FEDERATED LEARNING FOR PRIVACY-PRESERVING: A REVIEW OF PII DATA ANALYSIS IN FINTECH

Bibhu Dash, Pawankumar Sharma and Azad Ali

Dept. of Computer and Information Systems, University of the Cumberlands, KY, USA

ABSTRACT

There has been tremendous growth in the field of AI and machine learning. The developments across these fields have resulted in a considerable increase in other FinTech fields. Cyber security has been described as an essential part of the developments associated with technology. Increased cyber security ensures that people remain protected, and that data remains safe. New methods have been integrated into developing AI that achieves cyber security. The data analysis capabilities of AI and its cyber security functions have ensured that privacy has increased significantly. The ethical concept associated with data privacy has also been advocated across most FinTech regulations. These concepts and considerations have all been engaged with the need to achieve the required ethical requirements. The concept of federated learning is a recently developed measure that achieves the abovementioned concept. It ensured the development of AI and machine learning while keeping privacy in data analysis. The research paper effectively describes the issue of federated learning for confidentiality. It describes the overall process associated with its development and some of the contributions it has achieved. The widespread application of federated learning in FinTech is showcased, and why federated learning is essential for overall growth in FinTech.

KEYWORDS

FinTech, AI, federated learning, machine learning, cyber security, data privacy, PII data, differential privacy.

1. INTRODUCTION

The fields of machine learning and data science have been described to have an essential role in science. The development of science through data analysis is associated with the large amounts of data included across research studies. Therefore, the need for privacy is an ongoing principle for most individuals. Data collection has been a growing concept across the globe. This can be seen from the increase in software and applications focused on collecting data. Social media platforms have also been indicated to collect data from their users, providing better services [1]. Data usage has increased significantly, resulting in more data that requires privacy. Consumers and policymakers have therefore focused on privacy-related concepts. The General Data Protection Regulation (GDPR) is among the actions taken due to data protection across the globe. The move has been associated with the need for promoting the required types of developments and ensuring that there are significant achievements related to protected data. Google introduced the federated learning idea in 2017 [2]. The concept enabled data scientists to share their statistical models in analyzing data. Security was, however, a key aspect required in data sharing. Federated learning was, therefore, an effective model for enabling privacy for data analysts across the globe.

DOI: 10.5121/ijsea.2022.13401

2. Federated Learning - An Overview

Federated learning was a plan introduced by Google back in 2017, and the idea was associated with assisting data scientists with what they do [3]. Federated Learning is a machine learning method for training models on a large corpus of decentralized data. Federated Learning works on a scalable production framework for independent analysis and machine learning across many devices and domains. The primary function of federated learning was to ensure privacy and promote effectiveness in the work of these data scientists. Therefore, scientists could share statistical data analysis models on decentralized devices and servers with local data sets. This concept allowed the scientists access to better models that assist their work.

Since the developed systems are decentralized, there was no official control over the statistical models that the scientist shared. The idea allowed most scientists to be amazed by the capabilities of the entire federated learning system [3]. The main advantages also included that the data scientists were not required to send data online or to the cloud. Their only requirement was to obtain the statistical models from the federated learning and use them to analyze their data. Comparing the new decentralized system with federated knowledge and the traditional centralized machine learning techniques showcases a significant advancement in maintaining data security for different users. Research teams and data scientists can now ensure the privacy of their data through these systems. Therefore, the number of technology systems has promoted the need to maintain these types of systems. The overall promotion associated with these data types is that they ensure no concerns related to data privacy. The confidence induced towards researchers influences their general research in completing with complete confidence.

Federated learning has attained an increasing awareness of its functionality and capabilities. The technology tackles a leading problem for most data scientists across the globe. Research teams focused on sensitive information are also significantly assisted through the new technology, which promotes privacy for their data [4]. The removal of the centralized server technologies was also a leading aspect that has enabled this type of technology. More people have appreciated the new systems, ensuring better and more effective privacy for any sensitive research study or data analysis [5]. Federated learning has also been described as having minimal updates, providing the machine learning model is more effective. These unique characteristics have promoted the use of the federated learning tool and have attracted an increasing number of users from across the different parts of the globe.

Different companies and data analysis foundations have engaged in using the new system. A widely known example is Nvidia, which recently used Federated Learning (FL) on its autonomous driving platform [4]. This model was associated with promoting the required development of autonomous driving from the different parties available [6]. This model was also associated with having diverse geographical landscapes and driving situations that all need to be described. There are also increased data sets that people cannot define [6] (see Fig 1). Therefore, the data analysis is associated with OEMs who train their driving models individuals and send them to federated learning (FL), where the shared model becomes more effective than the system already available [7]. The background of federated learning showcases that it may be adopted in personally identifiable information (PII) data analysis [8]. The model may be adopted by organizations such as FinTech, which will attain more effective processes.

