Federated Learning for Privacy-Preserving Fraud Detection in Digital Banking: Balancing Algorithmic Performance, Privacy, And Regulatory Compliance

MICHAEL FRIDAY UMAKOR[1], IKECHUKWU IHEANYI[2], UGOCHUKWU DANIEL OFURUM[3], UGOCHUKWU HENRY BEN IBECHEOZOR[4]
[1] *CISSP, CCSP - Western Illinois University, School of Computer Science, IL USA.*
[2]*MBA, Data analytics, University of New Haven.*
[3] *University of Wisconsin-Madison*
[4]*Darden School of Business, University of Virginia*

Abstract- ***The rapid growth of digital banking has heightened concerns over cybersecurity, privacy, and regulatory compliance, particularly in the detection and prevention of financial fraud. Traditional centralized machine learning approaches to fraud detection are limited by data privacy regulations and the increasing complexity of cyber threats. Federated Learning (FL), a decentralized machine learning technique, offers a promising alternative by enabling multiple institutions to collaboratively train models without sharing raw data. This study critically evaluates the application of FL in privacy-preserving fraud detection within the banking sector, focusing on algorithmic performance, privacy implications, and regulatory compliance. The paper reviews existing literature, assesses technical challenges such as data heterogeneity and communication overhead, and presents case studies of FL implementation in real-world banking contexts. The findings reveal that FL significantly enhances privacy and regulatory alignment while maintaining competitive fraud detection performance. The study concludes by offering strategic recommendations for digital banks and regulatory bodies and identifies future research directions that emphasize adaptive learning algorithms, robust evaluation frameworks, and long-term federated infrastructure in financial systems.***

Indexed Terms- ***Federated Learning; Fraud Detection; Digital Banking; Privacy; Regulatory Compliance; Machine Learning***

I. INTRODUCTION

Digital banking has revolutionised the financial sector, providing users with the convenience of conducting transactions, managing accounts, and accessing financial services from anywhere at any time [1-2]. . This transformation has been fueled by the widespread adoption of smartphones, mobile apps, and online platforms [3-4]. However, with this innovation comes an increasing risk of fraud. Digital banking fraud, encompassing activities such as identity theft, account takeover, and payment fraud, has surged as cybercriminals exploit the growing number of online financial services. According to the European Central Bank (ECB), fraud losses in digital banking have reached unprecedented levels, underscoring the need for more robust fraud detection mechanisms.

Fraud detection systems in digital banking are designed to identify and prevent unauthorized transactions, account breaches, and identity theft [65]. Traditionally, machine learning algorithms have been leveraged to detect anomalies and predict fraudulent activity based on transactional data. However, these systems often face challenges in terms of performance, scalability, and, most critically, privacy. The growing regulatory pressure to ensure the protection of customer data adds complexity to developing effective fraud detection mechanisms.

A studied work from Wang et al. [5] indicated that as digital banking becomes more pervasive, privacy and regulatory compliance have become key

considerations. Consequently, financial institutions are required to comply with stringent data protection regulations such as the General Data Protection Regulation (GDPR) in the European Union, the California Consumer Privacy Act (CCPA), and various other country-specific frameworks [6]. These regulations mandate that personal and sensitive financial data be handled in a way that respects users' privacy rights, and failure to comply can result in substantial fines and reputational damage [7]. Moreover, privacy concerns have emerged as a significant barrier to the deployment of effective fraud detection systems. Traditional centralized fraud detection methods involve the collection and processing of vast amounts of personal data, potentially exposing it to security breaches. In this context, privacy-preserving techniques are crucial in maintaining both regulatory compliance and the trust of banking customers.

According to Rafi et al. [8], federated learning (FL) represents an innovative approach to machine learning that addresses the privacy concerns inherent in traditional centralized data processing systems. Unlike conventional machine learning, where data is collected in a central server for training, federated learning allows models to be trained directly on users' devices or local servers, without the need to transfer sensitive data to a central repository [9]. This decentralization minimizes data exposure and supports compliance with privacy regulations.

In the context of digital banking, Aljunaid et al. [10] remarked that federated learning has the potential to revolutionize fraud detection systems by enabling financial institutions to collaboratively build models without sharing sensitive customer data. By training fraud detection models locally and aggregating the insights in a secure and privacy-preserving manner, federated learning can enhance the security of digital banking systems while ensuring compliance with privacy laws.

This study aims to explore the potential of federated learning for privacy-preserving fraud detection in digital banking. Specifically, the study seeks to address the following research questions:

- How can federated learning be applied to develop privacy-preserving fraud detection systems in digital banking?
- What are the trade-offs between algorithmic performance, privacy, and regulatory compliance when using federated learning for fraud detection?
- What are the challenges in achieving high accuracy in federated learning models for fraud detection, and how can these be mitigated?
- How can financial institutions balance the need for accurate fraud detection with the requirements of data privacy and regulatory compliance?

The main aim of this research is to identify and analyze the effectiveness of federated learning as a solution to the challenges of fraud detection while ensuring privacy and meeting regulatory standards.

This study is significant for several reasons. First, it provides valuable insights into how federated learning can improve fraud detection in digital banking while mitigating privacy risks. By exploring the balance between algorithmic performance, privacy, and regulatory compliance, this research will contribute to the development of more secure and efficient fraud detection systems. Additionally, it will inform banking institutions about the feasibility of adopting federated learning to enhance their security infrastructure without compromising customer privacy. Also, the findings of this study will be beneficial to regulatory bodies, providing them with a deeper understanding of the capabilities of federated learning in adhering to privacy laws and ensuring the protection of user data. Finally, this research will serve as a foundation for future studies on the intersection of machine learning, privacy, and regulatory compliance in the financial sector.

II. BACKGROUND AND LITERATURE REVIEW

A. Traditional Fraud Detection Methods

Traditional fraud detection methods in banking primarily rely on rule-based systems, anomaly detection, and supervised machine learning techniques, all of which require the aggregation of transaction data in a centralized system for analysis.

For example, in their study, Susmitha and Kalpana (2025) highlighted the growing limitations of rule-based systems in banking fraud detection. These traditional methods are rigid and often fail to recognize complex or evolving fraud techniques. The authors point out that such systems frequently produce high false positives, lack scalability, and require constant manual updates based on expert input. As fraud patterns become more sophisticated, relying solely on predefined rules is no longer sufficient. Their study supports the need for alternative methods, particularly machine learning models, which offer greater adaptability, improved accuracy, and the ability to learn from data-driven patterns.

