancies and aggregation of common features among shared users across different datasets.

Additionally, in [18], a robust and effective Fraud Detection System (FDS) was developed for credit card transactions addressing the challenges posed by data imbalance and privacy concerns. The focus is on creating a system that can accurately identify fraudulent transactions while protecting sensitive customer data. A real-world dataset of European credit card transactions was used, and the dataset is highly imbalanced, with only 0.172% of transactions being fraudulent. The study successfully developed FFD that demonstrates significant improvements over traditional FDS methods. Overall, the key findings can be concluded as follows:

- The study incorporates an oversampling approach to mitigate the impact of imbalanced datasets, leading to more robust fraud detection models.
- FL-assisted framework considers factors of communication cost and convergence rate to optimize performance.
- FFD framework offers a promising approach to credit card fraud detection by combining the benefits of FL with effective data handling techniques.

While the study presents promising results, the future research remains on further strengthening privacy protections and evaluating performance on non-IID financial datasets.

3.1.4. Automated credit scoring

The framework in [19] used FL to enhance collaboration, where multiple banks can build accurate credit assessment models without exposing the individual datasets. FL ensures that sensitive customer data remains within the originating bank's infrastructure. By aggregating diverse datasets, FL helps create more holistic and accurate credit scoring models. Furthermore, combined with Explainable AI (XAI), FL provides transparency in the decision-making process, which meets the requirements for explainable and unbiased credit assessments.

The authors in [20] used PySyft to conduct an experiment on X FL and blockchain-based credit scoring systems to tackle the challenges of credit model sharing. The proposed system incorporates several innovative components to address the challenges of credit modeling, as follows:

- **D-SGD Algorithm**: A decentralized Byzantine fault-tolerant Stochastic Gradient Descent algorithm was implemented to optimize the FL process.
- **Incentive Mechanism**: The Shapley value and Delegated Proof of Stake (DPoS) were integrated to calculate and reward participant contributions. The game-theoretic concept can measure the marginal contribution of each participant to the overall outcome and accuracy, which ensure that contributions are fairly recognized. Meanwhile, DPoS, a consensus mechanism commonly used in blockchain technology, allows participants to delegate their voting power to elected representatives. In this study, participants with a higher delegated stake hold greater influence on the final model, thereby incentivizing active and meaningful engagement in the FL process.
- **System Evaluation**: "Give Me Some Credit" dataset was used for initial evaluation, along with six additional credit datasets from various sources (Germany, Taiwan, Australia, and China). Model performance was assessed using accuracy, precision, recall, F1 score, and AUC.
- **System Architecture and Workflow**: 1) A Centralized coordinator calculates contributions and records in blockchain transactions, then 2) Distributed participants perform local model training and send gradients to the coordinator. The increased number of participants leads to longer training times due to increased communication and computation demands.

In summary, FL-assisted automated credit scoring can offer four primary benefits in digital finance services, including:

- **Collaborative modeling and privacy**: promoting multi-party data support for comprehensive credit assessments while ensuring privacy.
- **Technological innovation**: supporting and leveraging blockchain for trustworthiness and XAI for fairness in credit scoring.
- **Observance of regulations**: ensuring reliable and unbiased credit assessments that meet regulatory requirements.
- **Framework development**: proposing an automated credit decision framework, which provides a comprehensive taxonomy of the features and a comparative analysis of the combined architectures.

3.2. Practical FL Findings for Asia's Digital Finance

From the abovementioned use case observation, our findings conclude the practical guides on platform selection based on criticalities, existing works' performance metrics, and output satisfaction. The choice of FL platform should be based on a careful evaluation of an organization's specific needs, including regulatory compliance, data sensitivity, elasticity and scalability requirements, and existing technology infrastructure, as follows:

