Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Overview of Research Progress and Challenges in Federated Learning

Wenxin Guo*

College of Economics and Management, Chang'an University, Xi'an, China

*Corresponding author email: 2022900892@chd.edu.cn

Abstract. In recent years, federated learning has become a hot research topic in the machine learning community. It aims to reduce the potential data security and privacy risks caused by the centralized training paradigm of traditional machine learning through local training and global aggregation. Although federated learning methods have been widely applied in numerous fields such as finance, healthcare, autonomous driving, and smart retail, there are still urgent issues to be addressed in the field of federated learning, including data privacy leakage, malicious node attacks, model security, and the trustworthiness of participants. By delving into and discussing federated learning, this paper aims to provide researchers and practitioners in related fields with a comprehensive understanding and the latest progress of this technology. Based on the characteristics of the data distribution of the parties involved in federated learning training, this paper categorizes existing federated learning methods into horizontal federated learning, vertical federated learning, and federated transfer learning. It also introduces representative federated learning algorithms under different types, including their design concepts, basic processes, and advantages and disadvantages. Combining different application scenarios, this paper further discusses the challenges of federated learning and looks forward to the future development direction of this topic.

Keywords: Federated Learning; Data Heterogeneity; Data Privacy; Node Attacks.

1. Introduction

With the rapid development of big data and cloud computing technologies, machine learning has gained widespread application in many fields [1-3]. However, traditional centralized machine learning models face challenges regarding data security and privacy protection. To tackle these issues, federated learning, a novel machine learning technique, has emerged. Federated learning is a machine learning framework based on cloud and edge computing, allowing models to be trained distributively across multiple devices or nodes without needing to send original data to a centralized server. In this manner, federated learning can enhance the efficiency and accuracy of machine learning while protecting data privacy.

Since its inception, federated learning has caused a significant stir in academic and industrial circles. In recent years, federated learning has developed rapidly, and applications related to federated learning have started to take root across various sectors. However, this has also led to a series of security issues, such as data privacy leaks, attacks by malicious nodes, model security, and the trustworthiness of participants. These security threats pose new challenges to the implementation of federated learning. In response, federated learning needs to adopt appropriate privacy protection measures to ensure the security and reliability of the federated learning system. Therefore, researching security and privacy protection technologies within federated learning is of great significance.

Scholars domestically and internationally have conducted extensive research on federated learning and data security issues [4-6]. Based on the characteristics of the data distribution among participants in federated learning training, existing federated learning methods mainly include horizontal federated learning, vertical federated learning, and federated transfer learning. (1) Horizontal federated learning involves stacking and aligning the data attribute features of participants, suitable for scenarios where there is significant overlap in the distribution features of participants' data but less overlap in user samples. (2) Vertical federated learning involves stacking and aligning the data

samples of participants, applicable to scenarios where there is a substantial overlap in participants' data user samples but relatively less overlap in attribute features. (3) Different from horizontal and vertical federated learning, federated transfer learning neither aligns the data's attribute features nor the user samples. It is essentially a categorization meant to supplement horizontal and vertical federated learning, applicable in scenarios where there is less overlap in both the attribute features and user samples of each participant's data.

Through in-depth research and discussion on federated learning, this paper aims to provide researchers and practitioners in related fields with a comprehensive understanding and the latest progress of this technology. This paper will introduce the basic concepts, principles, and methods of federated learning, as well as its applications in various fields and future development prospects. Specifically, after further elaborating on the limitations of the traditional machine learning paradigm, the core design ideas of federated learning will be firstly elucidated. Subsequently, representative federated learning methods will be discussed, including their research problems, motivations, basic processes, and advantages and disadvantages. The paper will also analyze application cases of federated learning in various fields and discuss its advantages and limitations in different scenarios. Lastly, the paper will look forward to the future development trends of federated learning, including improvements in privacy protection technologies and cross-domain applications.

2. Machine Learning and Deep Learning

Machine learning is an important branch of artificial intelligence, with the core idea of enabling machines to exhibit intelligent behavior like humans. By imitating the human learning process, machines can continuously recognize things and accumulate experience, thus discovering patterns in data and making predictions on new data to enhance the performance of machine systems. However, traditional machine learning training methods have limitations, namely the need to collect all the data to be trained on a single server node, placing high demands on the data processing timeliness and speed of the central server node. With the continuous growth of data volume and increasing complexity of models, a single server node cannot meet the increasingly complex model training requirements due to computational and storage capacity limitations.