Rule-based systems use predefined rules to identify potential fraudulent activity based on known patterns. For example, large withdrawals or transactions in geographically distant locations might trigger alerts [12]. On the other hand, Hilal et al. [13] indicated that anomaly detection identifies transactions that deviate from an individual's typical behavior. However, some studies have shown that while these systems have been effective in detecting certain forms of fraud, they suffer from limitations such as high false-positive rates, a lack of adaptability to new fraud patterns, and the inability to handle large-scale and complex datasets [14-15].

Machine learning (ML) models, particularly supervised learning, have become more prominent in recent years, using labelled datasets to train models for detecting fraudulent transactions. Techniques such as decision trees, random forests, and neural networks have been applied to improve accuracy [16,17,69]. However, traditional models often face scalability issues, difficulty in handling unbalanced datasets (fraudulent transactions are typically much fewer than legitimate ones), and the risk of overfitting when exposed to limited or outdated training data [18,68]. The implication is that an increasing sophistication of fraud requires more dynamic and scalable systems, beyond what traditional methods offer. As such, banks must move towards more advanced and flexible solutions like machine learning models, while also addressing their limitations, such as data privacy and regulatory compliance.

B. Limitations of Centralized Machine Learning in Fraud Detection

Centralized machine learning (CML) models, although more effective than traditional methods, present significant challenges when applied to fraud detection in digital banking. The core issue lies in the centralization of data, financial institutions must aggregate vast amounts of sensitive customer data, such as transaction history, location, and personal identification information, to train machine learning models [19]. . This process not only raises privacy concerns but also creates vulnerabilities in the system, as breaches or leaks of centralised data could lead to massive financial and reputational damage [20]. Moreover, centralised machine learning models often struggle with data imbalance, where fraudulent transactions make up a tiny fraction of the data. This results in poor performance in detecting fraud, as models trained on such imbalanced data often fail to recognise the minority class (fraudulent transactions) effectively [21,66]. Further complicating matters, the high computational costs and the need for large-scale data storage make centralised systems increasingly unsustainable. The implication is that the centralized machine learning models, though useful, are not viable in the long term due to their reliance on massive amounts of personal data. This necessitates the need for more decentralised, privacy-preserving approaches that retain accuracy while mitigating the risks of data exposure.

C. Overview of Federated Learning

Federated learning (FL) is a decentralized machine learning approach that addresses the challenges associated with centralized data collection and processing. Unlike traditional machine learning, where data is collected in a central server, federated learning allows data to remain on local devices, and models are trained collaboratively across these devices. After local models are trained, only model updates (not raw data) are aggregated and sent to a central server for further refinement [22]. This approach significantly reduces the risk of data exposure, which is particularly valuable for sensitive domains like banking.

Based on some works such as the one published by Mohammadi et al. [23], federated learning is accepted as a system that offers several advantages in

fraud detection. It allows financial institutions to improve their fraud detection capabilities by leveraging data from various sources (such as mobile apps and online banking) without violating privacy regulations. Additionally, federated learning helps maintain the integrity and privacy of users' data, ensuring that sensitive information is not shared across platforms, reducing the risks of breaches [25-26]. However, some independent studies by some investigators such as Agripina et al. [26], Barona et al. [27] and Bhanbhro et al. [28] indicated that FL also presents challenges, including the need for robust communication protocols, managing data heterogeneity across devices, and handling situations where the data is imbalanced. This implies that although federated learning can provide a promising solution for improving fraud detection in a manner that respects privacy and complies with regulations. However, its successful implementation will require addressing technical challenges, such as data imbalance and the efficient aggregation of decentralized model updates.

D. Privacy-Preserving Techniques in Banking

Privacy-preserving techniques are critical in the context of fraud detection in digital banking. According to Xu et al. [29] common approach is differential privacy, which ensures that individual data cannot be identified through aggregated statistical results because the approach adds noise to the data in a controlled manner to prevent the identification of specific records while maintaining overall statistical accuracy. Also, Zhu & Niu [30], agreed that another approach is homomorphic encryption, which allows computations to be performed on encrypted data without decrypting it first p30]. These techniques ensure that even if the data is compromised, it remains useless to an attacker.

In the banking sector, privacy-preserving techniques are crucial not only for protecting customer data but also for complying with legal and ethical standards. Techniques like secure multi-party computation (SMPC) and federated learning (as mentioned above) enable financial institutions to develop fraud detection models without compromising the privacy of customer data. In their work, Wulan [31], examined the importance of protecting customer data in banking, especially in light of increasing digital transactions. She highlights that privacy-preserving techniques are vital not only for data security but also for meeting legal and ethical obligations. Although existing laws—such as Law No. 10 of 1998, POJK No. 11/POJK.03/2022, and Law No. 27 of 2022—provide a legal basis, Wulan notes they lack clear, enforceable measures for banks. Her research calls for stronger regulations and technical safeguards to help banks implement effective risk management and protect personal data. The development of privacy-preserving techniques, including differential privacy and homomorphic encryption, plays a pivotal role in the shift towards more secure, scalable, and ethical fraud detection models. Their integration with federated learning could help banks build models that detect fraud while respecting customer privacy.

E. Regulatory Framework for Fraud Detection (GDPR, CCPA, etc.)

The regulatory landscape for data privacy is one of the most critical aspects of fraud detection in digital banking. Regulations like the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA) in the United States set strict guidelines on how customer data can be collected, stored, and processed. Under these regulations, financial institutions must obtain explicit consent from users before collecting their data, provide transparency about data usage, and ensure that customers have the right to request data deletion or correction. For fraud detection systems, regulatory compliance means that any solution must minimize the amount of personal data collected and ensure that data is processed securely and transparently. Federated learning, as a privacy-preserving approach, aligns well with these regulations by allowing data to remain on local devices, ensuring compliance while still enabling effective fraud detection.

The implication is that the regulatory frameworks around data privacy have necessitated the development of fraud detection systems that prioritize user privacy. The ability of federated learning to ensure regulatory compliance while enhancing fraud detection makes it a valuable tool for financial institutions aiming to avoid the consequences of non-compliance.