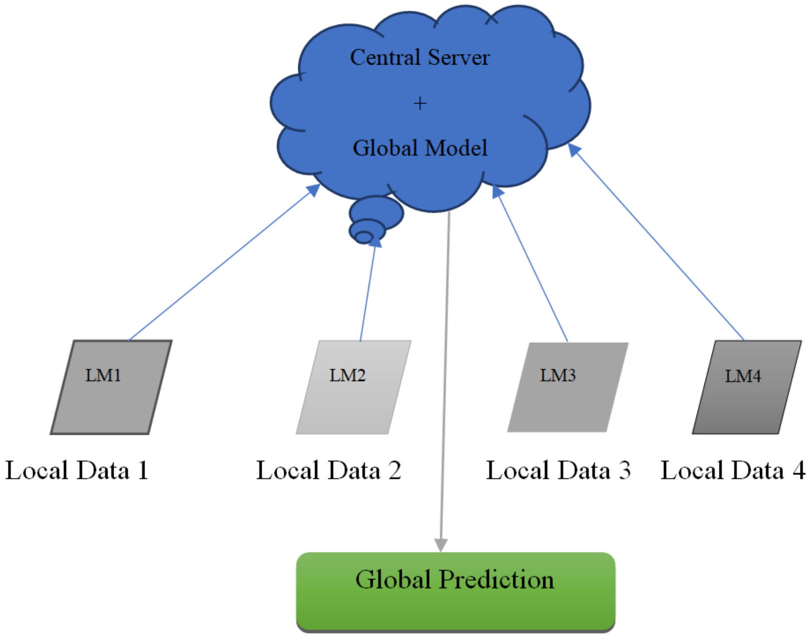

Fig 1. Centralized topology of Federated Learning (FL) on different local datasets
*LM – Local Model

3. Federated Learning and Privacy

By definition, "Federated learning is a machine learning setting where multiple entities (clients) collaborate in solving a machine learning problem, under the coordination of a central server or service provider. Each client's raw data is stored locally and not exchanged or transferred; instead, focused updates for immediate aggregation are used to achieve the learning objective" [1]. As described above, the primary use of federated learning was to promote privacy for data scientists and researchers. Centralized learning and data collection have been described as a potential risk for users' data [9]. Organizations significantly benefit from increased privacy which ensures they have avoided legal issues. The risks associated with organizations being sued have increased the need for using the decentralized model. The expanding community associated with federated learning indicates that it achieves the required level of privacy that organizations and scientists require.

3.1. Privacy Principles for Learning and Analytics

The principles of learning and analytics are essential concepts to describe when understanding whether the decentralized, federated learning system achieves different privacy principles and explains its popularity and growing community of users [10]. Bonawitz et al. (2021) indicate that privacy is a multifaceted concept [1]. Privacy could be described from an increased number of considerations. Bonawitz et al. (2021) consider three main ideas regarding achieving privacy [1]. The article highlights transparency, consent, data minimization, and anonymization are the focus areas for achieving privacy. The foundation of privacy is showcased to require performing these three aspects.

Consent and transparency are vital aspects associated with privacy. This element focuses on describing how both parties agree on using their data. It includes a description of how the first party approves their data use [11]. Federated learning achieves the concept of consent and transparency since parties from both sides understand what is required [12]. The available parties realize that technology is involved in sharing the resources necessary and avoiding a centralized model that will describe the data. The main focus is on the exact approach it uses to achieve this form of privacy.

Data anonymization is another principle that engages in output analysis. Since the technology involved is associated with attaining the required system. The final output is also significant to the first party while it is not considered by the second party [13]. The overall focus will attain the required level of privacy. The principle of data minimization is also achieved through the federated learning principle [12]. Data minimization focuses on ensuring access has been minimized to all individuals. Since federated learning limits access to the user alone, it indicates that federated learning achieves the required level of privacy.

The growing aspect of federated learning focuses on the concept of being combined with other techniques, promoting its push for privacy. The overall model ensures that the above principles have been achieved, enabling the essential decentralization characteristics [14]. The approach ensures that analysts and researchers have benefited from the list of its applications and settings available. The focus included also provides that direct access is available [15]. Direct access means that users will not lose any essential data or time when using the system model [12]. Therefore, including AI has promoted the required development and growth associated with these models. It ensures that users benefit from privacy from the applications included.

3.2. Federated Learning Settings and Applications

Understanding the effectiveness of federated learning requires an overall analysis of its settings and applications. The main characteristic of federated learning is that it keeps the available data in a decentralized environment. The system model is effective because it ensures learning via aggregation has also been achieved. The entire approach remains focused on whether the information is entirely private and safe. Each of these concepts requires that engaged research has been provided on its effectiveness in promoting the needed privacy requirements [11].

Therefore, the federated learning environment acts like a data center that ensures data is secure through different approaches focused on privacy [2]. The learning settings ensure that arbitrary data can be distributed and shuffled towards adequate data from the selected system. This engages confidence from the different users who appreciate using this model. These individuals ensure an overall focus associated with every individual having access to the data from any location and time required.

There are mainly two federated learning settings that are widely known. They include Cross-device FL and Cross-silo FL. The cross-device FL refers to a setting in which the clients have many devices which they can shift [16]. It also applies to IoT devices, effectively accessing the required services from different regions and devices. On the other hand, the cross-silo FL refers to where clients are in small organizations, meaning that the institutions do not have many employees or individuals requiring access to the information available [17]. The data set is also different for each decentralized data center, making a choice relevant to the company or organization [18].

Federated learning has been classified into two main concepts. The parts and features across the two types of systems are different. The categorization focuses on how data is distributed across the participating parties [19]. It refers to the features in which sample spaces are included and how they are accessed. Categorization is an essential aspect of the federated learning models since they refer to the available concept and how it is accepted and referred to by different parties. It describes the functional approach associated with reference and storing the data and how analysis is achieved. The organizations using the required processes may require different strategies, resulting in the need for different categorizations.

