Open Access Library Journal
2024, Volume 11, e12239
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Advances, Challenges & Recent Developments in Federated Learning

Nsie Erimola María Reina Agripina*, Blessed Shinga Mafukidze

Department of Computer Science, Hubei University of Technology, Wuhan, China
Email: *damarisnsie23@icloud.com

How to cite this paper: Agripina, N.E.M.R. and Mafukidze, B.S. (2024) Advances, Challenges & Recent Developments in Federated Learning. *Open Access Library Journal*, **11**: e12239.
https://doi.org/10.4236/oalib.1112239

Received: September 4, 2024
Accepted: October 18, 2024
Published: October 21, 2024

Abstract

This has led to the rise of a paradigm shift in machine learning called federated learning (FL) that allows for decentralized model training over distributed data sources. With FL, devices, servers, or edges train the model together without sharing their privacy-sensitive data, effectively addressing the arising data privacy regulation, data residency, and data silos types of issues, among many others. The FL ecosystem has also been through a series of significant developments, leading to the emergence of secure aggregation protocols and federated optimization techniques for better model convergence and performance, though there are still critical roadblocks such as data heterogeneity, communication overhead, and vulnerability to attacks. This paper aims to summarize the current progress, practical limitations, and future research directions on the application of FL, particularly in the healthcare, finance, and Internet of Things domains, as a means of preserving privacy and enhancing learning. The future entails the incorporation of edge computing, decentralized learning frameworks, and privacy-preserving techniques into the picture that has the potential to reshape today's state-of-the-art FL.

Subject Areas

Information Management, Machine Learning

Keywords

Federated Learning, Decentralized Technology, Machine Learning, Data

1. Introduction

Federated learning has emerged as a transformative paradigm in machine learning, offering a decentralized approach to collaboratively train models across distributed data sources while preserving data privacy and security. This innovative

framework enables organizations to leverage data from multiple devices, servers, or edge devices without centralizing sensitive information, addressing the challenges associated with data silos, privacy regulations, and data residency requirements. Recent interest in this field has been high, both from research and an applied perspective. This monograph outlines the distinctive features and difficulties of the federated learning environment, emphasizes significant practical limitations and considerations, and then lists a variety of fruitful research directions. This work's objectives are to draw attention to research issues that are both theoretically and practically significant and to encourage research on problems that could have significant real-world impact.

The International Data Corporation predicts that by 2025, there will be 80 billion Internet-connected devices, and the amount of data generated could reach 180 trillion gigabytes. On the other hand, those computation-intensive machine learning applications can finally be run on devices rather than centralized cloud data centers thanks to the evolution of powerful computation hardware design and efficient computing architecture, like parallel high-performance graphics processing units (GPUs), which also promotes the widespread use of machine learning. However, it is known that a machine learning model's training requires a significant amount of data, whereas the amount of data produced by a single device is constrained. In addition, because of the diverse property. The machine learning model created from one device is difficult to work with and desirable for others in terms of individual behavior characteristics.

1.1. Overview of Federated Learning (FL)

Federated Learning (FL) has emerged as a promising approach for collaborative model training across distributed devices while preserving data privacy and security. This field has witnessed significant advances, notable challenges, and ongoing developments that are shaping the future of decentralized machine learning. One of the key advances in FL research is the development of more efficient algorithms and protocols for secure aggregation of model updates from multiple clients. Secure aggregation techniques, such as encryption and differential privacy, help ensure that individual contributions remain private while still enabling collaborative model learning. Additionally, advancements in federated optimization methods, like Federated Averaging and Federated Learning with Bayesian Neural Networks, have improved model convergence and performance across distributed datasets McMahan (2016) [1].

Despite these advances, FL faces several challenges that need to be addressed to realize its full potential. One of the primary challenges is the heterogeneity of data distributions and quality across devices, which can lead to biases and model performance disparities. Research efforts are ongoing to develop robust federated learning frameworks that can mitigate the impact of data distribution mismatches and improve model generalization across diverse data sources. FL systems must also contend with issues related to communication and computational overhead,

as well as network failures and adversarial attacks that could compromise model privacy and integrity. To address these challenges, researchers are exploring novel solutions, such as decentralized learning scheduling algorithms, adaptive federated optimizers, and secure federated learning frameworks with differential privacy guarantees He (2020) [2]. Recent developments in FL research have focused on expanding the applicability of federated learning to new domains, such as healthcare, financial services, and IoT networks, where data privacy and regulatory compliance are paramount. Collaborative efforts between academia, industry, and policymakers are driving the development of standards and best practices for federated learning deployment in real-world settings. Additionally, the integration of edge computing capabilities and federated learning technologies is paving the way for decentralized AI applications that deliver personalized services and insights at the network edge Yang (2019) [3].

Federated learning continues to evolve as a transformative paradigm for collaborative and privacy-preserving machine learning. Advances in algorithms, challenges in data heterogeneity and security, and recent developments in new application domains are shaping the landscape of FL research and innovation, laying the foundation for the next generation of decentralized AI systems.

1.2. Exploring the Evolution of Edge Computing: A Focus on Federated Learning (FL)

Edge computing is not a brand-new idea. In fact, basic query processing across distributed, low-power devices has been studied for decades in the contexts of computing at the edge, fog computing, and query processing in sensor networks. In recent works, the idea of centrally training machine learning models while serving and storing them locally has also been explored. However, it is possible to use improved local resources on each device as the storage and computational capacities of distributed networks' devices increase. Due to this, federated learning, which investigates training statistical models on distant devices directly, is gaining popularity.

In order to dispatch data distribution and computing resources effectively and create an effective cooperation model, parameter servers, atypical distributed and centralized technology, primarily require a central server Ho (2013) [4]. There is a double communication overhead as a result of this type of centralized data processing technology.

The data must be copied if they are obtained for training from many databases that are distributed throughout the central server at first. Then, for distributed computation, a central server will distribute data to each distributed client. It intensifies the system's rigorous assessments for computing power, storage, and bandwidth. For cases in FL, each client is entirely independent, data is not allocated by the center and the training process is not directed by the server. Therefore, FL is an integrated technology that combines machine learning models with data fusion through decentralized collaboration.

2. Characteristics of Federating Learning (FL)

Universality for cross-organizational scenarios FL, as proposed by Google, is essentially an encrypted distributed machine learning technique that enables users to create a training model jointly while maintaining the underlying data locally. The original then According to Yang (2019) [3], the idea of FL has been expanded to include any privacy-preserving decentralized collaborative machine learning techniques. As a result, FL can handle both vertically partitioned data according to features and horizontally partitioned data according to samples in a collaborative learning environment. Cross-organizational enterprises could be integrated into the federal framework by expanding FL. For instance, a bank with information on customers' purchasing power could work with an online marketplace that has information on product features to suggest purchases.