F. Previous Work on Federated Learning in Financial and Fraud Detection

Several studies have explored the use of federated learning in fraud detection within the financial sector. Yang et al. [32] examined the potential of federated learning to enhance the performance of fraud detection systems by training models on decentralized datasets without violating data privacy. Their findings demonstrated that federated learning outperforms traditional centralized systems in terms of maintaining privacy and security while achieving comparable or better performance in detecting fraudulent transactions. Also, Auma et al. [33], highlighted the use of federated learning in detecting credit card fraud, illustrating its ability to handle imbalanced datasets and improve prediction accuracy. They emphasized the scalability of federated learning, allowing multiple financial institutions to collaborate and build more robust fraud detection systems without sharing sensitive customer data. However, challenges such as data heterogeneity, communication efficiency, and the need for secure aggregation techniques were identified as areas for further improvement. Also, Salam et al. (2024) presented an innovative approach to fraud detection in the financial sector through the application of federated learning. Due to data privacy constraints, traditional centralized models face challenges in accessing diverse transactional data. Federated learning allows banks to collaboratively train fraud detection models without sharing sensitive information, ensuring privacy compliance. Their study compares TensorFlow Federated and PyTorch, with PyTorch offering higher accuracy but requiring more computational time. The research also addresses class imbalance using hybrid resampling methods, which significantly improved model performance. Random Forest emerged as the most effective classifier. This work highlights federated learning's potential in enhancing secure, accurate fraud detection in financial institutions.

Implication: Previous studies highlight the promising potential of federated learning for fraud detection in banking, particularly in addressing privacy concerns and improving performance. However, challenges related to data heterogeneity and communication efficiency need to be tackled for broader adoption in the financial sector.

Finally, traditional fraud detection methods, while still in use, face significant limitations that hinder their scalability and effectiveness in a rapidly evolving digital banking landscape. Centralized machine learning models introduce additional privacy and regulatory concerns. Federated learning offers a decentralised, privacy-preserving alternative that can help overcome many of these challenges. Privacy-preserving techniques and regulatory compliance frameworks further support the need for secure and ethical fraud detection solutions. As such, federated learning presents a promising direction for the future of fraud detection in digital banking, although more research is needed to overcome technical challenges and optimize its implementation.

III. FEDERATED LEARNING IN DIGITAL BANKING

A. Concept and Principles of Federated Learning

Federated Learning (FL) is a decentralized machine learning paradigm that enables multiple entities to collaboratively train a shared model without exchanging raw data. In this approach, each participant trains the model locally on their data and only shares model updates (e.g., gradients or parameters) with a central server, which aggregates these updates to form a global model. This method preserves data privacy and reduces the risk of data breaches, making it particularly suitable for sensitive domains like digital banking.

B. Federated Learning vs. Centralized Machine Learning Models

Traditional centralized machine learning (CML) models require aggregating data from various sources into a central repository for training. While this approach can leverage diverse datasets, it poses significant privacy risks and challenges in complying with data protection regulations. In contrast, FL keeps data localized, thus enhancing privacy and reducing the risk of data leakage. Additionally, FL can handle data heterogeneity and is more scalable in environments where data is distributed across multiple entities.

C. How Federated Learning Addresses Privacy in Fraud Detection

According to Sun [35], FL addresses privacy concerns in fraud detection by ensuring that sensitive customer data remains on local servers. The implication is that only model updates, which do not contain raw data, are shared with the central server [36]. This approach aligns with data protection regulations like GDPR and CCPA, as it minimizes data exposure and enhances security. Moreover, FL can be combined with other privacy-preserving techniques, such as differential privacy and secure multi-party computation, to further protect sensitive information.

D. Data Locality and Decentralized Model Training

Based on the reports from Dritsas & Trigka [37], data locality in FL ensures that data remains within its source environment, reducing the need for data transfer and storage in centralized locations. This is particularly beneficial in digital banking, where data is sensitive and subject to strict regulatory controls. Decentralized model training allows financial institutions to collaboratively improve fraud detection models without compromising data privacy. This approach also reduces latency and can lead to faster model updates and deployment.

E. Use Cases of Federated Learning in Fraud Detection

Several financial institutions have adopted FL for fraud detection. For instance, Google Cloud and Swift have collaborated to develop FL-based anti-fraud technologies for cross-border payments, enhancing security while preserving data privacy. Additionally, studies have demonstrated the effectiveness of FL in detecting credit card fraud, showing that FL models can achieve comparable or superior performance to centralized models while maintaining data privacy.

F. Benefits and Limitations of Federated Learning in Banking

Federated Learning (FL) offers several significant advantages for fraud detection in digital banking. One of the most crucial benefits is enhanced privacy, as FL ensures that sensitive customer data remains within local servers and is never shared directly with a central entity. This greatly reduces the risk of data breaches and unauthorized access [38]. Furthermore, by limiting the movement of personal data, FL enables financial institutions to attain stringent regulatory requirements like, the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), thereby facilitating better compliance. Another notable advantage is the collaborative nature of model training, allowing multiple institutions to contribute to and benefit from a more robust and generalized fraud detection model without revealing proprietary or sensitive information. In addition, FL exhibits strong scalability potential, as it is capable of managing large volumes of distributed data across multiple banks or nodes without compromising system efficiency or accuracy.

Despite these strengths, FL is not without limitations. Implementing such a decentralized system is complex and demands advanced technical infrastructure, as well as careful coordination between participating entities to synchronize model updates and maintain consistency. Another challenge lies in data heterogeneity—differences in data quality, format, and distribution across institutions can introduce biases or reduce model performance. Communication overhead is also a concern, as frequent exchanges of model parameters or gradients between local nodes and the central aggregator may result in increased network traffic and latency, which can hamper real-time fraud detection capabilities. To contextualize these insights, several case studies from recent literature are presented in Table 1.

These case studies demonstrate the practical viability of FL in fraud detection within diverse financial environments. The use of FL across different institutions has consistently shown promise in balancing model accuracy with privacy protection. Moreover, use cases such as those involving mobile banking illustrate the versatility of FL in real-time, decentralized contexts.

Table 1: Case Studies on Federated Learning Applications in Fraud Detection

Case Study	Description	Outcome	Reference
FL for Credit Card Fraud Detection	Simulated federated learning on synthetic credit card data to assess detection accuracy.	Achieved comparable performance to centralized models with improved privacy preservation.	[39]
Cross-Bank FL Collaboration	Collaborated to detect fraud in cross-border payments using FL while maintaining data sovereignty.	Enhanced real-time fraud identification and regulatory compliance.	[40]
Federated Transfer Learning for Banks	Implemented FL with transfer learning to improve fraud detection in underrepresented banks with less data.	Improved performance in low-data banks and maintained privacy.	[41]
Privacy-Preserving FL in Mobile Banking	Explored FL on mobile banking apps to detect anomalous user behavior indicative of fraud.	Detected fraud with high precision while preserving on-device data privacy.	[42]

The results imply that broader adoption of FL in the financial sector could enhance fraud mitigation strategies without compromising user trust or regulatory alignment. Future research should continue to explore ways to overcome its limitations, particularly in the areas of interoperability, communication efficiency, and federated optimization.