3.3. Federated Learning Security and Privacy Challenges

FL has been described as having several challenges that impact its users. Several privacy and security challenges are associated with the continuous use of federated learning, FL. The users are affected by these challenges, which hinder the individuals from carefully executing the required applications and services of the federated learning tools [20]. The overall approach focuses on the privacy issues that may impact the users [21]. The main privacy issue is having raw data in a distributed machine learning setting.

Other concerns have been associated with sharing model information as the training procedure. The approach has trained the system and machine to become more valuable and approachable [22]. The overall focus remains on the potential leak, which may be included by the machine used. The leak could affect the individuals available and have a much more significant impact on all the individuals available. The issues associated with these practices include the continued risk of data loss and sensitive text information getting lost [23]. These challenges indicate the required focus on improving the federated learning system and ensuring people remain safe. Several approaches have already been developed to ensure privacy has been guaranteed, and the risks and challenges showcased are resolved.

The concept of differential privacy was developed as a randomized mechanism to ensure output distribution. The approach focuses on promoting the elements included and removing access to samples engaged in the systems [24]. It enables some focus to prevent access to the type of sample used in the machine learning process. Therefore, the learning methods are protected, which ensures differential privacy has been achieved. The approach can focus on privacy principles, indicating some effectiveness of the adopted system [25].

Another approach engaged has been the secure multiparty computation SMG. This approach focuses on a collaborative computational idea that ensures the functions included do not revolve around the required type of development [26]. It promotes a specific collaborative agreed function, ensuring leaking is avoided. The approach, therefore, achieves the required learning and analysis while maintaining security and privacy by preventing leaks. The process has since been described to work very effectively.

Both approaches described have the limitation of not being able to work more effectively in a scenario where the organization is larger. It limits the capabilities that could be achieved in the method which best applies [5]. Machine learning deployments may require an even more significant focus that the adopted approaches could limit [27]. Therefore, the overall limitation can be described as limiting to the main focus areas and how they apply to different individuals. However, Federated learning has been focused on reducing all the challenges it engages across by focusing on other updates that focus on removing these challenges [28]. The issues of privacy are therefore disregarded every time there is an update. Thus, Federated learning continues to be inexpensive through increased users, communication efficiency, and tolerance.

4. FEDERATED LEARNING FOR PRIVACY-PRESERVING

Machine learning has been described to have tremendous success in promoting AI applications. Cheng et al. (2020) explain that the practicality of AI has been pushed mainly by the success of machine learning [2]. The study indicated that speech recognition, self-autonomy, and computer vision applications had been promoted primarily through machine learning. Two main challenges impact the success and growth of AI [29]. They include data on isolated islands and the increase in PPA demand [2]. However, the conventional approaches focus on the older methods of using centralized data collection approaches, which then contribute toward crucial engagement in the field of AI. The main problems with AI approaches can therefore be described as isolation and data fragmentation which is also faced by the issue of privacy protection. The government has also focused on privacy protection, as showcased by the different laws [30]. This creates a challenge that must be resolved in the growing field of AI and machine learning researchers and data scientists. Resolving these challenges was introduced through the development of the FL.

The main advantage associated with promoting the use of FL has been the continued focus on promoting FL applications [31]. The overall emphasis on methods required to engage in data privacy contributes to the increased use of FL. The FL capabilities have been employed towards achieving the government requirements and promoting a specific type of development for parties available [2]. The entire approach remains focused on achieving privacy and the required service implementation and how well it promotes the needed type of development and applications [32]. Fundamentally fixing the issues associated with previous models has been the main focus of this approach and its uses in the larger community.

Federated learning, as described above, achieves several required principles of PPI. The main aim of promoting this type of development requires more focus on privacy [33]. Privacy contributes toward a large portion of the PPI required. Organizations mainly contribute to their approaches by focusing on some of these approaches and how well they can be engaged [2]. The homomorphic encryptions especially engage towards the required development, how well it promotes significant growth towards the rising community, and how they achieve the outcome.

Machine Learning, ML requires a stable environment for learning. Federated learning ensures that this environment has been created and privacy has been preserved [34]. The new FL platform ensures that privacy-related vulnerabilities can be resolved. Resolving these challenges is among the main functions associated with federated learning. Federated learning can therefore be defined to be an alternative to the cloud-centric ML approaches. The model has predominantly focused on resolving issues from models such as the traditional cloud-centric approach. Most of the applications can be seen to have already achieved this aspect.

Model learning is among the main concepts associated with the development needed for promoting the model algorithm. The decentralized and model learning approaches have achieved the security available [35]. Privacy is a leading concept required, mainly achieved from the model available. The procedure was carefully executed and ensured that people benefited significantly in the process [36]. Federated learning in an overall analysis has promoted the privacy required in the process [37]. FL settings and applications have been approached and achieved the requirements available [38][39][40].

4.1. Differential Privacy

Federated learning has the possibility of becoming more effective in the future as more of its challenges are eradicated. The possibilities associated with the general applications of federated learning have been described as adequate to promote more applications in the future [41]. The technologies related to this growth and development have been told to highly promote growth and developments within AI and cybersecurity [42]. Federated learning has therefore promoted the required set of developments globally and ensured that different people could work more effectively and promote specific developments within their lives [13][29].

