A survey on federated learning: evolution, applications and challenges

Lingyu Shi

Wuhan Britain China School, Wuhan, Hubei, China, 430000

www.shilingyu2023@163.com

Abstract. Federated learning, a machine learning technique that enables collaborative model training on decentralized data, has gained significant attention in recent years due to its potential to address privacy concerns. This paper explores the evolution, applications, and challenges of federated learning. The research topic focuses on providing a comprehensive understanding of federated learning, its advantages, and limitations. The purpose of the study is to highlight the importance of federated learning in preserving data privacy and enabling collaborative model training. The study conducted a literature review by systematically analyzing relevant papers from peer-reviewed journals, conference proceedings, and reputable sources. The results reveal that federated learning offers a promising solution for collaborative machine learning while addressing concerns related to data privacy and security. The study emphasizes the need for further research in optimizing communication protocols, scalability, and privacy-preserving techniques. Overall, this paper contributes to the understanding of federated learning and its potential for secure and efficient decentralized learning paradigms.

Keywords: federated learning, decentralized data, privacy preservation, collaborative model training, model aggregation.

1. Introduction

Federated learning (FL) is a machine learning (ML) technique that enables training models on decentralized data while preserving data privacy and security. This paper provides a comprehensive overview of FL, including its history, current state, and future directions.

The paper highlights the advantages of FL, its applications in smartphones, healthcare, and the Internet of Things (LoT), as well as its challenges and limitations. These challenges include non-IID data, systems heterogeneity, and privacy concerns [1]. The research methodology used in this paper is a literature review that summarizes and analyzes existing research on FL. This author conducted a systematic search of relevant papers published in peer-reviewed journals, conference proceedings, and other reputable sources. This author then reviewed and synthesized the findings of these papers to provide a comprehensive overview of FL.

The motivation for this study is to provide researchers with a comprehensive understanding of FL and its potential impact on various industries. This paper serves as a starting point for future research and a reference for identifying key trends, challenges, and opportunities in the field of FL. Overall, this paper provides valuable insights into the past, present, and future of FL and its potential impact on the field of ML.

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

2. Evolution of federated learning

2.1. Centralized learning

Centralized learning refers to the traditional approach of training ML models, where data is collected from various sources and then transferred to a central server for analysis and model training. This approach has been the dominant method for many years and has led to significant advancements in the field of ML.

The centralized learning model was first introduced in the 1950s, and it was initially used in simple tasks such as character recognition. Over time, with the increase in computational power, the centralized learning model was used for more complex tasks.

Despite its successes, centralized learning has several limitations. One of the major shortcomings is its dependency on data centralization, which raises concerns about data privacy, ownership, and control. Centralized approaches require data to be transferred to a central server, which can be time-consuming and costly. Additionally, centralized learning can suffer from performance issues due to high network traffic and latency, which can negatively impact the model's accuracy and speed [2].

In response to these limitations, researchers have been exploring alternative approaches to ML, such as on-site ML.

2.2. Distributed on-site learning

Distributed on-site learning is becoming more popular due to the risks involved in sending private data to a centralized entity. This approach involves deploying a pre-trained or generic ML model to each device, which can then personalize it by training on its local data, making predictions, or running inference computations. This approach offers privacy advantages since data does not leave its host device. On-device intelligence has already been used in many applications, including medical applications, skin cancer detection, smart classrooms, and neural network-assisted services. However, this approach limits the generation of local models to each user's experience without benefiting from the data of their peers [3]. To address this, federated learning (FL) has been developed, which allows users' computations to be federated while still maintaining privacy.

2.3. Federated learning

2.3.1. Key concepts. In 2016, FL was introduced by Google researchers as a way to move the training task to the device itself while also federating local models and learning. This approach aims to provide a privacy-preserving ML framework and has gained momentum in both academia and industry. Unlike other approaches that involve sending private data to the server, performing ML tasks on devices without peer's data, or precluding direct access to raw data, FL enables on-device ML training while federating locally trained ML models [4]. This approach minimizes data communication overhead by keeping raw data on the devices and aggregating locally computed model updates, which ensures data privacy.

2.3.2. Basic steps. The basic steps of FL involve a series of key processes to enable collaborative model training on decentralized data while preserving privacy. These steps can be summarized as follows:

Initialization: The process begins by initializing a global model either with pre-trained weights or random parameters.

Client Selection: A subset of clients or devices is selected to participate in the training process. This selection can be based on various criteria such as device availability, network conditions, or representative data.

Model Distribution: The global model is distributed to the selected clients. Each client receives a copy of the model and performs local training using its own data.

Local Training: Each client independently trains the model using its local data. This training can involve multiple iterations and updates to improve the model’s performance.

Model Aggregation: After local training, the clients send their model updates, such as gradients or weights, to a central server or aggregator.

Global Model Update: The central server aggregates the received model updates from the clients, typically by averaging or weighted averaging, to obtain an updated global model.

Iteration: Steps from Client Selection to Global Model Update are repeated for multiple iterations to further refine the global model collaboratively.