- **Regulatory Compliance**: **FATE** and **IBM FL** stand out as the top suggestions for regulated financial institutions in Asia, where compliance, security, and data locality are critical. These platforms provide the necessary tools to ensure adherence to regulatory requirements while supporting scalable and secure operations.
- **Data Sensitivity**: **PySyft** is highly recommended for financial institutions focusing on privacy-preserving techniques, an increasingly important consideration in Asia's expanding FinTech sector. Its advanced security features make **PySyft** an optimal platform for organizations handling sensitive data.
- **Elasticity and Scalability**: **Flower** and **FedML** are the most suited for financial organizations that require customizable solutions and the ability to operate across diverse and rapidly evolving markets in Asia. These platforms are particularly beneficial for institutions that need to integrate multiple data sources and ML frameworks while maintaining a scalable infrastructure.
- **Possible Existing Infrastructure Integration**: **TFF** and **NVIDIA FLARE** are particularly well-suited for institutions that already have specific existing platforms in place. **TFF** is ideal for organizations that have systems already deployed using TensorFlow, as it integrates with TensorFlow and Keras, allowing for a smooth transition to FL. This makes it a strong choice for financial institutions that rely heavily on TensorFlow for their ML/DL models and need to extend their capabilities to include FL. **NVIDIA FLARE**, on the other hand, is a powerful option for organizations that require federated analytics and lifecycle management with a focus on scalable, elastic workflows. Its open-source SDK provides the flexibility to build custom FL solutions while offering robust management dashboards for easier orchestration. This is particularly useful for financial institutions that need advanced analytics capabilities and a user-friendly interface for managing complex federated learning deployments.

Beyond this platform selection approach, there are key considerations as follows:

- **Hybrid Approach**: Consider combining the strengths of different platforms through a hybrid approach. For instance, using PySyft for privacy-sensitive data and FATE or IBM FL for large-scale and production-level deployments.
- **Continuous Evaluation**: The FL landscape is rapidly evolving. Regular evaluation of available platforms is essential to stay updated with the latest advancements and select the most suitable option.
- **Data Quality and Preparation**: The success of FL heavily depends on the quality and consistency of data. Data preparation and cleaning are crucial steps before implementing FL.

- **Deployment Cost**: Evaluate the cost implications of each platform, including licensing fees, cloud infrastructure, and maintenance.
- **Community Support**: Consider the availability of community support, documentation, and tutorials for each platform, which are a long-term sufficiency for developers.
- **Interoperability**: Evaluate the platform's ability to integrate with existing systems and data sources.
- **Benchmarking**: Conduct thorough benchmarking to compare the performance of different platforms in real-world scenarios.