Differential privacy is a technique that allows researchers and database analysts to extract useful information from databases that contain people's personal information (PII) without revealing it. This is possible by introducing a small amount of distraction into the data provided by the database or dataset [39]. A differential privacy method, for example, inserts random data into an original data set during collecting, concealing individual data points before further anonymizing on a server. The introduced distraction should be broad enough to protect privacy while remaining confined to keep the information presented to analysts valuable.

Differential privacy can be employed in various contexts, such as recommendation systems, social networks, and location-based applications. It works well with large amounts of data, but scaling is complex unless other approaches, such as federated learning, are applied. This is especially important in sectors where data is sensitive, and privacy is critical [39]. Healthcare, financial services (FinTech), and the Internet of Things as it scales throughout our homes and cities are just a few instances of how differentiated privacy produces better results.

5. PRESERVING PII DATA IN FINTECH

FinTech is among the technologies that essentially appreciate cyber security. FinTech mainly comprises technologies and software focused on financial applications [39]. The increased usage of these technologies is primarily supported by the growing need for developing these technologies [38]. The field of finance has been improved over the last few decades by engaging technologies in the area. This approach has been associated with increased effectiveness in the financial world [14]. More people have included technologies across financial departments to ensure that it is more secure.

FinTech requires that an extensive amount of security be applied. Security is a significant concern for most financial wizards [41]. The required level of security ensures that people safely deposit their money across different banks and send the money when needed [42]. Business owners especially require these technologies since they promote the required type of growth across the community [36]. Therefore, the technology associated with finance applications is an essential part of the world. Its applications have been described to have a specific meaning in achieving growth for companies [43]. Algorithms associated with promoting security have been developed to ensure that the entire field of finance has been appreciated and is secure for all types of transactions. It has also contributed to reducing fraud across most departments of finance.

Estevez (2020) indicates FinTech first emerged in the 21st century and was applied in the back-end systems of big financial institutions [41]. The technologies proved very effective, resulting in an appreciation of the technology across consumer-oriented services. The industries now associated with FinTech include education, retail banking, fundraising, non-profits, and investment companies [44]. Each of these companies and industries now achieves a specific type of growth by ensuring financing effectiveness [45]. The overall applications of this type of development have applied trust across the finance sectors of these organizations.

The use of crypto is another significant development in FinTech. FinTech allows for sending and receiving crypto to be available and capable for different financial companies [26]. Promoting this type of development ensures that people can effectively use different types of finances [46]. The increased technological development across this technology has effectively promoted a specific result for most individuals. The growth included has been described as having a needed type of development and a particular type of growth. AI has promoted this type of application for most users.

Federated learning has been one of the critical models applied to ensure that machine learning has been engaged in promoting FinTech. Federated learning has been described as a practical approach to increasing the applications of AI and machine learning [47]. Machine learning ensures that devices are made more applicable and that they have attained the required level of security. The federated machine learning concept promotes the necessary development level in FinTech [48]. Financial experts believe that the development of machine learning has an overall promotion of the required growth in financial situations. This type of development is described as having a huge potential in ensuring financial technology has become effective for all parties.

Federated machine learning has been described as a new evolution that makes people's lives easier [49]. The solution has been related to effectively promoting a specific type of development for objectives and responsibilities in the financial realm. Ligade (2020) suggests federated machine learning is a new concept that has also been developed in response to the available development and how it impacts people in general [40]. It describes the exact type of influence that people may acquire for the required product development [34]. Taking advantage of federated machine learning has been described as a correct approach when dealing with FinTech (see Fig 2). A fleet of autonomous devices in FinTech may require an up-to-date model to investigate process and system behavior to function correctly while constructing aggregate models. Due to the private nature of particular instances, accessing the data and each device's limited connectivity may be impossible. Federated Learning approaches can aid in developing models that can adapt to changing conditions while protecting user privacy. When dealing with devices with censors or IoT connectivity, these learning techniques add more value in delivering better output instead of connecting them centrally.

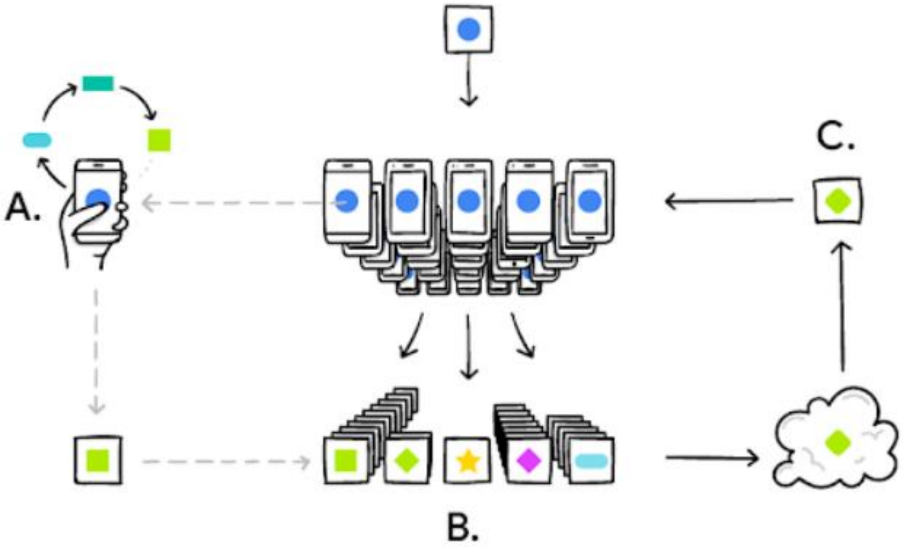

Fig 2. Federated Learning model in FinTech (A-Personalized TinyML Model, B-Updates are gathered from different users, C- Consensus global model) [40].