2.1. Decentralized Training

Federated Learning (FL) refers to the process of training machine learning models on multiple devices or servers without the need to bring all the data to a central location. Instead of sending raw data to a centralized server for model training, FL distributes the learning process across individual devices or edge servers, where data is stored and processed locally. In a decentralized training setup, each device or server participates in model training by computing model updates based on its local data and then sending only the updated model parameters to a central server or aggregator. The central server aggregates these partial model updates from multiple devices to create a global model that reflects insights from the diverse data sources participating in the FL process. By keeping data decentralized and localized on individual devices, decentralized training helps protect the privacy of sensitive information. Personal data remains under the control of the user and is not exposed to potential security risks associated with centralized data storage or transmission. This approach also enables users to retain ownership and control over their data while still benefiting from the collective intelligence captured in the global model. This is illustrated in Figure 1.

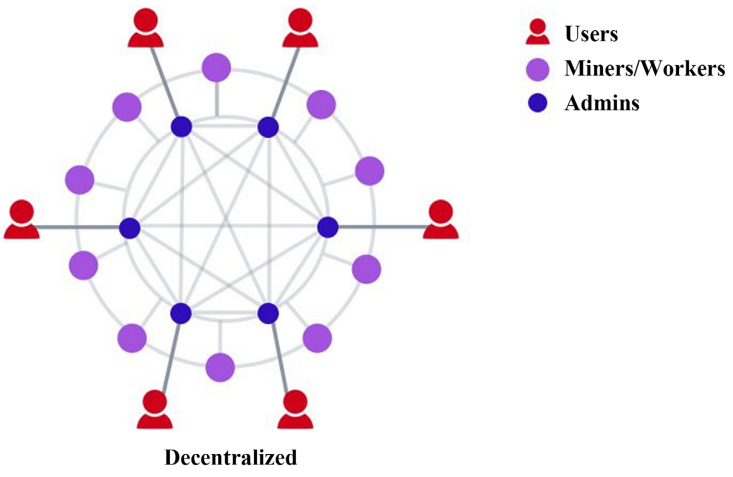


Figure 1. Beltrán *et al.* (2023) [5], Decentralized Technology.

Decentralized training in FL promotes collaboration and knowledge sharing among participants without compromising data privacy, allowing organizations to leverage distributed data sources for model improvement and performance enhancement. Additionally, this approach enhances scalability and efficiency by overcoming challenges related to data silos, network latency, and data privacy regulations, making FL an attractive solution for privacy-conscious and data-sensitive applications in various domains

2.2. Equality and Empowerment in Federated Learning for Shared Prosperity

All parties benefit from equal standing and definite dominion in this framework for cooperation in order to achieve shared prosperity. In terms of equality, the person who has access to a large amount of data typically holds the dominating position in collaborative training that is spread. Therefore, the preference for companies with large amounts of data or photographs with various types of labels could have a negative impact on the growth of collaborative learning in the industrial field. Small and medium-sized businesses lack motivation for cooperative training in deep learning networks because institutions with big data can manipulate the prediction model. However, due to equality among all parties in FL, the standing of these clients with modest data sets would be elevated.

In conclusion, as shown in Figure 1. FL is a decentralized solution that enables dispersed customers or organizations to autonomously train a collaborative model while maintaining localized data. This approach can assist business organizations in sharing collaboration without transferring any raw data.

3. Classification of Federated Learning (FL)

3.1. Vertical Federated Learning

This method is employed when each device has a data set with unique attributes derived from sample examples. For example, Vertical FL can be used to create a shared machine learning model across two organizations that have data about the same group of individuals but distinct feature sets. In healthcare, one entity may have electronic health records containing patient demographics and medical history, while another entity may have lab test results. By using Vertical Federated Learning, both entities can collaborate to train a model that predicts patient outcomes or diagnoses without sharing their sensitive data.

3.2. Horizontal Federated Learning

Horizontal Federated Learning is a form of Federated Learning where multiple clients or devices with similar data distributions collaborate to train a global machine learning model while keeping their data locally. In horizontal Federated Learning, each client or device has its own data. This method is utilized when each device has a data set with the same feature space but a separate collection of sample instances. This sort of learning, in which the participating mobile phones have

various training data with the same properties, is used in the initial FL Google keyboard use case. This is illustrated in Figure 2.

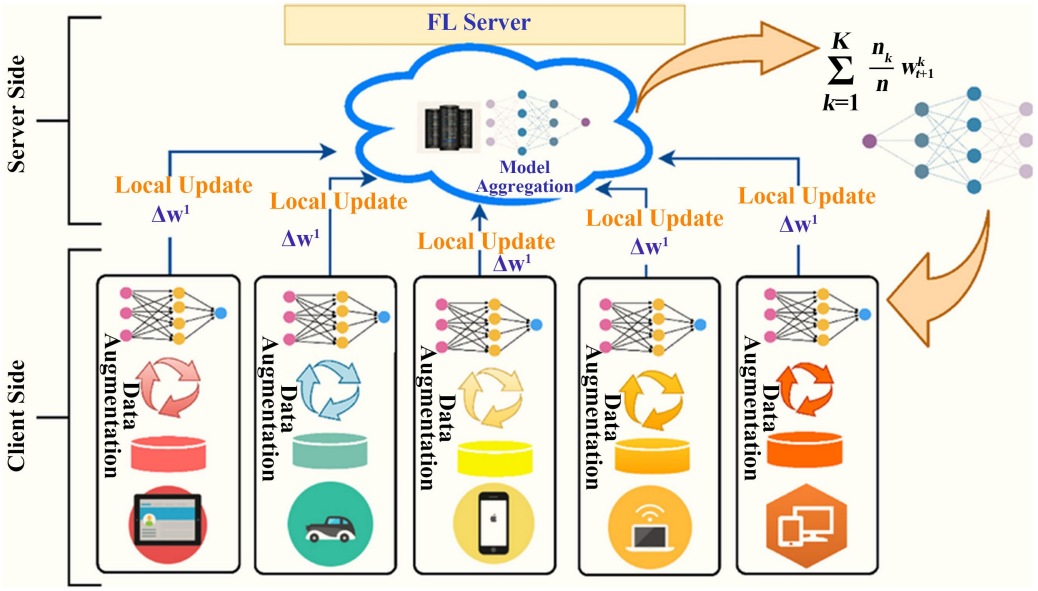


Figure 2. Shaheen *et al.* (2024) [6], Application of Federated learning: Taxonomy, Challenges, Research Trends.

Horizontal Federated Learning is particularly useful when dealing with multiple distributed data sources that share similar characteristics but cannot be combined due to privacy or regulatory constraints.