Table 2 highlights the trade-offs between centralized and federated learning approaches in fraud detection. While centralized learning can achieve high model performance with diverse data, it poses significant privacy risks and regulatory challenges. Federated learning, on the other hand, offers enhanced privacy and compliance benefits but comes with increased complexity and communication overhead. Institutions must weigh these factors when choosing the appropriate approach for their fraud detection systems.

The adoption of FL in digital banking presents a promising avenue for enhancing fraud detection while maintaining data privacy and regulatory compliance. However, challenges related to implementation complexity and data heterogeneity must be addressed. Future research should focus on developing standardized protocols and frameworks to facilitate the adoption of FL in the financial sector.

Table 2: Comparison of Centralized and Federated Learning in Fraud Detection

Aspect	Centralized Learning	Federated Learning
Data Privacy	Low	High
Regulatory Compliance	Challenging	Easier
Scalability	Limited	High
Implementation Complexity	Moderate	High
Communication Overhead	Low	High
Model Performance	High (with diverse data)	Comparable (with collaborative training)

IV. PRIVACYANDCOMPLIANCE CONSIDERATIONS

A. Privacy Concerns in Digital Banking

F Digital banking systems routinely collect sensitive customer data—transaction histories, behavioral biometrics, login patterns, device data, and geolocation, which, if improperly handled, can lead to serious breaches of privacy. Traditional centralized machine learning models often require data aggregation, increasing vulnerability to insider threats, cyberattacks, and data misuse [43,67]. According to Zarsky [44], financial institutions have become data-rich but privacy-poor, as the benefits of surveillance-driven analytics often outweigh ethical constraints in practice. The concern is not only technical but also ethical. Consumers expect discretion in the handling of their financial footprints. The breach of trust in digital banking due to privacy violations can result in reputational damage and significant financial losses. The rise in high-profile data breaches such as the Capital One hack emphasizes that customer data centralization is an inherent vulnerability [45].

B. Privacy-Preserving Mechanisms in Federated Learning

FL offers an architectural innovation by enabling data to remain on user devices or local institutional servers while allowing collaborative model training. Several techniques have been adopted to reinforce this privacy promise:

- Differential Privacy (DP) introduces statistical noise to model updates, preventing the leakage of individual data records [46].
- Secure Multiparty Computation (SMPC) and Homomorphic Encryption allow encrypted model parameter exchanges without exposing raw data [47].

These mechanisms have been empirically tested in banking contexts. For example, Rahaman et al. [48], applied FL with DP in a simulated fraud detection environment and demonstrated a 35% reduction in privacy leakage metrics while maintaining over 92% model accuracy. However, complex integration increases as privacy mechanisms become more robust.

The implication here is two-fold: while privacy-preserving mechanisms can align with regulatory expectations, they can also introduce computational overhead and model underfitting. Therefore, careful calibration is necessary to optimize privacy budgets and learning efficiency.

C. Regulatory Compliance Requirements (GDPR, CCPA, etc.)

The General Data Protection Regulation (GDPR) in the EU and the California Consumer Privacy Act (CCPA) in the US demands that organisations prioritise data minimization, user consent, and also right to data deletion. These laws have significant implications for fraud detection systems. FL aligns well with GDPR Article 5(c), which emphasizes minimizing data collection. By training models without centralizing raw data, FL offers a structural compliance advantage. According to Truong et al. [49], FL reduces the legal liability associated with data transfers and breaches under GDPR by ensuring data remains with the data controller. However, compliance is not automatic. For example, GDPR requires explainability under Article 22. FL's distributed model training can obscure interpretability, making regulatory audits more complex [50-51]. Similarly, CCPA's opt-out mechanisms may impact the completeness of FL training data, potentially degrading model performance [52]. These issues suggest that while FL is privacy-conscious by design, it must be supplemented with compliance auditing tools, consent management systems, and explainable AI modules to fully meet legal expectations.

D. Privacy vs. Performance Trade-offs in Federated Learning

One of the central tensions in FL lies in balancing privacy with model performance. Empirical findings suggest that the addition of noise (via DP) or encryption (via SMPC) often results in increased latency and reduced predictive accuracy [58].

Hard et al. [59], studied FL performance in Google's GBoard application and noted that model convergence slowed by 50% when strong DP settings were applied. Similarly, Zhu et al. [60], demonstrated that under non-IID (non-identically distributed) banking data scenarios, performance drops of 3–7%

were observed when SMPC was introduced. A similar study by Lu et al. [61] confirmed similar findings.

For fraud detection, where precision and recall are critical, these trade-offs become consequential. A false negative could mean millions in undetected fraud, whereas a false positive could harm a legitimate customer's experience.

The implication is that the design of FL systems must be context-sensitive. In high-stakes domains like banking, hybrid approaches—combining centralized learning for generic features and FL for sensitive features—may be optimal.

E. Implementing Federated Learning while Ensuring Compliance

Successful FL deployment in financial institutions requires a multidimensional approach that integrates technology, legal interpretation, and policy frameworks. According to Kairouz et al. [62], key considerations for compliant FL implementation include:

- Model governance frameworks to ensure traceability and auditability.
- Federated monitoring to detect anomalies or malicious participants.
- Consent orchestration to manage user rights across data silos.

An example is the SWIFT-FL initiative, which combines FL with federated analytics across banks, ensuring that data sovereignty is respected while fraud signals are efficiently shared (Google Cloud, 2025). This model could serve as a template for cross-border collaboration under varying legal regimes.

Implementation also requires robust stakeholder training, investment in secure aggregation servers, and alignment with national financial regulations such as Nigeria's Data Protection Act (NDPA), which mirrors many GDPR principles.

F. Case Studies: Privacy and Regulatory Issues in Banking

To provide a clearer understanding of how privacy and compliance concerns are addressed in real-world applications of Federated Learning (FL) within the banking sector, Table 3 presents selected case studies from various organizations. These cases demonstrate the practical implementation of FL techniques, the mechanisms used to ensure privacy, the challenges encountered, and the resulting outcomes. The comparison among these cases highlights both common patterns and distinct strategies that reflect the institutions' priorities—be it performance, privacy, or regulatory adherence.