Data-protection laws heavily protect FinTech. Most of these laws focus on promoting people and how well they engage in several factors [50]. The overall approach has been described as having a specific requirement when dealing with people's finances [23]. Financing, therefore, protects a lot of people from the different approaches taken by federated learning [49]. Federated learning has adhered to most of these laws and ensured that it is more effective and that each of these considerations has been made.

Global privacy demands that the model updates created at each round be private to all untrusted third parties other than the central server, as shown in Figure 3, while local privacy requires that the changes be secret to the server as well. Individual model changes are safeguarded via a secure aggregation approach. Although the central server cannot access any local modifications, it may nevertheless view the identical aggregated results at each round. Secure aggregation is a lossless approach that can keep original correctness while maintaining a high level of secrecy; nevertheless, the resulting method incurs considerable additional transmission costs. Other works use differential privacy in federated learning and provide global differential privacy; these systems have a lot of hyper parameters that impact communication and accuracy and must be carefully selected [50].

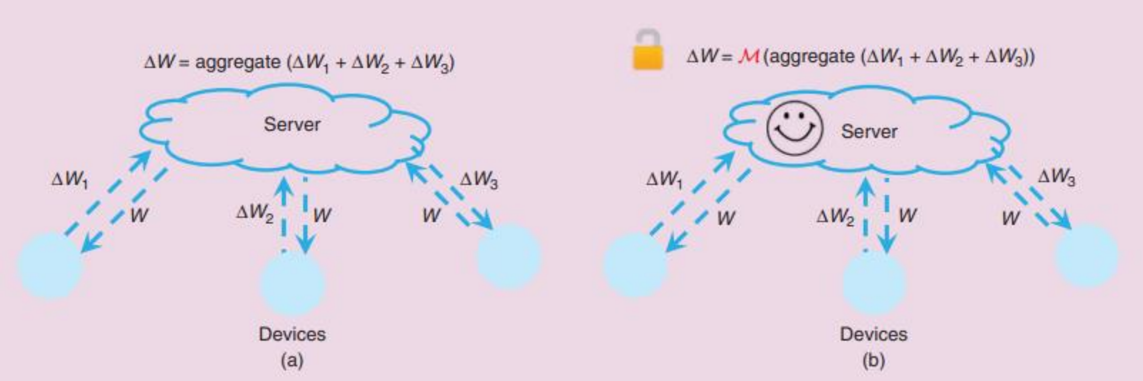
ΔW = aggregate ($\Delta W_1 + \Delta W_2 + \Delta W_3$)
Server
ΔW_1
W
ΔW_2
W
W
ΔW_3
Devices
(a)
$\Delta W = \mathcal{M}$(aggregate ($\Delta W_1 + \Delta W_2 + \Delta W_3$))
Server
ΔW_1
W
ΔW_2
W
W
ΔW_3
Devices
(b)

Fig 3. Illustration of privacy enhancing mechanisms in FL. Mdenotes a randomized mechanism used to privatize the data. (a) FL without additional privacy protection mechanisms, (b) FL with global privacy with trusted server [54].

The different features and applications of federated learning have been found to effectively promote the required growth across the financing department around the globe [46]. The technology associated with financial solutions has integrated machine learning and AI at an increasingly positive rate [51,52]. This inclusion has continued to ensure that people have been promoted and that effective results have been achieved. Ligade (2020) indicates that the entire focus on these applications has especially attained positive results as indicated by current financial technologies [40]. Machine learning has ensured that more developments are being acquired daily and that security and privacy are leading aspects [48]. FinTech has allowed machine learning to be included when dealing with the required data. Therefore, the data contained is promoted through the available federated learning and its effectiveness. Federated learning can thus be described as effectively promoting a specific type of development within financial solutions [19]. Continued federated learning development ensures machine learning and financial technologies' success in AI and building a sustainable carbon-free ecosystem [53].

6. Challenges

There are many core challenges dealing and operating with federated learning techniques. Those challenges are outlined below.

- ***Expensive Communication***

  It is critical to design communication-efficient approaches that deliver short messages or model changes repeatedly as part of the training process, rather than sending the complete data set over the network. Because this process is carried out in millions of little devices, connection, bandwidth, and power are crucial for sustaining these activities. The best two options for making these processes more efficient and reducing communication in phases are: 1) lowering the total number of communication rounds; and 2) reducing the amount of the messages communicated at each round [54].