3.2.1. Cross-Silo Federated Learning

When there are fewer participating devices and they are accessible throughout all rounds, Cross-Silo Federated Learning is used. The training data may be in FL format, either horizontally or vertically. Cross-silo is mostly utilized for scenarios involving organizations. Works like making use of cross-silo to help FL create their model. In Cross-Silo Federated Learning, each entity trains a local model on its own data and then shares only model updates with other entities instead of raw data. These model updates are aggregated and used to improve the global model, which benefits from the diversity and richness of data from multiple sources without compromising data privacy or security. For example, in the context of mobile device manufacturers collaborating to improve predictive text suggestions, each manufacturer may have data on the language usage patterns and typing habits of their users. By utilizing Cross-Silo Federated Learning, these manufacturers can collaboratively train a more accurate predictive text model by leveraging the unique insights from each dataset without sharing raw user data.

3.2.2. Cross-Device Federated Learning

Scenarios with a large number of participating devices use Crossdevice Federated Learning. Client selection and incentive designs are two prominent strategies required to support this kind of FL. In Cross-Device Federated Learning, the model

is trained locally on each device using its respective data, and model updates are then aggregated and shared among devices. This allows for personalized model training while benefiting from the diversity of data from various devices. By sharing only model updates and not raw data, privacy and security are maintained. For example, consider a user who uses a smartphone, tablet, and laptop for various tasks such as web browsing, email, and social media. These devices can collaborate through Cross-Device Federated Learning to train a personalized, predictive text model that adapts to the user's typing patterns and language preferences across all devices.

4. Practical Uses of Federated Learning

Federated Learning offers a versatile and privacy-preserving approach to collaborative model training, making it suitable for a wide range of applications across various industries where data privacy and security are paramount. Federated learning is crucial for supporting distributed training data applications that deal with privacy-sensitive data. Theoretically, it appears to be a great strategy for resolving issues with data stored on a single central server or location that arise in typical machine learning models. This is illustrated in Figure 3.

Figure 3. Farooq *et al.* (2023) [7], an example application of federated learning for the task of next-word prediction on mobile phones.

4.1. Practical Use in Autonomous Vehicle

Federated learning is used to create self-driving cars since it can make predictions in real-time. The data may contain realtime updates on the state of the roads and traffic, enabling continual learning and quicker decision-making. This might lead to a safer and more fun self-driving car experience. A prospective field for the application of federated machine learning is the automotive industry. But at the moment, research is the only thing being done in this area. Federated learning may reduce training time for predicting the steering angle of self-driving cars, according to one study.

4.2. Practical Use in Healthcare

For machine learning applications, electronic health records (EHR) are regarded as the primary source of healthcare data. If only the scant data present in a particular hospital is used to train ML models. The predictions may become somewhat biased as a result.

As a result, training with more data is necessary to make the models more generalizable. This can be accomplished by exchanging data among businesses. Sharing patient electronic health information between hospitals may not be possible due to the sensitive nature of healthcare data. Federated learning may be an alternative for creating a collaborative learning model for healthcare data in such circumstances. This method may leverage a significant quantity of data from multiple healthcare databases and devices to construct AI models while remaining compliant with rules.

4.3. Practical Use in Finance

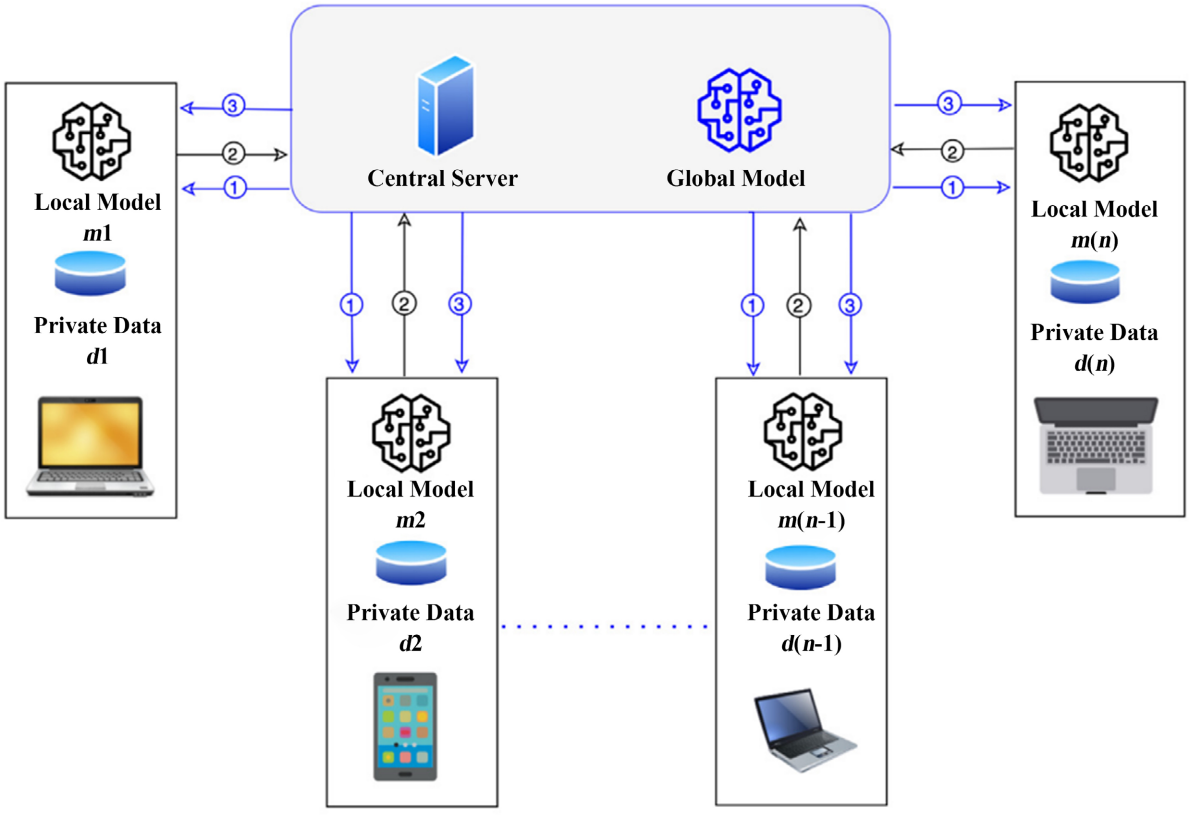
In the banking industry, federated learning is used most effectively for loan risk assessment. The majority of the time, banks employ whitelisting strategies to exclude clients by using credit card records from the central banks. Elements like taxation and reputation etc., can be used to manage risk by working with other financial institutions and e-commerce businesses. Since it is risky for businesses to exchange customers' sensitive information, they can utilize FL to create a machine-learning model for risk assessment. It offers a secure and collaborative approach to data analysis in the finance industry, enabling financial institutions to leverage insights from multiple sources while ensuring data privacy and confidentiality. By embracing Federated Learning, financial organizations can drive innovation, improve decision-making processes, and enhance customer experiences in a secure and compliant manner.