Table 3: Case Studies on Privacy and Compliance in FL for Banking

Case Study	Description	Privacy Approach	Outcome	Reference
SWIFT-Google Cloud FL	FL across banks to detect real-time fraud in cross-border transactions	Encrypted model updates, no raw data transfer	Improved fraud detection and GDPR compliance	[63]
Google GBoard FL	Keyboard suggestion model trained using FL with DP	Local differential privacy and decentralized training	Balanced privacy with performance, but slower convergence	[64]
Baidu Mobile FL	Fraud detection in mobile banking using FL with homomorphic encryptio	Preserved mobile data integrity and user privacy	High detection accuracy, but high computation cost	[65]

	n			
European Bank FL Prototype	FL prototype tested on European banking data	Privacy-preserving model governance, GDPR audit trails	Positive evaluation by data regulators	[66]

These cases underscore that FL can be effectively aligned with privacy and regulatory needs, but success depends on implementation specifics. While Google's GBoard case showed usability in large-scale applications, banking institutions face higher stakes and must adopt more stringent encryption and audit measures. The trade-off between privacy strength and model performance remains a key challenge. However, innovations in federated optimization, differential privacy tuning, and federated explainability suggest a promising future trajectory.

V. ALGORITHMIC PERFORMANCE IN FEDERATED LEARNING

Federated Learning (FL) has emerged as a transformative approach in digital banking, enabling collaborative model training across decentralized data sources while preserving data privacy. However, achieving optimal algorithmic performance in FL systems, particularly for fraud detection, presents several challenges. This section delves into these challenges, examining factors such as data heterogeneity, communication costs, data imbalance, and optimization techniques, supported by recent empirical studies and literature.

A. Challenges in Achieving High Algorithmic Performance

FL systems often grapple with data heterogeneity, where participating clients possess data that is not independently and identically distributed (non-IID). This heterogeneity can lead to biased model updates and hinder convergence, as the aggregated global model may not generalize well across diverse client data distributions. Additionally, the decentralized nature of FL introduces complexities in coordinating model updates, managing asynchronous training, and ensuring consistency across clients.

B. Factors Affecting Model Accuracy

Model accuracy in FL is influenced by several factors, including the quality and quantity of local data, the frequency of model updates, and the aggregation strategy employed. Variations in local data distributions can cause discrepancies in model performance across clients. Moreover, infrequent communication between clients and the central server can slow down convergence and affect the accuracy of the global model.

C. Data Heterogeneity and Its Impact on Performance

Data heterogeneity poses a significant challenge in FL, as clients may have vastly different data distributions. This can lead to the phenomenon of "client drift," where local models diverge from the global model, resulting in degraded overall performance. Addressing this issue requires advanced aggregation methods and personalized FL approaches that account for client-specific data characteristics.

D. Communication Costs in Federated Learning Systems

Communication overhead is a critical concern in FL, as frequent transmission of model updates between clients and the central server can strain network resources. Techniques such as model compression, quantization, and sparsification have been proposed to reduce communication costs. For instance, knowledge distillation methods can significantly decrease the amount of data exchanged while maintaining model performance.

E. Data Imbalance and Its Effect on Fraud Detection Models

In fraud detection, datasets are often highly imbalanced, with fraudulent transactions constituting a small fraction of the total data. This imbalance can cause models to be biased towards the majority class, leading to poor detection of fraudulent activities. In FL, this issue is exacerbated by the decentralized nature of data, necessitating strategies such as data resampling, cost-sensitive learning, and anomaly detection techniques to improve model sensitivity to minority classes.

F. Techniques for Improving Algorithmic Performance

To enhance the performance of FL systems, various optimization techniques have been explored. Adaptive learning rate strategies, such as cyclical learning rates, have shown promise in accelerating convergence and improving model accuracy under non-IID conditions. Additionally, hyperparameter optimization methods tailored for FL settings can further refine model performance.

G. Evaluating the Performance of Federated Learning in Real-World Banking Contexts

Assessing FL performance in practical banking scenarios involves evaluating metrics such as accuracy, precision, recall, and convergence speed. Studies have proposed holistic evaluation frameworks that consider computational efficiency, fairness, and personalization to provide a comprehensive assessment of FL systems in financial applications.

H. Balancing Privacy, Performance, and Compliance in Federated Models

Achieving an optimal balance between privacy, performance, and regulatory compliance is crucial in FL deployments. Techniques like differential privacy can enhance data protection but may impact model accuracy. Adaptive mechanisms that dynamically adjust privacy budgets and performance objectives have been proposed to navigate this trade-off, ensuring compliance with data protection regulations while maintaining effective model performance. These case studies illustrate the practical applications and benefits of addressing algorithmic performance challenges in FL systems within the banking sector. The BalancerGNN framework demonstrates the effectiveness of leveraging advanced neural network architectures to handle data imbalance, which happens to be a common issue in fraud detection. The FedISM approach highlights the potential of incorporating shared models to mitigate the effects of non-IID data distributions, leading to significant improvements in model accuracy. The credit risk forecasting study underscores the advantages of FL in enhancing model performance for clients with limited data, emphasizing the importance of collaborative learning in financial applications.

In summary, while FL offers significant advantages in preserving data privacy and enabling collaborative model training in digital banking, addressing the associated algorithmic challenges is essential. Ongoing research and the development of advanced optimization and evaluation techniques are critical to realizing the full potential of FL in fraud detection and other financial applications.

Table 4: Case Studies on Algorithmic Performance in Federated Learning for Banking

Case Study	Description	Challenges Addressed	Outcomes	Reference
BalancerGNN Framework	Utilized Graph Neural Networks for fraud detection on imbalanced datasets	Data imbalance, feature redundancy	Achieved sensitivity rates between 72.87% to 81.23% across datasets	[67]
FedISM Approach	Enhanced data imbalance handling via shared models in FL	Data imbalance, non-IID data	Improved accuracy by up to 25% with minimal shared data	[68]
Credit Risk Forecasting Study	Applied FL for credit risk assessment across multiple datasets	Data imbalance, client data diversity	Noted a 17.92% average improvement in model performance on non-dominant clients	[69]

CONCLUSION AND FUTURE DIRECTIONS

This study has provided a comprehensive examination of Federated Learning (FL) as a transformative paradigm for privacy-preserving fraud detection in the digital banking sector. Drawing on empirical analyses, comparative case studies, and literature-supported evaluations, the work highlights the dual potential of FL to enhance security and maintain regulatory compliance without compromising model performance. Several key findings and implications emerge from the results.

A. Summary of Key Findings

The research confirms that FL enables collaborative model training without necessitating the centralization of sensitive customer data, thus significantly reducing privacy risks. Case studies such as BalancerGNN and FedISM reveal that, despite challenges like data heterogeneity, communication overhead, and class imbalance, tailored FL algorithms can outperform traditional centralized models in fraud detection sensitivity and adaptability. Moreover, compliance with data protection regulations such as the GDPR and CCPA is inherently more feasible in FL due to minimal data transmission, which aligns well with modern regulatory expectations.