- ***Managing System Heterogeneity***

  Because of differences in hardware (CPU and memory), network connectivity (3G, 4G, 5G, and Wi-Fi), and power supply, the storage, computing, and communication capabilities of each device in federated networks may change. With millions of devices interconnected, it's not surprising that a few go down throughout the process, causing interferences by dropping out. These system-level properties increase issues like straggler mitigation and fault tolerance. Hence, FL techniques should: 1) promote a modest level of engagement; 2) use heterogeneous technology; and 3) be resilient enough to establish strategies for dealing with lost devices [55].

- ***Handling Statistical Heterogeneity***

  Devices regularly create and gather data in a very non identically dispersed manner across the network; for example, mobile phone users employ language in a variety of ways when doing a next-word prediction job. This presents a significant barrier in managing and utilizing same data for worldwide forecast. As a result, there are various options, such as learning distinct local models at the same time using multitask learning frameworks [55]. There is also a tight relationship between leading methodologies for federated learning and meta learning in this regard. Both the multitask and meta learning approaches provide customized or device-specific modeling, which is frequently a more natural method to dealing with statistical heterogeneity in data for greater customization.

7. CONCLUSION

Finally, we learned that federated learning has numerous usability that boosts FinTech. Federated Learning necessitates adopting new tools and a new way of thinking by machine learning practitioners: model building, training, and evaluation without direct access to or labeling raw data, with communication costs as a limiting factor. FinTech is one of the industries that can benefit from engaging federated learning apps. The report goes into great detail about federated learning and how it affects people's lives today. The examination of how FinTech has benefited from federated learning demonstrates the various ways in which people gain from enhanced privacy. When it comes to FinTech, privacy is the primary reason why federated learning is used. Most organizations and data scientists are highlighted, as well as how they have profited from the growing use of federated learning. Therefore, the continued development and research in federated learning promises growth in FinTech.

We can see that the number, quality, and distribution of data significantly impact the FL output quality. It will be fascinating to see how the differential privacy component affects the federated learning output in FinTech in the future.

REFERENCES

[1] Bonawitz Kallista, Peter Kairouz, Brendan McMahan, and Daniel Ramage. Federated Learning and Privacy: Building privacy-preserving systems for machine learning and data science on decentralized data. https://queue.acm.org/detail.cfm?id=3501293

[2] Cheng Yong, Yang Liu, Tianjian Chen, Qiang Yang. 2020. Federated Learning for Privacy-Preserving AI. CACM. https://cacm.acm.org/magazines/2020/12/248796-federated-learning-for-privacy-preserving-ai/fulltext

[3] Bonawitz, K. et al. Practical secure aggregation for privacy-preserving machine learning. In Proceedings of ACM SIGSAC CCS'17 (Nov. 2017).

[4] DLA Piper. Data protection laws of the world: Full handbook (Jan. 2020); https://bit.ly/354nDiC

[5] Kairouz, P. et al. Advances and open problems in federated learning. (Dec. 2019); arXiv preprint arXiv:1912.04977

[6] Liu, Y., Chen, T., and Yang, Q. Secure federated transfer learning. In Proceedings of IJCAI'19 (Aug. 2019).

[7] McMahan, H.B., Moore, E., Ramage, D., and y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of AISTATS'17 (Apr. 2017).

[8] Richardson, A., Filos-Ratsikas, A., and Faltings, B. Rewarding high-quality data via influence functions. (Aug. 2019); arXiv preprint arXiv:1908.11598

[9] Yang, Q. et al. Federated machine learning: Concept and applications. ACM Trans. Intell. Syst. Technol. (TIST) (Feb. 2019).

[10] Yang, Q. et al. Federated Learning. Morgan & Claypool, Dec. 2019.

[11] H. Zheng, H. Hu and Z. Han, "Preserving User Privacy for Machine Learning: Local Differential Privacy or Federated Machine Learning?", IEEE Intelligent Systems, vol. 35, no. 4, pp. 5-14, 2020. Available: 10.1109/mis.2020.3010335.

[12] M. Kim and J. Lee, "Information-theoretic privacy in federated submodel learning", ICT Express, 2022. Available: 10.1016/j.icte.2022.02.008.

[13] H. Fang and Q. Qian, "Privacy Preserving Machine Learning with Homomorphic Encryption and Federated Learning", Future Internet, vol. 13, no. 4, p. 94, 2021. Available: 10.3390/fi13040094.

[14] Woubie, Abraham, and Tom Backstrom. "Federated Learning For Privacy-Preserving Speaker Recognition". IEEE Access, vol 9, 2021, pp. 149477-149485. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/access.2021.3124029.

[15] Park, Jaehyoung, and Hyuk Lim. "Privacy-Preserving Federated Learning Using Homomorphic Encryption". Applied Sciences, vol 12, no. 2, 2022, p. 734. MDPI AG, doi:10.3390/app12020734.

[16] Śmietanka, Małgorzata et al. "Federated Learning For Privacy-Preserving Data Access". SSRN Electronic Journal, 2020. Elsevier BV, doi:10.2139/ssrn.3696609.

[17] Jere, Malhar S. et al. "A Taxonomy Of Attacks On Federated Learning". IEEE Security &Amp; Privacy, vol 19, no. 2, 2021, pp. 20-28. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/msec.2020.3039941.

[18] Jiang, Xue et al. "Comprehensive Analysis Of Privacy Leakage In Vertical Federated Learning During Prediction". Proceedings On Privacy Enhancing Technologies, vol 2022, no. 2, 2022, pp. 263-281. Walter De Gruyter Gmbh, doi:10.2478/popets-2022-0045.