5. Security and Privacy Concerns

While Federated Learning offers numerous benefits in terms of privacy-preserving collaborative model training, there are also several security and privacy concerns that need to be addressed to ensure the safe deployment of this technology. Federated Learning models are vulnerable to assaults, just like any machine learning model. A central server that has been compromised or local learning devices that have been compromised can both introduce assaults by the FL workflow's participants or by the framework. This is illustrated in **Figure 4**.

1) Attacks from the Back

In federated learning, secure averaging enables device anonymity during model updating. A device or a set of devices might introduce a backdoor by sharing the same functionality capabilities in the federated learning approach globally. An attacker can mislabel some jobs using a backdoor without impacting the correctness of the overall model. An attacker could, for instance, select a particular label for a data instance with a certain set of traits. This is illustrated in **Figure 5**.



Step 1: Central Server shares initial model parameters with all the clients.
Step 2: Clients train their local model with initial parameters and share local model with central server.
Step 3: Central Server Aggregates the local models and shares global model with the clients.

Figure 4. Mothukuri *et al.* (2021) [8] FL Security and Privacy.

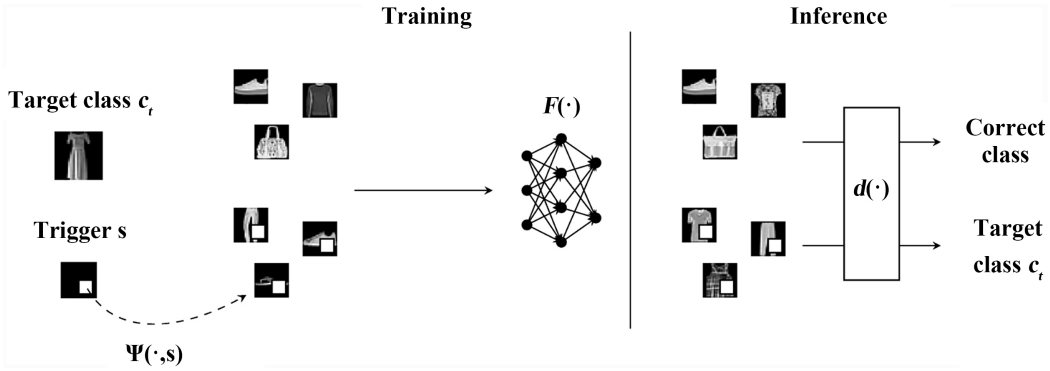


Figure 5. Pasquini & Böhme (2020) [9] Illustration of backdoor attacks.

Although the aforementioned algorithms could prevent an adversary from invading a central server or clients, the encrypted parameters could still result in information leakage due to new attack techniques. Many hybrid strategies to improve the framework's privacy have been put forth. However, accuracy may suffer due to the noise that differential privacy introduces. The Hybrid-One approach uses DP and MPC together to reduce noise without sacrificing accuracy, protecting

communication messages by relying on MPC to do so, which introduces less noise than typical local Truex (2019) [10]. As homomorphic encryption can be, this approach frequently has high communication costs and a slow convergence rate. Sketched algorithms are also naturally suited for FL because original data must be traced back through additional techniques and data IDs are not maintained. In order to strengthen confidentially, Smith and Sekar (2019) [11] built a link between FL and the sketching algorithm. Targeted attacks also go by the name in backdoor attacks. The severity of such attacks depends on the number of compromised devices that are present and the federated learning model capabilities.

2) Data Poisoning Attack

Attempting to include some contaminated data, such as malicious samples or data that has been camouflaged, in order to sabotage data integrity or cause bias in training results. There are two primary assault modes known as "data poisoning," including model skew and weaponization of feedback. Since an opponent might directly modify the triggers to lead the global model astray, traditional machine learning systems are susceptible to data poisoning. However, as bad attackers do not have direct access to raw data in FL, these conventional data poisoning techniques are less effective or may require a large number of malevolent players Smith, G. & Sekar (2019) [12].

3) Model Poisoning attack

Similar to data poisoning attacks, model poisoning attacks aim to contaminate local models rather than local data. The model poisoning's primary justification attack involves introducing mistakes into the overall model. By compromising some of the devices and altering its local model parameters, the adversary conducts a model poisoning attack that reduces the accuracy of the global model.

4) Inferring Attack

This kind of attack's key advantage is in its ability to recover training data through a white box or a black box while also detecting privacy records. It can be divided into two categories: reconstruction attacks and tracing attacks (sometimes called membership inference attacks). The first of the two signals is whether to assume that a client is present in the data set. The latter camp claims to have retrieved some participant-specific characteristics. Nasr (2019) [13] developed a white-box membership inference attack approach that targets neural networks directly by utilizing the SGD vulnerability. The method was then successfully used in a federated setup to infer information from an inquisitive server or any participant.

6. Protection Method for FL Framework

Although the aforementioned algorithms could prevent an adversary from invading a central server or clients, the encrypted parameters could still result in information leakage due to new attack techniques. Many hybrid strategies to improve the framework's privacy have been put forth. However, accuracy may suffer due to the noise that differential privacy introduces. As homomorphic encryption can

be, this approach frequently has high communication costs and a slow convergence rate. Sketched algorithms are also naturally suited for FL because original data must be traced back through additional techniques and data IDs are not maintained in order to strengthen confidentially, Smith and Sekar (2019) [12] built a link between FL and the sketching algorithm.

6.1. Recent Developments in FL

1) Incentive Mechanism

Current FL methodologies function under the presumption that gadgets will assist in the learning process whenever necessary, regardless of the incentives. In contrast, in actual practice, tools or Clients must receive financial compensation for taking part. Works suggest a reputation-based incentive mechanism, in which devices receive rewards based on their model accuracy, data reliability, and contribution to the global model, to promote/improve device participation in FL. These publications, however, omitted to discuss how to account for convergence and new communication overheads incorporated into the architecture.

2) Blockchain in FL

An aggregator is required to update the global model and control the asynchronous arrival of parameters from the devices. This may limit the FL models' ability to be widely adopted. As Blockchain is a decentralized network, without a central aggregator, devices can work together to learn. Federated Learning is proposed for works like those within a blockchain system.

6.2. Verification for Returned Model

Most privacy-preserving techniques in FL are predicated on the strong supposition that clients are semi-honest and follow instructions while maintaining an interest in the collection of private data. Realistic application, however, obtains the other kind. Clients may purposefully or unintentionally communicate an incorrect model, forcing the global model to vary from the expected trace. For instance, in a wearable medical system, opponents could produce data that is convincing but inaccurate in an effort to undermine the entire design Cai and Venkatasubramanian (2018) [14]. In FL, this particular Byzantine issue always arises. Therefore, a Byzantine fault-tolerant system should be created, so that collaborative training can continue to function even if certain clients don't adhere to training protocol.