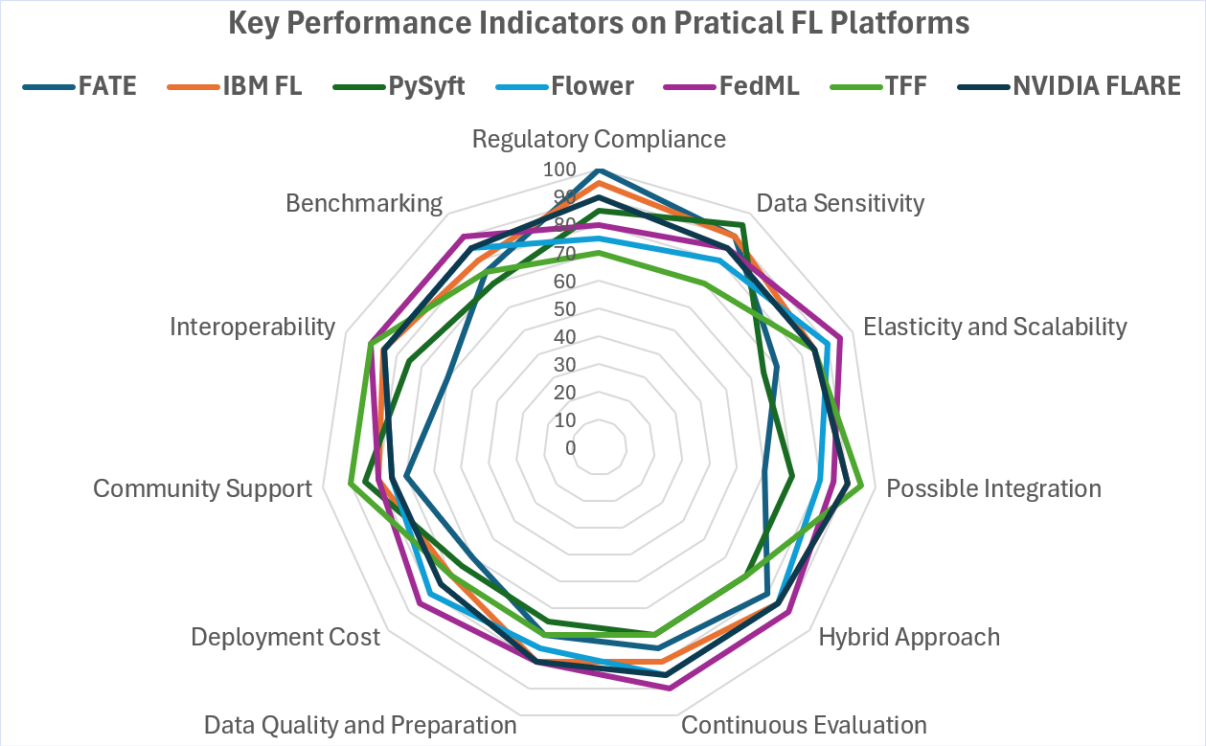

Figure 2. Observation and Weight Adjustment for Key Performance Indicator (KPI)-Based Platform Recommendation.

Figure 2 illustrates the key considerations on each platform. By carefully considering these factors, financial institutions can select the most appropriate FL platform to drive innovation and enhance digital finance operations while ensuring data privacy and security.

4. Conclusion and Policy Recommendation

As digital finance continues to evolve, the integration of FL approaches offers a promising solution to the challenges of data privacy and security. By enabling collaborative model training without the need to centralize data, FL addresses many of the privacy concerns associated with traditional AI models. The ongoing development of FL platforms, methodologies, and use cases will play a crucial role in shaping the future of Asia's digital finance. In this context, institutions and researchers must work together to develop and implement policies that support the adoption of FL. By doing so, we can ensure that digital finance applications are both innovative and secure, paving the way for a more efficient and privacy-preserving financial system.

4.1. Policy Recommendations

Figure 3. The Potential Stakeholder and Respective Roles for Policy Recommendations.

4.1.1. Policymakers

With the role of initiating legislative and regulatory frameworks, the key focus areas are to enable the development and safe use of FL technologies in finance, as follows:

- **Adapting Existing Regulations:** Existing data privacy regulations, such as the EU General Data Protection Regulation (GDPR) [21] and the Personal Data Protection Law in Japan [22], should serve as reference points for developing specific frameworks suitable for the Asian context. The adaptation involves revising data flow models to align with the unique data handling practices of FL and local capabilities in Asia, which clarify data ownership within FL collaborations, and establishing clear legal frameworks for data sharing and liability.
- **Data Ownership and Control**: Policymakers should address issues such as data sharing agreements, participant rights and responsibilities, and mechanisms for resolving disputes in the FL guidelines. For data management and collection efficiencies, policymakers are responsible for creating national strategies to expand broadband infrastructure, especially to rural and underserved areas. They can enact legislation that encourages private sector investment and allocate public funds to bridge the digital divide.
- **Regulatory Sandbox**: A controlled environment for testing FL applications should be established by policymakers in collaboration with relevant financial institutions and industry stakeholders, such as banks, FinTech companies, and academic institutions. The proposed sandbox would provide a safe space for experimentation, which allows participants to test and refine FL models while ensuring robust data security and privacy safeguards are in place.
- **Curriculum Integration**: Policymaker should be responsible for integrating AI and FL topics into national education curricula, which can work with academic institutions and industry experts to design and implement these changes in higher education, vocational training, and certification programs, for propelling the digital talent readiness.
- **Ethical Guidelines**: Policymakers in collaboration with industry stakeholders should establish clear and enforceable ethical guidelines for third-party developers involved in FL projects within the financial sector. These guidelines should address issues such as fairness, transparency, accountability, and non-discrimination in the development and deployment of FL models.

- **Enhanced Network Quality**: Particularly, in developing country members, policymakers should create favorable policies that incentivize telecom operators to upgrade networks and roll out better technologies that support low-latency, high-bandwidth operations. The policy should be bonded with technology providers to innovate and deploy technology solutions, such as network optimization tools and infrastructure upgrades, to improve communication between devices in future privacy-aware AI-empowered systems. Furthermore, policymakers should encourage investment in edge computing by providing tax incentives, subsidies, or co-funding initiatives for the private sector.