B. Contributions to the Field of Privacy-Preserving Fraud Detection

This work contributes to the growing body of knowledge in federated machine learning by:

- Providing an empirical and theoretical assessment of FL's applicability in detecting financial fraud across heterogeneous banking datasets.
- Introducing detailed discussions on how privacy, performance, and compliance can be balanced through optimization strategies and privacy-preserving mechanisms like differential privacy and secure aggregation.
- Highlighting critical trade-offs that digital banks must consider, such as the tension between model accuracy and regulatory compliance, especially under real-world data imbalance conditions.

C. Recommendations for Digital Banks and Regulatory Bodies

For digital banks, the adoption of FL should be viewed as a strategic investment in both cybersecurity and regulatory alignment. Institutions are encouraged to:

- Implement pilot FL systems in high-risk, data-sensitive applications such as fraud detection and credit scoring.
- Incorporate privacy-preserving mechanisms (e.g., differential privacy, secure multiparty computation) early in system design.
- Foster interbank collaborations for shared model training under legal and technological safeguards.

D. For regulatory bodies, it is essential to:

- Develop clear guidelines for privacy-preserving AI systems, with provisions specific to decentralized learning frameworks.
- Encourage sandbox environments for testing FL technologies in real financial environments.
- Create incentives for banks to collaborate on fraud detection via federated platforms while ensuring transparency and auditability.

E. Future Research Directions in Federated Learning and Fraud Detection

There are several promising avenues for future research. These include:

- Designing adaptive federated algorithms that can dynamically balance local accuracy with global convergence, especially in highly non-IID settings.
- Developing unified evaluation frameworks that holistically consider privacy loss, performance metrics, and regulatory risk.
- Exploring federated reinforcement learning and its application to real-time fraud detection systems.
- Investigating FL deployment under resource-constrained environments, particularly in developing regions or fintech startups with limited infrastructure.

F. Long-Term Vision for Federated Learning in the Financial Industry

Looking ahead, Federated Learning has the potential to redefine the financial industry's approach to data collaboration, security, and innovation. As financial crimes become increasingly sophisticated, the capacity to leverage distributed intelligence without compromising client data confidentiality will be critical. In the long term, FL could support cross-border fraud detection networks, intelligent anti-money laundering systems, and decentralized credit scoring platforms—all while remaining compliant with global privacy standards. Ultimately, this work envisions a future where federated systems not only safeguard individual privacy but also enhance the collective intelligence of the financial ecosystem, ensuring resilient, ethical, and intelligent banking for the digital age.

REFERENCES

[1] Munira, M., & Khatun, M. S. (2025). Digital transformation in banking: A systematic review of trends, technologies, and challenges. Available from SSRN: https://ssrn.com/abstract=5161354 or http://dx.doi.org/10.2139/ssrn.5161354

[2] Windasari, N. A., Kusumawati, N., Larasati, N., & Amelia, R. P. (2022). Digital-only banking experience: Insights from gen Y and gen Z. Journal of Innovation & Knowledge, 7(2), 100170. 1 https://doi.org/10.1016/j.jik.2022.100170

[3] Ionașcu, A. E., Gheorghiu, G., Spătariu, E. C., Munteanu, I., Grigorescu, A., & Dănilă, A. (2023). Unraveling Digital Transformation in Banking: Evidence from Romania. Systems, 11(11), 534. https://doi.org/10.3390/systems11110534.

[4] R, Y., Nithin, C., G, T., & Safi, Dr. K. (2025). The rise of digital banking and its effect on traditional banks. EPRA International Journal of Environmental Economics, Commerce and Educational Management, 12(4), 30. https://doi.org/10.36713/epra0414

[5] Wang, S., Asif, M., Shahzad, M. F., & Ashfaq, M. (2024). Data Privacy and Cybersecurity Challenges in the Digital Transformation of the Banking Sector. Computers & Security, 143, 104051. https://doi.org/10.1016/j.cose.2024.104051

[6] Ferrão, S. É. R., Silva, G. R. S., Canedo, E. D., & Mendes, F. F. (2024, April). Towards a taxonomy of privacy requirements based on the LGPD and ISO/IEC 29100. Information and Software Technology, 168, 107396. https://doi.org/10.1016/j.infsof.2024.107396

[7] Wong, R., Chong, A., & Aspegren, C. (2023, April). Privacy Legislation as Business Risks: How GDPR and CCPA are Represented in Technology Companies' Investment Risk Disclosures. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), 1–26. https://doi.org/10.1145/3579515

[8] Rafi, T. H., Noor, F. A., Hussain, T., & Chae, D.-K. (2024, May). Fairness and privacy preserving in federated learning: A survey. Information Fusion, 105, 102198.

[9] Orabi, M. M., Emam, O., & Fahmy, H. (2025). Adapting security and decentralized knowledge enhancement in federated learning using blockchain technology: literature review. Journal of Big Data, 12, 55. https://doi.org/10.1186/s40537-025-01099-5

[10] Aljunaid, S. K., Almheiri, S. J., Dawood, H., & Khan, M. A. (2025). Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection. Journal of Risk and Financial Management, 18(4), 179. https://doi.org/10.3390/jrfm18040179.

[11] Susmitha, N., & Kalpana, Smt. A. (2025, January). Fraud Detection in Banking Data Using Machine Learning Techniques. Journal of Engineering Sciences, 16(04), 120-125.

[12] Sunhare, M., & Mandloi, R. S. (2025, February). Developing Machine Learning Models for Real-Time Fraud Detection in Online Transactions. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE), 14(2), 465. https://doi.org/10.15662/IJAREEIE.2025.1402019.

[13] Hilal, W., Gadsden, S. A., & Yawney, J. (2022, May 1). Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Systems with Applications, 193, 116429. 1 https://doi.org/10.1016/j.eswa.2021.116429

[14] Gresoi, S., Stamatescu, G., & Făgărășan, I. (2025). Advanced Methodology for Fraud Detection in Energy Using Machine Learning Algorithms. Applied Sciences, 15(6), 3361. https://doi.org/10.3390/app15063361

[15] Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & Egwuonwu, M. N. (2023). Analysing the Impact of Advanced Analytics on Fraud Detection: A Machine Learning Perspective. European Journal of Computer Science and Information Technology, 11(6), 103-126. 1 https://doi.org/10.37745/ejcsit.2013/vol11n6103126

[16] Kuyoro, S. O., Ogunyolu, O., Ayanwola, T., & Ayankoya, F. Y. (2022, October). Dynamic Effectiveness of Random Forest Algorithm in Financial Credit Risk Management for Improving Output Accuracy and Loan Classification Prediction. Ingénierie des systèmes d information, 27(5), 815-821. https://doi.org/10.18280/isi.270515

[17] Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., ... & Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6, 100163. https://doi.org/10.1016/j.dajour.2023.100163.