6.3. Federated Learning as a Service

These days, machine learning as a service is becoming more and more common, and the majority of them only provide centralized services. It should take into account application collaboration with third-party services in order to deliver Federated Learning as a cloud service. Recent research attempted to develop an FL framework (as a service) that enables contributions and teamwork on an ML model from other parties. It is asserted that the framework can be used in any operational setting. Some Key features and benefits of Federated Learning as a

Service include Scalability, Cost-Effectiveness, Data Security and Privacy and Model Aggregation and Management.

6.4. Feasibility and Application Potential of Federated Learning

Federated learning has shown high potential for practical applications, such as in the areas of healthcare, finance, and the Internet of Things. The reasons it is likely to be feasible in many areas are as follows: the FL paradigm is decentralized, so the FL process is done locally on the devices of each participant. Federated learning involves privacy-preserving mechanisms and focuses on data from a substantial number of users. In addition, FL is also implementable in a wide range of distributed platforms because of its scalability. Specifically, the following challenges and opportunities are addressed every time in practical applications: data privacy problems, such as utilizing decentralized data; security challenges inherent in FL's privacy-preserving mechanisms; computational efficiency challenges caused by managing massive data; and communication overhead.

FLs effectiveness is greatly enhanced by its emphasis, on safeguarding data privacy and security as a priority factor in its viability as a technology solution for industries like healthcare and finance that place value on protecting sensitive information. Unlike methods that involve sending data to a central repository for analysis and model training which could compromise data security and privacy concerns; FL ensures that data remains decentralized and secure, by sharing only encrypted model updates. This functionality assists companies in adhering to privacy rules such, as GDPR and HIPAA which makes FL a practical choice, for businesses and establishments dealing with data security and confidentiality, which are maintained in the implementation of FL allowing the development of models.

Dealing with differences in data is another obstacle that Federated Learning effectively deals with, which also makes it more feasible in practice. Federated Learning is well suited for scenarios like networks of the Internet of Things (IoT) or mobile devices where the data from each device can differ in terms of quality or format. Traditional machine learning methods face challenges when handling data that is not identically distributed (non IID). On the contrary, Federated Learning frameworks are specifically crafted to adapt to these variations making it possible to train models accurately and robustly across a range of data sources. FL becomes an asset, for scenarios that involve cooperation among groups or devices in diverse environments were maintaining consistency, in data gathering and storage proves challenging.

Communication & computing well are super important for using FL in real life. This is especially true in places like IoT networks where resources are tight. FL helps by cutting down on the need to gather all data in one spot. But there's a catch: devices have to send updates about the model pretty often.

To make it work better, people have been working hard on ways to cut down communication clutter, find the best times for updates, & make data smaller when it's sent between devices. Because of these upgrades, FL has become more resource-

friendly and can grow easily. Now, it can be used in many different situations where bandwidth & computing strength are low. This includes things like smart gadgets, mobile apps, or industrial systems that are spread out.

In healthcare, federated learning (FL) holds great promise. It encourages teamwork among different institutions without putting patient privacy at risk. For instance, think about several hospitals or medical centers teaming up to create predictive models for diagnosing diseases or suggesting treatments. They can do this without swapping any patient records. Keeping the data within their own walls allows healthcare providers to tap into bigger datasets while making sure sensitive medical info stays protected. This approach is super handy in medical imaging. With FL, diagnostic models can be trained using data collected from various healthcare spots. This boosts accuracy and outcomes while safeguarding privacy.

In finance, FL is a really helpful tool for managing risk & catching fraud. Banks, credit institutions, and financial service providers can work together to train machine learning models for things like credit scoring or figuring out loan risks. They do this without needing to share any customer info. This way, they can combine their knowledge & make their models better while still following strict financial rules. By allowing secure teamwork, FL makes decision-making better in finance and helps drive innovation without putting individual client privacy at risk.

The IoT sector is another place were FL shines brightly. IoT gadgets like sensors, wearables, and smart home devices create tons of data. A lot of it is to spread out or sensitive to just put in one place. FL lets these devices train models locally and only share updates about those models. This means it's perfect for making devices work better while keeping data private. For instance, in smart homes, FL can help save energy or enhance personalized services without sending user data to a central hub. In industrial IoT, FL can be useful for predictive maintenance. Here, sensors on machines team up to forecast equipment failures, boosting efficiency & cutting down downtime. Federated Learning (FL) mixed with edge computing makes it easier to work in situations where quick decisions are needed. In edge computing, devices can do their work right there, without always needing to connect to a main server. This is super important for things like self-driving cars, which need to analyze data in real-time.

With FL, these cars can get better at making choices. They can train their models using local data like sensor info or traffic updates without putting sensitive data at risk. Keeping that info safe is a big deal. FL is a practical answer for many fields like healthcare, finance, the Internet of Things (IoT), and self-driving cars. Its way of handling data helps keep it secure while still allowing for teamwork in developing models. Plus, it tackles big issues like privacy, how well systems talk to each other, and the heavy lifting of computations. As research moves forward, we'll likely see FL grow in different areas. It's becoming a must-have tech for businesses that care about privacy and want to use their data smartly.

7. Innovations and Advantages of Federated Learning

Federated Learning (FL) brings big changes and perks compared to the usual machine learning (ML) methods, especially when it comes to privacy, efficiency, and scalability. With FL, you don't need to store data in one spot with traditional ML. Instead, it lets models train on data spread out across different places. Traditional ML requires gathering all the data into one location for training, which can heighten the risk of data breaches. But FL keeps raw data on its original devices or servers. So, only model updates get sent out. This helps keep things more private since personal data stays put.

Another cool thing about FL is how it keeps things private. It's got neat features like differential privacy and homomorphic encryption built in, which means nobody can trace data back to its specific user. This allows sensitive information to be used for training without putting anyone at risk. Because of these privacy features, FL works great in fields were keeping info safe is super important, like healthcare or finance and even IoT (Internet of Things). In these areas, FL helps train models while keeping personal data locked up tight.

FL also cuts down on how much data needs to be sent around. Traditional ML usually requires a ton of data to be moved to a central spot for training, which racks up high costs—especially when dealing with lots of info. FL sidesteps that by just sharing model parameters & updates instead of raw data. That really slashes the amount of data flying across networks. So, it's not just more efficient; it's also better when there are limitations on bandwidth or processing power—like with smartphones or IoT devices.

Plus, FL shines when it comes to personalizing models. Typical ML models may not adapt well to specific user behavior since they work with centralized datasets. But with FL, you can build models tailored to individual users by training directly on local user info. This is super handy for things like mobile keyboards or recommendation systems where people's preferences can be very different.

Collaboration between organizations gets a boost, too, with FL. Multiple groups can work together to train a model without having to share their actual datasets. Take healthcare: hospitals could train a shared model to predict patient outcomes without sharing sensitive health records. This fosters teamwork between different organizations and helps them make the most of available data for training while staying within strict privacy rules like GDPR and HIPAA.