4.1.2. Financial Institutions

With the roles of leading the exploration and integration of FL technologies in finance, the key focuses are to ensure the research capability in privacy, handling reward/trust-aware techniques, and cybersecurity, as follows:

- **Research and Development:** Financial regulatory authorities should actively promote research and collaboration by establishing a dedicated working group or forum that includes stakeholders from academia, industry, and policymakers. The proposed group should focus on privacy-aware learning technologies, responsible AI techniques, and the ethical implications of FL in finance. Additionally, funding for research grants should be raised to facilitate the formation of research consortia and organize workshops and conferences to encourage knowledge sharing. The financial institutions should invest in employee training programs to ensure that internal developers and other staff are well-versed in FL technologies and security standards.
- **Industry Collaboration:** Financial institutions should encourage and facilitate collaboration among each other, technology providers, and relevant stakeholders at both national and regional levels. Joint efforts are crucial for developing best practices, mitigating risks, improving the infrastructure setup, and fostering AI innovation with FL solutions, which allows for a comprehensive approach that considers diverse regulatory environments and market conditions.
- **Risk Assessment and Management**: Financial institutions must implement robust risk assessment and management frameworks for all AI development and deployment activities, including FL projects.
- **Cybersecurity Awareness and Training**: Financial institutions should foster a culture of cybersecurity awareness among all employees, by providing regular cybersecurity training programs to educate employees on data security best practices and potential threats.

4.1.3. Technology Providers

With the roles of developing and delivering secure, affordable, and reliable FL solutions for financial institutions, the key focuses are non-biased and well-regulated on data utilization and protection, which can break down the main policies as follows:

- **Comprehensive Security Standards**: Technology providers should mandate the development and implementation of comprehensive security standards for all data management practices within FL projects. These standards should address data security throughout the entire FL lifecycle, encompassing data collection, storage, processing, and transmission.
- **Data Protection Measures**: Third-party developers must implement robust data protection measures to safeguard customer privacy within FL projects. These measures may include data anonymization, pseudonymization, differential privacy techniques, and secure data aggregation methods.
- **Access Controls and Encryption**: Stringent access controls and data encryption measures must be implemented to protect sensitive customer data within FL projects, which includes employing

Role-Based Access Control (RBAC), encryption at rest and in transit, and secure key management practices.

- **Privacy Impact Assessments (PIAs)**: Technology providers should mandate the implementation of PIAs for all FL-based applications developed by third-party providers. These PIAs should thoroughly evaluate potential privacy risks associated with the FL model and propose appropriate mitigation strategies.
- **Transparency and Accountability**: Third-party developers must be transparent about their data collection and usage practices in the context of FL projects. They should also be held accountable for the development and deployment of AI-based features, ensuring they comply with relevant data privacy regulations and ethical guidelines.
- **Affordability and Localized Solutions**: Technology providers should focus on creating cost-effective FL technologies that address the unique needs of developing countries and their limitations in terms of infrastructure and expertise.

4.1.4. International Organizations

With the roles of providing legal guidance, technical assistance, funding, and knowledge-sharing platforms to support the implementation of FL technologies, we can break down the policy recommendations as follows:

- **Technical Assistance and Infrastructure Investment**: The international organizations shall support developing countries by providing resources to build the necessary digital infrastructure, including improving internet access and edge computing capabilities. Furthermore, funding on research initiatives should be considered to focus on exploring privacy-aware FL technologies for financial applications, with a focus on developing regions.
- **Capacity Building and Knowledge Sharing:** The offer on capacity-building programs should be improved within local expertise in AI and FL, helping developing countries integrate FL technologies in their financial systems. The collaboration with financial institutions and technology providers shall be improved by collaborating with educational institutions to fund training programs and workshops for the staff.
- **Ethical and Legal Guidance**: International organizations should closely collaborate with policymakers and technology providers to ensure ethical standards are maintained across regions, focusing on international best practices in data privacy and security. In consultation with regulators and technology providers, the standard organizations should develop harmonized FL models to ensure interoperability, particularly for cross-border FL applications.
- **Data Subject Rights**: Regulations should be updated to ensure that data subjects (customers) retain their fundamental data rights (e.g., access, rectification, erasure), particularly within the cross-region FL context. This may require the development of novel mechanisms to enable individuals to exercise these rights in a distributed FL environment.

Reference

[1] Huang, R. H., & Li, X. (2023). China's pursuit of central bank digital currency: Reasons, prospects, and implications. Banking and Finance Law Review, 39(3).