[19] Bonawitz, Kallista et al. "Federated Learning And Privacy". Queue, vol 19, no. 5, 2021, pp. 87-114. Association For Computing Machinery (ACM), doi:10.1145/3494834.3500240.

[20] Gálvez, Rafa et al. "Less Is More: A Privacy-Respecting Android Malware Classifier Using Federated Learning". Proceedings On Privacy Enhancing Technologies, vol 2021, no. 4, 2021, pp. 96-116. Walter De Gruyter Gmbh, doi:10.2478/popets-2021-0062.

[21] Treleaven, Philip et al. "Federated Learning: The Pioneering Distributed Machine Learning And Privacy-Preserving Data Technology". Computer, vol 55, no. 4, 2022, pp. 20-29. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/mc.2021.3052390.

[22] Chamikara, M.A.P. et al. "Privacy Preserving Distributed Machine Learning With Federated Learning". Computer Communications, vol 171, 2021, pp. 112-125. Elsevier BV, doi:10.1016/j.comcom.2021.02.014.

[23] Qin, YangJie et al. "Privacy-Preserving Federated Learning Framework In Multimedia Courses Recommendation". Wireless Networks, 2022. Springer Science And Business Media LLC, doi:10.1007/s11276-021-02854-1.

[24] Ouadrhiri, Ahmed El, and Ahmed Abdelhadi. "Differential Privacy For Deep And Federated Learning: A Survey". IEEE Access, vol 10, 2022, pp. 22359-22380. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/access.2022.3151670.

[25] Mahmood, Zeba, and Vacius Jusas. "Blockchain-Enabled: Multi-Layered Security Federated Learning Platform For Preserving Data Privacy". Electronics, vol 11, no. 10, 2022, p. 1624. MDPI AG, doi:10.3390/electronics11101624.

[26] Kim, Sungwook. "Incentive Design And Differential Privacy Based Federated Learning: A Mechanism Design Perspective". IEEE Access, vol 8, 2020, pp. 187317-187325. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/access.2020.3030888.

[27] Zhang, Wu Ming et al. "Privacy-Preserving Federated Learning With Collusion-Resistance In Mobile Crowdsensing". SSRN Electronic Journal, 2022. Elsevier BV, doi:10.2139/ssrn.4104451.

[28] Asad, Muhammad et al. "A Critical Evaluation Of Privacy And Security Threats In Federated Learning". Sensors, vol 20, no. 24, 2020, p. 7182. MDPI AG, doi:10.3390/s20247182.

[29] Jiang, Changsong et al. "PFLM: Privacy-Preserving Federated Learning With Membership Proof". Information Sciences, vol 576, 2021, pp. 288-311. Elsevier BV, doi:10.1016/j.ins.2021.05.077.

[30] Kim, Kyongjin, and Sengphil Hong. "The Data Processing Approach For Preserving Personal Data In Fintech-Driven Paradigm". International Journal Of Security And Its Applications, vol 10, no. 10, 2016, pp. 341-350. NADIA, doi:10.14257/ijsia.2016.10.10.30.

[31] Macpherson, Martina et al. "Artificial Intelligence And Fintech Technologies For ESG Data And Analysis". SSRN Electronic Journal, 2021. Elsevier BV, doi:10.2139/ssrn.3790774.

[32] Arner, Douglas W. et al. "Fintech And Regtech: Enabling Innovation While Preserving Financial Stability". SSRN Electronic Journal, 2017. Elsevier BV, doi:10.2139/ssrn.3211708.

[33] Kantarcioglu, Murat, and Wei Jiang. "Incentive Compatible Privacy-Preserving Data Analysis". IEEE Transactions On Knowledge And Data Engineering, vol 25, no. 6, 2013, pp. 1323-1335. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/tkde.2012.61.

[34] Wang, J. Christina, and Charles B. Perkins. "How Magic A Bullet Is Machine Learning For Credit Analysis? An Exploration With Fintech Lending Data". SSRN Electronic Journal, 2019. Elsevier BV, doi:10.2139/ssrn.3928076.

[35] Zhai, Yimeng. "Analysis Of Fintech Regulation Based On G-Sibs Fintech Index". Journal Of Finance Research, vol 4, no. 1, 2020, p. 69. Synergy Publishing Pte. Ltd., doi:10.26549/jfr.v4i1.3250.

[36] Sevilmiş, Fehmi, and Hulusi Karaca. "Performance Analysis Of SRF-PLL And DDSRF-PLL Algorithms For Grid Interactive Inverters". International Advanced Researches And Engineering Journal, 2019, pp. 116-122. International Advanced Researches And Engineering Journal, doi:10.35860/iarej.412250.

[37] Najaf, Khakan et al. "Var And Market Value Of Fintech Companies: An Analysis And Evidence From Global Data". Managerial finance, vol 47, no. 7, 2020, pp. 915-936. Emerald, doi:10.1108/mf-04-2020-0169.

[38] Dash, B., & Ansari, M. F. (2022). An Effective Cybersecurity Awareness Training Model: First Defense of an Organizational Security Strategy.

[39] B., Kirti, and B. M. "New Approach To Reduce The Data Loss In Privacy Preserving Data Analysis". International Journal Of Computer Applications, vol 155, no. 12, 2016, pp. 21-24. Foundation Of Computer Science, doi:10.5120/ijca2016912504.