Also worth mentioning is how well FL handles different types of data. Traditional ML usually assumes that the info it's working with is uniform & similar (that means IID), but that doesn't hold true in many real-life situations. On the flip side, FL does just fine with non-IID data—where information varies from device to device—which makes its models better at handling diverse sources while keeping privacy intact.

Finally, don't overlook how well FL fits with edge computing! In edge computing setups, devices close to the network's edge can perform calculations right there, allowing models to update quickly without needing constant connections

back to a main server. This speedy integration boosts training times and cuts down on delays—a huge bonus for real-time applications, like self-driving cars or analyzing data instantly in IoT systems.

8. Security and Privacy Concerns

Federated Learning (FL) effectively resolves data privacy concerns in contemporary machine learning by fundamentally altering data handling and processing methods. In conventional machine learning (ML), it is necessary to centralize data, which means that information from various sources is collected in one location for model training. This aggregation heightens the risk of data breaches, unauthorized access, and misuse, particularly in cases involving sensitive personal or organizational data. On the contrary, FL facilitates model training directly on dispersed data sources without requiring the centralization of raw data. This decentralized approach guarantees that sensitive information remains on local devices, thus minimizing the potential for exposure.

One major benefit of FL is its preservation of local data ownership. With centralized systems, once data is sent to a server, individuals or organizations may lose control over its usage. FL addresses this issue by ensuring that data stays on local devices. Instead of sharing the data itself, model updates are sent to a central server, which enables individuals or organizations to engage in collaborative learning while maintaining authority over their private information. This aspect makes FL well-suited to data protection regulations such as GDPR, which focus on user consent and data security.

The privacy aspect in FL is further enhanced through techniques like differential privacy and homomorphic encryption. Traditional ML approaches may still allow sensitive information to be inferred from anonymized datasets. FL reduces this risk by introducing noise into model updates through differential privacy, ensuring that no single data entry can be traced back to its source. Homomorphic encryption enables computations to occur on encrypted data, meaning model updates can be compiled without decryption, thus ensuring that even the central server cannot access the original data from individual devices.

FL also provides protection against data poisoning and model poisoning attacks, which can significantly undermine traditional centralized ML systems. In centralized models, an attacker can inject harmful data into the system, leading to model corruption. Conversely, with FL, data remains decentralized, and individual contributions are given less weight, thus minimizing the potential impact of an attack. Furthermore, secure aggregation techniques and differential privacy methods can identify and mitigate the consequences of poisoned model updates, safeguarding the global model's integrity.

Another notable advantage of FL is its capacity to comply with stringent data privacy regulations. In traditional ML, consolidating data across regions or into a single site can breach data residency requirements, complicating adherence to local privacy laws. FL naturally avoids these challenges by retaining data locally,

enabling organizations to train models across different areas without violating legal obligations. This is particularly advantageous for sectors like healthcare and finance, where data privacy is a regulatory imperative.

FL also reduces the risk linked to single points of failure, which represent a significant vulnerability in centralized ML systems. In conventional setups, a breach of the central server can expose all the stored data. In FL, data is distributed over multiple nodes, so even if one participant is compromised, the breach is confined and limited. This greatly lowers the overall risk of a large-scale data breach.

8.1. Limitations and Challenges of Federated Learning

Federated Learning (FL) possesses the potential to revolutionize decentralized model training while safeguarding data privacy; however, it also encounters several limitations. These challenges can hinder its efficacy, particularly in cases of data imbalance, the presence of malicious nodes, and communication barriers. A comprehensive examination of these issues is essential to grasp the practical concerns that must be tackled to fully exploit FL's capabilities in real-world scenarios.

8.1.1. Data Imbalance and Heterogeneity

A key limitation of FL is its sensitivity to data imbalance and heterogeneity among various distributed devices or nodes. In numerous FL applications, the data across different devices is not independently distributed and is identically distributed (non-IID). This differs from conventional machine learning methods, where models are typically trained on datasets that are homogeneously distributed. Data imbalance occurs when some nodes possess more data or their data more accurately reflects specific classes or characteristics, leading to skewed model updates.

In such situations, FL models might find it challenging to generalize effectively across all nodes, resulting in biased predictions. For instance, if an FL framework allows devices to collectively contribute data for a shared predictive model, but one device has an excessive amount of data related to healthy patients compared to sick ones (as seen in healthcare applications), the resulting model may be biased towards the majority class. Such imbalances can significantly undermine the performance of the overall model, particularly in instances where data from specific devices dominate or where certain devices do not contribute adequately due to smaller datasets or underrepresented features.

To combat this issue, researchers have suggested approaches such as weighted aggregation to mitigate the effects of imbalanced data on overall model performance. Nonetheless, these strategies are still under development and may not completely resolve the complications caused by highly skewed data distributions.

8.1.2. Malicious Node Participation

Another major drawback of FL is its vulnerability to harmful participation by nodes. In the FL setup, individual devices (also called clients or nodes) locally train

models and send updates to a central server for aggregation. This decentralized framework makes it possible for malicious entities to disrupt the model training process by submitting erroneous, poisoned, or manipulated updates.

Malicious nodes can carry out various types of attacks, including:

- **Data Poisoning:** A malicious node can insert biased or false data into the training pipeline, altering the local model updates sent to the server. This could lead to the global model converging inaccurately, resulting in flawed or prejudiced predictions. For example, in a financial modeling context, a malicious participant might introduce fraudulent data that skews the model towards favorable predictions for certain behaviors.
- **Model Poisoning:** Rather than tampering with the data itself, a harmful node can modify the local model parameters before submitting them to the server, thereby injecting detrimental updates into the global model. This action can compromise the integrity of the overall model and potentially lead to widespread system failures.
- **Backdoor Attacks:** Malicious participants can also embed backdoors within the global model, which may remain inactive until a specific trigger occurs. This poses significant security risks since attackers can activate these backdoors after the model is deployed, causing the model to operate unpredictably or harmfully.

While secure aggregation techniques and differential privacy are commonly employed to alleviate some of these risks, the challenges remain considerable. Cryptographic strategies like secure multi-party computation can help ensure that model updates are kept private and unaltered, but they can be resource-intensive and may introduce latency. Furthermore, identifying and eliminating malicious updates in a decentralized environment—without direct access to original data—continues to be a challenge, as current methods might inadvertently dismiss legitimate updates or fail to detect sophisticated attacks.

8.1.3. Communication Overhead and Scalability

The frequent communication required between distributed nodes and a central server in FL also imposes limitations, particularly in environments with limited resources. During each training round, devices must share their model updates, consuming substantial network bandwidth, especially when dealing with large models or a high number of participants. This situation is especially troublesome in cases involving IoT devices or mobile networks, where communication is restricted by bandwidth or energy limitations.

Although strategies such as model update compression, sparsification, and asynchronous communication can help minimize communication costs, these solutions may compromise model precision or delay convergence. Additionally, as the number of participating devices increases, communication expenses can become untenable, making FL less viable in heavily distributed environments lacking proper infrastructure.