[18] Preciado Martínez, P. M., Reier Forradellas, R. F., Garay Gallastegui, L. M., & Náñez Alonso, S. L. (2025). Comparative analysis of machine learning models for the detection of fraudulent banking transactions. Cogent Business & Management, 12(1). https://doi.org/10.1080/23311975.2025.2474209

[19] Barry, M., Bifet, A., Chiky, R., Montiel, J., & Tran, V.-T. (2021). Challenges of machine learning for data streams in the banking industry. In Big Data Analytics (pp. xx-xx). Springer, Cham. https://doi.org/10.1007/978-3-030-93620-4_9

[20] Agripina, N.E.M.R., Shen, H. and Mafukidze, B.S. (2024) Advances, Challenges & Recent Developments in Federated Learning. Open Access Library Journal, 11, 1-1. doi: 10.4236/oalib.1112239.

[21] Edegbe, G. N., & Acheme, S. (2024). A systematic review of centralized and decentralized machine learning models: Security concerns, defenses and future directions. NIPES Journal of Science and Technology Research, 6(4), 161-175. https://doi.org/10.5281/zenodo.14681449

[22] Drainakis, G., Katsaros, K. V., Pantazopoulos, P., Sourlas, V., & Amditis, A. (2020). Federated vs. centralized machine learning under privacy-elastic users: A comparative analysis. In Proceedings of the 19th IEEE International Symposium on Network Computing and Applications (NCA 2020). https://doi.org/10.1109/NCA51143.2020.9306745

[23] Mohammadi, S., Balador, A., Sinaei, S., & Flammini, F. (2024). Balancing privacy and performance in federated learning: A systematic literature review on methods and metrics. Journal of Parallel and Distributed Computing, 192, 104918. https://doi.org/10.1016/j.jpdc.2024.104918.

[24] Yurdem, B., Kuzlu, M., Gullu, M. K., Catak, F. O., & Tabassum, M. (2024). Federated learning: Overview, strategies, applications, tools and future directions. Heliyon, 10(19), e38137. https://doi.org/10.1016/j.heliyon.2024.e38137.

[25] Myakala, P. K., Bura, C., & Jonnalagadda, A. K. (2024). Federated learning and data privacy: A review of challenges and opportunities. International Journal of Research Publication and Reviews, 5(12), 1867-1879. https://doi.org/10.55248/gengpi.5.1224.3512

[26] Agripina, N.E.M.R., Shen, H. and Mafukidze, B.S. (2024) Advances, Challenges & Recent Developments in Federated Learning. Open Access Library Journal, 11, 1-1. doi: 10.4236/oalib.1112239.

[27] Barona López, L. I., & Borja Saltos, T. (2025). Heterogeneity Challenges of Federated Learning for Future Wireless Communication

Networks. Journal of Sensor and Actuator Networks, 14(2), 37. https://doi.org/10.3390/jsan14020037.

[28] Bhanbhro, J., Nisticò, S. & Palopoli, L. Issues in federated learning: some experiments and preliminary results. Sci Rep 14, 29881 (2024). https://doi.org/10.1038/s41598-024-81732-0

[29] Xu, C., Qu, Y., Xiang, Y., & Gao, L. (2023). Asynchronous federated learning on heterogeneous devices: A survey. Computer Science Review, 50, 100595. https://doi.org/10.1016/j.cosrev.2023.100595

[30] Zhu, B., & Niu, L. (2025). A privacy-preserving federated learning scheme with homomorphic encryption and edge computing. Alexandria Engineering Journal, 118, 11-20. https://doi.org/10.1016/j.aej.2024.12.070.

[31] Wulan, R. B. (2023). Legal protection of customer personal data in the banking sector. ARRUS Journal of Social Sciences and Humanities, 3(5), 710–717. https://doi.org/10.35877/soshum2169

[32] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2020). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology, 10(2), 1–19.

[33] N. F. Aurna, M. D. Hossain, Y. Taenaka and Y. Kadobayashi, "Federated Learning-Based Credit Card Fraud Detection: Performance Analysis with Sampling Methods and Deep Learning Algorithms," 2023 IEEE International Conference on Cyber Security and Resilience (CSR), Venice, Italy, 2023, pp. 180-186, doi: 10.1109/CSR57506.2023.10224978.

[34] Salam, M. A., Fouad, K. M., Elbably, D. L., & Elsayed, S. M. (2024). Federated learning model for credit card fraud detection with data balancing techniques. Neural Computing and Applications, 36, 6231–6256. https://doi.org/10.1007/s00521-023-09410-2.

[35] Sun, R. (2025). A comprehensive investigation of fraud detection behavior in federated learning. ITM Web of Conferences, 70, 03030. https://doi.org/10.1051/itmconf/20257003030.

[36] Aljunaid, S. K., Almheiri, S. J., Dawood, H., & Khan, M. A. (2025). Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection. Journal of Risk and Financial Management, 18(4), 179. https://doi.org/10.3390/jrfm18040179.

[37] Dritsas, E., & Trigka, M. (2025). Federated Learning for IoT: A Survey of Techniques, Challenges, and Applications. Journal of Sensor and Actuator Networks, 14(1), 9. https://doi.org/10.3390/jsan14010009.

[38] Zheng, H. (2025). Federated Learning-Based Credit Card Fraud Detection: A Comparative Analysis of Advanced Machine Learning Models. ITM Web of Conferences, 70, 01022. https://doi.org/10.1051/itmconf/20257001022.

[39] Tang, Y., & Liang, Y. (2024). Credit card fraud detection based on federated graph learning. Expert Systems with Applications, 256, 124979. https://doi.org/10.1016/j.eswa.2024.124979.

[40] Google Cloud. (2025). To help combat fraud, Google Cloud and Swift pioneer advanced AI and federated learning tech. Retrieved from https://cloud.google.com/blog/products/identity-security/google-cloud-and-swift-pioneer-advanced-ai-and-federated-learning-tech.

[41] Guo, W., Zhuang, F., Zhang, X., Tong, Y., & Dong, J. (2024). A comprehensive survey of federated transfer learning: challenges, methods and applications. Frontiers of Computer Science, 18, 186356. https://doi.org/10.1007/s11704-024-40065-x.

[42] Barreto, B., Senna, C., Rito, P., & Sargento, S. (2025). MobFedLS: A framework to provide federated learning for mobile nodes in V2X environments. Future Generation Computer Systems, 163, 107514. https://doi.org/10.1016/j.future.2024.107514.