[40] Ligade Maheshwar (2020). Federated machine learning for fintech. https://medium.com/techwasti/federated-machine-learning-for-fintech-b875b918c5fe

[41] Estevez Eric. (2020). Financial Technology – Fintech. What Is Financial Technology – Fintech? https://www.investopedia.com/terms/f/fintech.asp

[42] Ferdinand, Al-Lawati, Draper, Nokleby, 2020 N. Ferdinand, H. Al-Lawati, S.C. Draper, M. Nokleby Anytime minibatch: exploiting stragglers in online distributed optimizationarXiv preprint arXiv:2006.05752 (2020)

[43] Al-Rubaie, M., and Chang, J.M. Privacy-preserving machine learning: Threats and solutions. IEEE Security and Privacy (Apr. 2019).

[44] Shi, Yuan, and Xianze Xu. "Deep Federated Adaptation: An Adaptative Residential Load Forecasting Approach With Federated Learning". Sensors, vol 22, no. 9, 2022, p. 3264. MDPI AG, doi:10.3390/s22093264.

[45] Nayak, Sanjib Kumar et al. "Modeling And Forecasting Cryptocurrency Closing Prices With Rao Algorithm-Based Artificial Neural Networks: A Machine Learning Approach". Fintech, vol 1, no. 1, 2021, pp. 47-62. MDPI AG, doi:10.3390/fintech1010004.

[46] Ghimire, Bimal, and Danda B. Rawat. "Recent Advances On Federated Learning For Cybersecurity And Cybersecurity For Federated Learning For Internet Of Things". IEEE Internet Of Things Journal, vol 9, no. 11, 2022, pp. 8229-8249. Institute Of Electrical And Electronics Engineers (IEEE), doi:10.1109/jiot.2022.3150363.

[47] Bao, Hong, and David Roubaud. "Recent Development In Fintech: Non-Fungible Token". Fintech, vol 1, no. 1, 2021, pp. 44-46. MDPI AG, doi:10.3390/fintech1010003.

[48] Fang, Haokun, and Quan Qian. "Privacy Preserving Machine Learning With Homomorphic Encryption And Federated Learning". Future Internet, vol 13, no. 4, 2021, p. 94. MDPI AG, doi:10.3390/fi13040094.

[49] Bazarbash, Majid. "Fintech In Financial Inclusion: Machine Learning Applications In Assessing Credit Risk". SSRN Electronic Journal, 2019. Elsevier BV, doi:10.2139/ssrn.3404066.

[50] Brendan McMahan, H., Ramage, D., Talwar, K., & Zhang, L. (2017). Learning differentially private recurrent language models. *ArXiv e-prints*, arXiv-1710.

[51] Loh, Leonard Kin Yung et al. "An Ensembling Architecture Incorporating Machine Learning Models and Genetic Algorithm Optimization for Forex Trading". Fintech, vol 1, no. 2, 2022, pp. 100-124. MDPI AG, doi:10.3390/fintech1020008.

[52] Dash, B. (2021). A hybrid solution for extracting information from unstructured data using optical character recognition (OCR) with natural language processing (NLP).

[53] Sharma, P., & Dash, B. THE DIGITAL CARBON FOOTPRINT: THREAT TO AN ENVIRONMENTALLY SUSTAINABLE FUTURE. IJCSIT, doi: 10.5121/ijcsit.2022.14302.

[54] Aytaç, K., & Korçak, Ö. (2021). IoT based intelligence for proactive waste management in Quick Service Restaurants. *Journal of Cleaner Production*, *284*, 125401.

[55] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50-60.

AUTHORS

Bibbu Dash is an Architect-Data and Analytics in a Fortune 100 financial organization at Madision, WI. He is currently a Ph.D. student in Information Technology at the University of the Cumberlands, Kentucky. Bibhu has completed his Master of Engineering in Electronics and Communication Engg., and MBA from Illinois State University, Normal, IL. Bibhu's research interests are in the areas of AI, Cloud Computing, Big Data and Blockchain technologies.

Pawankumar Sharma is a Senior Product Manager for Walmart at San Bruno, California. He is currently on his Ph.D. in Information Technology at the University of the Cumberlands, Kentucky. Pawankumar Sharma has completed his Master of Science in Management Information Systems from the University of Nebraska at Omaha in 2015. He also holds another Master of Science in Information Systems Security from the University of the Cumber lands, Kentucky and graduated in 2020. His research interests are in the areas of Cyber security, Artificial Intelligence, Cloud Computing, Neural Networks, Information Systems, Big Data Analytics, Intrusion Detection and Prevention.

Azad Ali, D.Sc., Professor of Information Technology, has more than 30 years of combined experience in the areas of financial and information systems. He holds a bachelor's degree in Business Administration from the University of Baghdad, an MBA from the Indiana University of Pennsylvania, an MPA from the University of Pittsburgh, and a Doctor of Science in Communications and Information Systems from Robert Morris University. Dr. Ali's research interests include service-learning projects, web design tools, dealing with isolation in doctoral programs, and curriculum development. Azad has been involved in mentoring doctoral students to complete their doctoral dissertations and has so far mentored five students to complete their dissertations.