Model Deployment: Once the desired level of convergence is achieved, the final global model can be deployed for inference on new data.

3. Applications of federated learning

3.1. Smart phones

FL has found various applications in the context of smartphones, leveraging the power of distributed on-device training while preserving user privacy. One prominent example is in the field of next-word prediction, which aims to enhance the typing experience on smartphones. By utilizing FL, personalized language models can be trained directly on users' devices, considering their individual writing habits, frequently used words, and contextual information [4]. This enables accurate and context-aware next-word suggestions, improving typing efficiency and user satisfaction.

Another application of FL on smartphones is face detection. With the increasing demand for facial recognition capabilities in mobile applications, FL allows for the development of robust and accurate face detection models. By training these models directly on users' devices using their locally stored facial data, privacy concerns associated with uploading sensitive facial information to a central server can be mitigated [1]. The aggregated knowledge from these decentralized models can then be utilized to enhance face detection algorithms, enabling secure and efficient facial recognition features on smartphones.

3.2. Organizations: smart healthcare

FL has significant implications for smart healthcare, revolutionizing the way healthcare services are delivered and improving patient outcomes. One notable application is the use of FL to learn from health data and facilitate healthcare services, particularly in intelligent imaging for disease detection. By leveraging decentralized data from various healthcare providers and devices, FL enables the development of robust and accurate imaging models while preserving patient privacy. This allows for enhanced diagnostic capabilities, early disease detection, and improved treatment planning [5].

In addition, community-based FL algorithms have been employed in healthcare to predict mortality and hospital stay time. By federating data from multiple healthcare institutions and leveraging ML techniques, predictive models can be trained to identify risk factors and predict patient outcomes. This approach enables personalized and proactive care, facilitating timely interventions and resource allocation to improve patient prognosis and optimize healthcare delivery [3].

To provide high-performance models in smart healthcare, clustering techniques are employed to group homogeneous patients with similar characteristics. Each cluster then creates its personalized local and global models, allowing for tailored healthcare solutions. This approach ensures that the models are more accurate and effective, taking into account the specific needs and characteristics of each patient cluster. This personalized FL approach enhances the accuracy of predictions, treatment recommendations, and overall healthcare outcomes [3].

3.3. Internet of things: automated vehicles

FL also holds immense potential in the domain of automated vehicles, revolutionizing transportation systems and enhancing safety and efficiency.

Unmanned aerial vehicles (UAVs) can greatly benefit from FL techniques. By leveraging data collected from distributed UAVs, FL enables the development of robust models for navigation, object detection, and collision avoidance. The decentralized nature of FL allows UAVs to learn from each other's experiences and adapt to changing environmental conditions, enhancing their autonomy and overall performance [3].

In the context of electric vehicles (EVs), FL can be applied to provide effective energy demand forecasting services for charging station (CS) providers. By aggregating data from multiple Evs and charging stations, FL algorithms can be used to predict and optimize energy demand, ensuring efficient utilization of charging infrastructure. This enables CS providers to dynamically manage resources, minimize energy wastage, and enhance the overall charging experience for EV users [6].

Autonomous vehicles, which are a key component of the future transportation landscape, can benefit greatly from FL approaches. By combining data from various vehicles, FL enables the development of accurate and robust models for perception, decision-making, and control. This collaborative learning approach enhances the safety and reliability of autonomous vehicles by leveraging the collective knowledge gained from diverse driving scenarios and environments [6].

FL can also play a significant role in the car insurance industry. By leveraging decentralized data from various vehicles, FL algorithms can effectively identify risks, predict compensation costs, and provide personalized insurance services. This approach allows insurance providers to tailor their offerings based on individual driving behavior, leading to fairer pricing, improved risk assessment, and enhanced customer satisfaction [6].

4. Challenges of federated learning

4.1. Non-IID data

One of the significant challenges in FL is the presence of non-IID (non-identically distributed) data across the participating clients. Non-IID data refers to the scenario where the data distributions among the clients are different, leading to variations in the statistical properties of the data [7]. This challenge can manifest in various ways:

Feature distribution skew (covariate shift): Clients may have different feature distributions, meaning that the input variables in their datasets exhibit variations. This discrepancy in feature distributions poses challenges in building accurate and generalizable models.

Label distribution skew (prior probability shift): The distribution of labels (target variables) among the clients may differ. This can occur when certain classes are overrepresented or underrepresented in specific clients' datasets. This label distribution skew poses challenges in learning unbiased models.

Same label, different features (concept shift): Clients may have different feature representations for the same label. This occurs when different clients use distinct feature engineering techniques or collect data from diverse sources. As a result, the features associated with the same label may vary across clients, making it challenging to generalize the learned models.

Same features, different labels: Conversely, clients may have different labels assigned to the same set of features. This can occur due to variations in labeling criteria or subjective interpretations of the data. Such inconsistencies in label assignments hinder the creation of consistent and accurate models.