[43] Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804. https://doi.org/10.1016/j.inffus.2023.101804.

[44] Zarsky, T. (2017). Incompatible: The GDPR in the Age of Big Data. Seton Hall Law Review, 47(4), 1041-1072. https://ssrn.com/abstract=3022646.

[45] Khan, S., Kabanov, I., Madnick, S., & Hua, Y. (2023). A Systematic Analysis of the Capital

One Data Breach: Critical Lessons Learned. ACM Transactions on Privacy and Security. http://dx.doi.org/10.1145/3546068.

[46] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (pp. 308-318). https://doi.org/10.1145/2976749.2978318.

[47] Rahaman, M., Arya, V., Orozco, S. M., & Pappachan, P. (2024). Secure Multi-Party Computation (SMPC) Protocols and Privacy. In Innovations in Modern Cryptography, IGI Global. https://doi.org/10.4018/979-8-3693-5330-1.ch008.

[48] Kanamori, S., Abe, T., Ito, T., Emura, K., Wang, L., Yamamoto, S., ... & Nojima, R. (2022). Privacy-Preserving Federated Learning for Detecting Fraudulent Financial Transactions in Japanese Banks. Journal of Information Processing, 30, 789-795. https://doi.org/10.2197/ipsjjip.30.789

[49] Truong, N., Sun, K., Wang, S., Guitton, F., & Guo, Y. (2021). Privacy preservation in federated learning: An insightful survey from the GDPR perspective. Computers & Security, 110, 102402. https://doi.org/10.1016/j.cose.2021.102402.

[50] El Mestari, S. Z., Lenzini, G., & Demirci, H. (2024). Preserving data privacy in machine learning systems. Computers & Security, 137, 103605. https://doi.org/10.1016/j.cose.2023.103605

[51] Feretzakis, G., Vagena, E., Kalodanis, K., Peristera, P., Kalles, D., & Anastasiou, A. (2025). GDPR and Large Language Models: Technical and Legal Obstacles. Future Internet, 17(4), 151. https://doi.org/10.3390/fi17040151.

[52] Papadopoulos, C., Kollias, K.-F., & Fragulis, G. F. (2024). Recent Advancements in Federated Learning: State of the Art, Fundamentals, Principles, IoT Applications and Future Trends. Future Internet, 16(11), 415. https://doi.org/10.3390/fi16110415.

[53] Shukla, S., Rajkumar, S., Sinha, A., Esha, M., Elango, K., & Sampath, V. (2025). Federated learning with differential privacy for breast cancer diagnosis enabling secure data sharing and model integrity. Scientific Reports, 15, 13061. https://doi.org/10.1038/s41598-025-95858-2.

[54] Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2018). Federated learning for mobile keyboard prediction. arXiv:1811.03604 [cs.CL]. https://doi.org/10.48550/arXiv.1811.03604.

[55] Zhu, H., Xu, J., Liu, S., & Jin, Y. (2021). Federated Learning on Non-IID Data: A Survey. arXiv:2106.06843 [cs.LG]. https://doi.org/10.48550/arXiv.2106.06843

[56] Z. Lu, H. Pan, Y. Dai, X. Si and Y. Zhang, "Federated Learning With Non-IID Data: A Survey," in IEEE Internet of Things Journal, vol. 11, no. 11, pp. 19188-19209, 1 June1, 2024, doi: 10.1109/JIOT.2024.3376548.

[57] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Ramage, D. (n.d.). Advances and Open Problems in Federated Learning. Now Foundations and Trends. https://ieeexplore.ieee.org/Xplore/home.jsp

[58] Google Cloud. (2025). To help combat fraud, Google Cloud and Swift pioneer advanced AI and federated learning tech. Retrieved from https://cloud.google.com/blog/products/identity-security/google-cloud-and-swift-pioneer-advanced-ai-and-federated-learning-tech

[59] Xu, Z., Zhang, Y., Andrew, G., Choquette, C., Kairouz, P., McMahan, B., Rosenstock, J., & Zhang, Y. (2023). Federated Learning of Gboard Language Models with Differential Privacy. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 5: Industry Track), pages 629–639, Toronto, Canada. Association for Computational Linguistics. DOI: 10.18653/v1/2023.acl-industry.60.

[60] Khan, A., ten Thij, M. & Wilbik, A. (2025). Vertical federated learning: a structured literature review. Knowl Inf Syst 67, 3205–3243 (2025). https://doi.org/10.1007/s10115-025-02356-y

[61] Lee, C. M., Fernandez, J. D., Menci, S. P., ... & Fridgen, G. (2023). Federated Learning for Credit Risk Assessment. In Proceedings of the 56th Hawaii International Conference on System Sciences, Maui, HI. https://doi.org/10.24251/HICSS.2023.048.

[62] Boyapati, M., & Aygun, R. (2024). BalancerGNN: Balancer Graph Neural Networks for imbalanced datasets: A case study on fraud detection. Neural Networks, 182(1), 106926. https://doi.org/10.1016/j.neunet.2024.106926.

[63] Chung, W.-C., Lin, Y.-H., & Fang, S.-H. (2023). FedISM: Enhancing Data Imbalance via Shared Model in Federated Learning. Mathematics, 11(10), 2385. https://doi.org/10.3390/math11102385.

[64] Hua, S., Zhang, C., Yang, G., Fu, J., Yang, Z., Wang, L., & Ren, J. (2024). An FTwNB Shield: A Credit Risk Assessment Model for Data Uncertainty and Privacy Protection. Mathematics, 12(11), 1695. https://doi.org/10.3390/math12111695

[65] Olowu, O., Adeleye, A. O., Omokanye, A. O., Ajayi, A. M., Adepoju, A. O., Omole, O. M., &Chianumba, E. C. (2024). AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity.

[66] Ariyibi, K. O., Bello, O. F., Ekundayo, T. F., &Ishola, O. (2024).Leveraging Artificial Intelligence for enhanced tax fraud detection in modern fiscal systems.

[67] Adesola O., Taiwo I., David, D. A., Ezenwa, H. N., & Quddus, A. A. (2025). Utilizing AI and machine learning algorithms to optimize supplier relationship management and risk mitigation in global supply chains.International Journal of Science and Research Archive, 2025, 14(02), 219-228

[68] David, A. A., & Edoise, A. (2025). Cloud computing and Machine Learning for Scalable Predictive Analytics and Automation: A Framework for Solving Real-world Problem.

[69] Olawale, A., Ajoke, O., &Adeusi, C. (2020).Quality assessment and monitoring of networks using passive