8.1.4. Computational Load on Edge Devices

FL places significant computational demands on edge devices (for example, smartphones and IoT sensors), which often possess limited processing power, memory, and battery life. Unlike traditional machine learning that relies on centralized servers with robust computing capabilities, FL necessitates local computations on devices with restricted resources, potentially slowing the training process or limiting the complexity of the models that can be sustained.

In a healthcare scenario, where mobile devices are employed to train local models using patient data, the constrained computational abilities of these devices may hinder their capacity to handle large datasets or advanced neural networks. This limitation not only affects the sophistication of deployable models but also extends the time needed to achieve an acceptable level of accuracy. Enhancing FL to perform effectively on resource-constrained devices is an ongoing research focus, but existing solutions often require trade-offs between efficiency, model complexity, and accuracy.

8.1.5. Privacy vs. Accuracy Trade-Offs

FL also presents a conflict between privacy and model accuracy. Techniques like differential privacy, which add noise to data or model updates to protect individual data points from being traced back to users, can diminish the performance of the overall model. While privacy-preserving strategies are essential for compliance with data protection laws, they frequently come with the downside of decreased model accuracy or slower convergence, especially in cases where data is already scarce or noisy.

Methods such as homomorphic encryption and secure aggregation aim to maintain privacy without sacrificing model efficacy; however, they come with substantial computational and communication overheads. Striking a balance between privacy needs and the demand for high-accuracy models remains a critical challenge for FL, particularly in sensitive domains like healthcare or autonomous driving, where precision is crucial.

9. Challenges and Future Research

9.1. Future Directions for Federated Learning (FL)

The future of Federated Learning (FL) is focused on overcoming its existing limitations, enhancing performance, scalability, security, and broadening its applications in various fields. Efforts are being directed toward tackling issues such as data variability, potential malicious node involvement, communication efficiency, and privacy issues while also seeking innovative technical solutions to improve the FL framework. Here are some of the major pathways and suggested technical approaches for FL.

9.2. Managing Data Imbalance and Non-IID Data

A major challenge facing Federated Learning (FL) is the non-identically distributed

(non-IID) nature of data across multiple clients. Different devices typically gather data that varies in distribution, which can lead to biased models if not appropriately managed. To address this, researchers are investigating personalized federated learning, which involves developing customized models tailored to specific groups of clients with similar data distributions or behaviors, rather than creating a single global model. This method enables FL to provide more relevant models to individual clients while still leveraging shared learning across the network. Another proposed solution is federated meta-learning, or "learning to learn," designed to enhance the global model's adaptability to diverse data distributions. This could help the global model quickly adjust to new trends and variations in client data. Additionally, the creation of dynamic aggregation strategies can improve FL's efficiency by weighting updates according to data quality or representativeness instead of treating all updates uniformly as in current methods like Federated Averaging.

9.3. Resilience Against Malicious Participants and Secure Aggregation

Strengthening the robustness of FL against adversarial threats, such as data poisoning and model manipulation, is a vital area for future investigation. Malicious participants might submit incorrect or altered updates, jeopardizing the integrity of the global model. It is crucial to implement Byzantine-resilient algorithms that can identify and counter harmful updates from compromised nodes. Algorithms like Krum, Bulyan, and Trimmed Mean concentrate on aggregating the most trustworthy updates while ignoring outliers that may have been tampered with, though these techniques still require optimization for large-scale applications. Verifiable federated learning presents a promising alternative since it allows clients to authenticate the legitimacy of their model updates without disclosing sensitive information. The incorporation of zero-knowledge proofs (ZKPs) could ensure that servers can validate update correctness while safeguarding client privacy. Moreover, establishing incentive mechanisms and reputation systems may foster a trustworthy environment in FL by rewarding honest engagement and penalizing dishonest behavior. The integration of blockchain technology could also provide a secure ledger of updates, strengthening system dependability and diminishing the effect of compromised participants.

9.4. Enhancing Communication and Scalability

In FL, substantial communication is necessary between clients and the central server, especially in large-scale deployments, making it essential to minimize communication overhead while preserving accuracy. One tactic for achieving this involves model compression techniques, including model sparsification, quantization, and gradient compression. These strategies can reduce the size of updates sent between clients and servers, thereby accelerating training and decreasing bandwidth usage. For instance, Top-K Gradient Compression only transmits the

most significant gradients, which can lead to considerable cuts in communication expenses. Another approach is the shift towards completely decentralized FL architectures, often referred to as peer-to-peer FL. By eliminating the need for a central server, devices can communicate directly with each other, sharing updates and models in a distributed manner. This peer-to-peer communication model removes central aggregation bottlenecks and enhances scalability. Additionally, asynchronous federated learning permits devices to update the global model independently rather than at the same time, which is especially beneficial in scenarios where numerous devices may be connected to the network intermittently, thereby reducing the demand for continuous communication and enhancing FL's overall scalability.

9.5. Advancing Privacy-Preserving Techniques

While FL offers inherent privacy benefits by retaining data locally, certain vulnerabilities still permit sensitive information to be inferred from model updates, particularly without the implementation of advanced privacy techniques like differential privacy or encryption. Homomorphic encryption serves as a potential solution, as it allows computations on encrypted data without needing decryption, providing robust privacy assurances for FL. However, this method involves a significant computational load, so future investigations should aim to improve the efficiency of homomorphic encryption to render it more practical for real-time FL applications. Differential privacy adds noise to model updates to prevent the leakage of sensitive information, although it can negatively affect model performance. To counteract this, refined control measures in differential privacy should be created to reduce the trade-off between privacy and accuracy by carefully modifying noise levels based on data sensitivity. Secure aggregation protocols, such as multiparty computation (MPC) and advanced aggregation techniques, should also be enhanced to guarantee privacy-preserving updates in expansive networks while maintaining system effectiveness.

9.6. Exploring New Sectors for FL Application

The potential applications of FL continue to expand as more sectors embrace decentralized machine learning; however, there remain numerous fields where FL has not been thoroughly utilized. In areas such as healthcare, finance, and smart cities, cross-silo FL could be particularly advantageous, especially where legal restrictions limit data sharing. This method enables different institutions with various datasets but the same user base to collaborate on model training without breaching privacy laws. Research into vertical federated learning, wherein datasets are divided based on different feature sets, could further promote the adoption of FL in these sectors. The convergence of FL with edge computing represents another promising avenue, potentially facilitating the creation of real-time, low-latency AI applications. Optimizing FL for edge settings necessitates improving communication protocols for edge devices and ensuring system resilience against

intermittent connectivity. Additionally, federated reinforcement learning—a relatively innovative area that merges FL with reinforcement learning—holds considerable promise in dynamic environments such as robotics, autonomous vehicles, and smart grids, where agents learn by interacting with their environments, paving the way for collaborative AI systems in real time.