[2] Martin, K. (2018). The penalty for privacy violations: How privacy violations impact trust online. Journal of Business Research, 82.

[3] Wolff, J., & Atallah, N. (2021). Early GDPR penalties: Analysis of implementation and fines through May 2020. Journal of Information Policy, 11.

[4] McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. Y. (2017). Communication-efficient learning of deep networks from decentralized data. In Proceedings of the International Conference on Artificial Intelligence and Statistics.

[5] Mittal, S., Jindal, P., & Ramkumar, K. (2021). Data privacy and system security for banking on clouds using homomorphic encryption. In 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 1-6.
[6] Imteaj, A., & Amini, M. H. (2022). Leveraging asynchronous federated learning to predict customers' financial distress. Intelligent Systems with Applications, 14, 200064.
[7] TensorFlow/Federated. (n.d.). Available from: https://github.com/tensorflow/federated.
[8] FederatedAI/FATE. (n.d.). Available from: https://github.com/FederatedAI/FATE.
[9] Ziller, A., Trask, A., Loardo, A., Wagner, B., Nounahon, J., Passerat-Palma, J., et al. (2021). PySyft: A library for easy federated learning. In Federated Learning Systems (pp. 111–139). Springer. https://doi.org/10.1007/978-3-030-70604-3_5
[10] Ludwig, H., Baracaldo, N., Thomas, G., Zhou, Y., Anwar, A., Rajamoni, S., et al. (2020). IBM federated learning: An enterprise framework white paper V0.1. arXiv preprint. https://doi.org/10.48550/arXiv.2007.10987
[11] Roth, H. R., Chen, Y., Wen, Y., Yang, I., Xu, Z., Hsieh, Y., et al. (2023). Nvidia FLARE: Federated learning from simulation to real-world. arXiv preprint. https://doi.org/10.48550/arXiv.2210.13291
[12] He, C., Li, S., So, J., Zeng, X., Zhang, M., Wang, H., et al. (2020). FedML: A research library and benchmark for federated machine learning. arXiv preprint. https://doi.org/10.48550/arXiv.2007.13518
[13] Beutel, D. J., Topal, T., Mathur, A., Qiu, X., Parcollet, T., de Gusmao, P., et al. (2022). Flower: A friendly federated learning research framework. arXiv preprint. https://doi.org/10.48550/arXiv.2007.14390
[14] Pourroostaei Ardakani, S., Du, N., Lin, C., Yang, J.-C., Bi, Z., & Chen, L. (2023). A federated learning-enabled predictive analysis to forecast stock market trends. Journal of Ambient Intelligence and Humanized Computing, 14, 4529–4535.
[15] Kong, L., Zeng, G., & Brintrump, A. (2024). A federated machine learning approach for order-level risk prediction on supply chain financing. International Journal of Production Economics, 268, 109095. https://doi.org/10.1016/j.ijpe.2023.109095
[16] Long, G., Tan, Y., Jiang, J., & Zhang, C. (2020). Federated learning for open banking. In Lecture Notes in Computer Science (pp. 240–254).
[17] Abadi, A., Doyle, B., Gini, F., Guinamard, K., Murakonda, S. K., Liddell, J., Mellor, P., Murdoch, S. J., Naseri, M., Page, H., Theodorakopoulos, G., & Weller, S. (2024). Starlit: Privacy-preserving federated learning to enhance financial fraud detection. IACR Cryptology ePrint Archive.
[18] Yang, W., Zhang, Y., Ye, K., Li, L., & Xu, C.-Z. (2019). FFD: A federated learning-based method for credit card fraud detection. In Lecture Notes in Computer Science (pp. 18–32).
[19] Jovanovic, Z., Hou, Z., Biswas, K., & Vallipuram, M. (2024). Robust integration of blockchain and explainable federated learning for automated credit scoring. Computer Networks, 110303–110303.
[20] Yang, F., Abedin, M. Z., & Hajek, P. (2023). An explainable federated learning and blockchain-based secure credit modeling method. European Journal of Operational Research.
[21] European Parliament, Council of the European Union. (2016). General data protection regulation (GDPR). Technical Report, The European Parliament and The Council of The European Union. https://eur-lex.europa.eu/eli/reg/2016/679/oj
[22] Sakai, M., & Oshima, Y. (2020). The enforcement of personal data protection law in Japan. Global Privacy Law Review, 1(3).