Quantity skew or unbalancedness: The amount of data available to different clients may vary significantly. Some clients may have large datasets, while others may have limited data. This quantity skew poses challenges in achieving fair and representative model updates across all clients [7].

To address non-IID data challenges, several strategies can be employed. Data sharing allows clients to share subsets of their data to create a more balanced dataset. Data augmentation techniques can be used to artificially expand datasets and introduce diversity [7]. Algorithm-based approaches, like Federated Averaging [1], can adaptively weigh client contributions based on their data distributions. However, these methods do not solve the problem completely.

4.2. Systems heterogeneity

Systems heterogeneity is a significant challenge in FL, arising from the variability in hardware, network connectivity, and power among the participating devices. Clients in an FL setting can differ in terms of their hardware capabilities, such as CPU power and memory capacity. Additionally, network connectivity may vary from 3G, 4G, 5G, or Wi-Fi, leading to differences in communication speed and reliability. Furthermore, power availability, influenced by battery levels, adds another dimension of heterogeneity [1].

The presence of systems heterogeneity poses several challenges in FL. Firstly, the network size and systems-related constraints often result in only a small fraction of devices being active at any given time, even in large-scale deployments. For instance, in a network of a million devices, only a few hundred may be active simultaneously [1]. Moreover, individual devices may exhibit unreliability, leading to dropout events during iterations due to connectivity issues or energy constraints.

To address systems heterogeneity in FL, various techniques can be employed. Asynchronous communication allows devices to update models independently, accommodating different connectivity and power constraints. Active device sampling selects a subset of responsive devices for model updates. Fault tolerance mechanisms ensure continuity in the presence of device failures [1]. These strategies can solve part of the problem.

4.3. Privacy

In federated learning, communicating model updates during the training process can inadvertently reveal sensitive information, posing a risk to privacy. Despite the use of aggregation techniques, there is a possibility of unintentional information leakage. This vulnerability arises because model updates may contain implicit details about the data used for training, potentially exposing individual user characteristics or sensitive patterns [1].

These risks can manifest in different ways. For instance, an adversary with access to the model updates could analyze the changes over time and infer specific details about the training data or the users involved. Similarly, a compromised central server could gain insights into the private information of the participating devices by examining the aggregated updates.

To mitigate these privacy risks, various methods and techniques are employed. Secure computations, such as homomorphic encryption or secure multi-party computation, allow for collaborative model training on encrypted data without exposing the underlying information. Additionally, federated learning frameworks incorporate mechanisms to reduce the reliance on a trusted central server, minimizing the exposure of sensitive data during communication [7]. These methods can solve the problem to some extent.

5. Conclusion

In conclusion, this paper has provided an overview of federated learning, exploring its evolution, applications, and challenges. Federated learning offers a promising solution for collaborative model training while preserving data privacy. The discussed applications highlight their potential in smartphones, healthcare, and automated vehicles. Further research can focus on optimizing communication protocols, scalability, and privacy-preserving techniques. Future studies may explore applications in finance, energy, and social media, as well as integration with technologies like blockchain and edge computing. Overall, federated learning has the potential to revolutionize collaborative machine learning, and continued research can unlock its full potential for secure and efficient decentralized learning paradigms.

References

[1] Li, Tian, et al. “Federated Learning: Challenges, Methods, and Future Directions.” IEEE Signal Processing Magazine, vol. 37, no. 3, 2020, pp. 3–9, https://doi.org/10.1109/msp.2020.2975749.

[2] Zhu, Hangyu, et al. “Federated Learning on Non-IID Data: A Survey.” Neurocomputing, vol. 465, 2021, pp. 1–22, https://doi.org/10.1016/j.neucom.2021.07.098.

[3] Abdulrahman, Sawsan, et al. “A Survey on Federated Learning: The Journey from Centralized to Distributed on-Site Learning and Beyond.” IEEE Internet of Things Journal, vol. 8, no. 7, 2021, pp. 5478–5486, https://doi.org/10.1109/jiot.2020.3030072.

[4] Abhishek V A, et al. “Federated Learning: Collaborative Machine Learning Withoutcentralized

Training Data." International Journal of Engineering Technology and Management Sciences, 2022, pp. 355–357, https://doi.org/10.46647/ijetms.2022.v06i05.052.

[5] Nguyen, Dinh C., et al. "Federated Learning for Internet of Things: A Comprehensive Survey." IEEE Communications Surveys & Tutorials, vol. 23, no. 3, 2021, pp. 15–20, https://doi.org/10.1109/comst.2021.3075439.

[6] Zheng, Zhaohua, et al. "Applications of Federated Learning in Smart Cities: Recent Advances, Taxonomy, and Open Challenges." Connection Science, vol. 34, no. 1, 2021, pp. 11–14, https://doi.org/10.1080/09540091.2021.1936455.

[7] Kairouz, Peter, et al. "Advances and Open Problems in Federated Learning." Foundations and Trends® in Machine Learning, vol. 14, no. 1–2, 2021, pp. 18–58, https://doi.org/10.1561/2200000083.