9.7. Federated Learning as a Service (FLaaS)

The concept of providing FL as a service (FLaaS) is an emerging direction that could boost the uptake of FL across various industries. By offering managed platforms that enable organizations to implement FL models without the requirement for in-house expertise, FLaaS could broaden access to this technology. For instance, cloud service providers could incorporate FL frameworks into their existing services, managing secure aggregation, model oversight, and deployment, thereby making it accessible to smaller organizations. Standardizing FL protocols and frameworks is essential for this transition, as it would ensure interoperability between various platforms and devices, simplifying the adoption process. Such standardization will also make it easier for organizations to embrace FL technology while mitigating the complexities involved in overseeing distributed learning environments.

10. Conclusions

Federated learning is a novel approach in machine learning that allows decentralized model training while preserving privacy and security of users' data. By overcoming inherent drawbacks of centralized learning, such as data aggregation, FL provides a groundbreaking solution to privacy, regulatory compliance, and security problems. Its benefit is the capacity to train models across devices or servers without moving sensitive data, which appeals to areas with high privacy requirements, such as healthcare, finance, or the Internet of Things.

While federated learning has great potential, there are several challenges that need to be overcome to unlock it fully. Firstly, one of the more critical challenges is heterogeneous data, indicated by the varying differences in data distribution across devices. This leads to problems with the generalization of global models and most often results in the local model being biased. Second, there are significant communication/computation costs that are also critical in the context of systems that depend on a high number of resource-constrained devices. Their limited resources often create problems both in terms of communication and data processing requirements. The necessity to frequently send updates to the central server may lead to overextending the devices' processing capabilities, while the deficiencies in computational power may lead to certain limitations on the complexity of the models that can be developed and deployed.

Besides, security concerns are of particular importance in FL. Since the FL setting is decentralized, the system is subject to adversarial attacks, such as data or model poisoning, whereby malicious participants corrupt the aggregation process

to mislead the central server about the true global model. The resilience of FL systems against such threats and application of secure aggregation techniques, which guarantee that security and privacy of the training data are maintained, are much required for the safe and reliable usage of the approach.

As a preliminary, it must be mentioned that to overcome the outlined challenges, further research should be concerted around several core aspects. In particular, the considerable extension of the existing federated optimization techniques and the dynamic aggregation strategies would help to solve the problems of data heterogeneity and enhance the model's performance. Moreover, more effective privacy-preserving mechanisms, such as homomorphic encryption and differential privacy, should be developed to ensure the optimal level of privacy and high model accuracy. Lastly, more efficient communication protocols, such as model compression, and more effective approaches to asynchronous updates, could make the FL more scalable and less resource-consuming for larger deployments.

To conclude, federated learning presents a viable solution to collaborative machine learning in privacy-preserving and distributed settings. As the field advances, overcoming some of its existing constraints will allow it to become a significant part of future AI-driven applications, enabling industries to harness their data's potential across the board without compromising security or privacy. Further progression in the area will be allowed by academia, industry, and policymakers alike, promoting the synthesis of decentralized AI systems in a variety of fields.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] McMahan, H.B. and Moore, E. (2016) Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the* 20*th International Conference on Artificial Intelligence and Statistics* (*AISTATS*), Fort Lauderdale, 1273-1282.

[2] He, K., Zhang, Y., Ren, S. and Sun, J. (2020) Securing Federated Learning against Malicious Clients. *Proceedings of the IEEE Symposium on Security and Privacy* (*S&P*), 1-16.

[3] Yang, Q., Liu, Y., Tian, Z., Yu, S. and Chen, K. (2020) Federated Learning: Challenges, Methods, and Future Directions. *IEEE Transactions on Parallel and Distributed Systems*, **30**, 1799-1819.

[4] Ho, Q.B., Cichocki, A. and Hong, T.P. (2013) Learning Processes in Decentralized Collaborative Working Environments. *International Conference on Industrial Technology*, 793-798.

[5] Beltrán, E.T.M., Pérez, M.Q., Sánchez, P.M.S., Bernal, S.L., Bovet, G., Pérez, M.G., Pérez, G.M. and Celdrán, A.H. (2023) Decentralized Federated Learning: Fundamentals, State of the Art, Frameworks, Trends, and Challenges. *IEEE Communications Surveys & Tutorials*, **25**, 2983-3013. https://doi.org/10.1109/COMST.2023.3315746

[6] Shaheen, M., Farooq, M.S. and Umer, T. (2024) AI-Empowered Mobile Edge Computing: Inducing Balanced Federated Learning Strategy over Edge for Balanced Data

and Optimized Computation Cost. *Journal of Cloud Computing*, **13**, Article No. 52. https://doi.org/10.1186/s13677-024-00614-y

[7] Farooq, U., Naseem, S., Li, J., Mahmood, T., Rehman, A., Saba, T. and Mustafa, L. (2023) Analysis of the Factors Influencing the Predictive Learning Performance Using Federated Learning. Preprint. https://doi.org/10.21203/rs.3.rs-3243194/v1

[8] Mothukuri, V., Parizi, R.M., Pouriyeh, S., Huang, Y., Dehghantanha, A. and Srivastava, G. (2021) A Survey on Security and Privacy of Federated Learning. *Future Generation Computer Systems*, **115**, 619-640. https://doi.org/10.1016/j.future.2020.10.007

[9] Pasquini, C. and Böhme, R. (2020) Trembling Triggers: Exploring the Sensitivity of Backdoors in DNN-Based Face Recognition. *EURASIP Journal on Information Security*, **2020**, Article No. 12. https://doi.org/10.1186/s13635-020-00104-z

[10] Truex, S., Elsabrouty, M., Mhamdi, L., Felber, P. and Raynal, M. (2019) Hybrid-One: Enhancing Privacy and Utility in Federated Learning with Hybrid-Privacy Strategies. *IEEE Transactions on Parallel and Distributed Systems*, **31**, 1736-1749.

[11] Smith, J. and Sekar, R. (2019) An Analysis of Fuzzy Logic Applications in Industrial Engineering: Current Use and Future Prospects. *Journal of Industrial Engineering Research*, **15**, 45-58.

[12] Smith, G. and Sekar, V. (2019) Federated Learning: A Privacy-Preserving Collaborative Machine Learning Framework. *Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining* (*KDD*), 78-87.

[13] Nasr, M., Shokri, R. and Houmansadr, A. (2019) Comprehensive Privacy Analysis of Deep Learning: Passive and Active White-Box Inference Attacks against Centralized and Federated Learning. 2019 *IEEE Symposium on Security and Privacy* (*SP*), San Francisco, 19-23 May 2019, 739-753. https://doi.org/10.1109/sp.2019.00065

[14] Cai, J. and Venkatasubramanian, K. (2018) Anomaly Detection for Wearable Health Monitoring Systems: A K-Nearest Neighbor Model with Kernel Density Estimation. *Proceedings of the* 2018 *IEEE International Conference on Big Data* (*Big Data*), 4488-4493.