To address these issues, Distributed Machine Learning (DML) has emerged. DML technology typically disperses the training process across multiple subprocessor nodes for parallel processing. In actual execution, the central processor is responsible for distributing and scheduling different subsets of training data or parts of the model to various subprocessor nodes. At each processing node, models and data are computed in parallel; when each independent node completes its calculation, the central processor receives the parameters returned by these computing nodes and aggregates them to generate the final result. Although distributed optimization technology has helped machine learning algorithms achieve higher computational performance and training effects, this technology has not completely moved away from the centralized training mode. The training process incurs significant communication overhead and additional costs for synchronization. For example, if a computing node fails or data is lost, the subsequent work of all other computing nodes is also forced to halt. Moreover, due to centralized storage and uniform distribution of data, issues of data privacy leaks and security threats have become increasingly prominent. These challenges have forced traditional machine learning methods to simultaneously consider large-scale computation, distributed optimization, and the protection of privacy and data security.

Deep learning is an important research direction in the field of machine learning, aiming to enable machines to mimic human analytical learning capabilities, thereby recognizing and understanding data such as text, images, and sound. It achieves complex function approximation between input and output through self-learning of large samples' features, thereby enhancing the accuracy of classification or prediction. The core of deep learning lies in neural networks, a system that simulates the structure of human brain neurons. It consists of multiple layers, including the input layer, hidden layers, and output layer, connected by weights and biases. When an input signal enters the neural

network, it is transmitted layer by layer, undergoing iterative correction to achieve the desired output result. The training process of deep learning relies on the backpropagation algorithm, which optimizes network performance by comparing the error between network output and true labels, reversing the error signal, and updating network parameters accordingly. Additionally, activation functions and loss functions play crucial roles in deep learning. Activation functions increase the network's nonlinearity, while loss functions measure the difference between network output and true labels, optimizing the network by minimizing loss. Deep learning has achieved remarkable results in multiple fields, such as computer vision, natural language processing, speech recognition, recommendation systems, medical image recognition, financial risk control, intelligent manufacturing, and genomics. It has provided new methods and ideas for solving complex pattern recognition problems, propelling the rapid development of artificial intelligence technology. With continuous technological advancements and data accumulation, deep learning will play an increasingly important role in the future. It will continue to drive the development of artificial intelligence technology, bringing more convenience and innovation to humanity. Simultaneously, deep learning will face some challenges and issues, such as data privacy, algorithm fairness, and interpretability, which researchers need to continually explore and resolve.

3. Federated Learning

This section focuses on analyzing the limitations of centralized training and the solutions offered by distributed machine learning. Based on this, the section introduces the latest advancements in the field of federated learning from three perspectives: horizontal federation, vertical federation, and federated transfer, depending on the distribution of feature space and sample space of training data on the sub-servers.

3.1. Horizontal Federal Learning

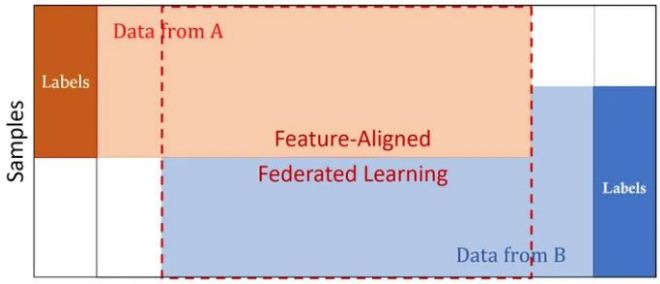

Figure 1. Horizontal Federal Learning Framework

One of the representative algorithms of horizontal federated learning is the Federated Averaging algorithm, FedAVG [7], as shown in **Figure 1**. FedAVG is a federated learning algorithm based on gradient averaging. It allows multiple devices to collaboratively train a machine learning model in a distributed environment while protecting users' data privacy. Specifically, the working principle of FedAVG: the server initializes a global model and sends it to all devices participating in the training. Each device trains the model using its local dataset and calculates new local model parameters. The devices upload the trained local model parameters to the server. To protect user privacy, devices do not need to upload original data, only model parameters. After receiving the model parameters uploaded by all devices, the server calculates the average of these parameters to obtain new global model parameters. The server sends the updated global model parameters back to each device, which then uses the new global model parameters for the next round of training. This process is repeated until the model converges or reaches a preset number of communication rounds. Since each device trains using its local dataset, FedAVG can support training on heterogeneous devices. At the same time, because only model parameters need to be transmitted and not original data, FedAVG can reduce communication overhead.

3.2. Vertical Federal Learning

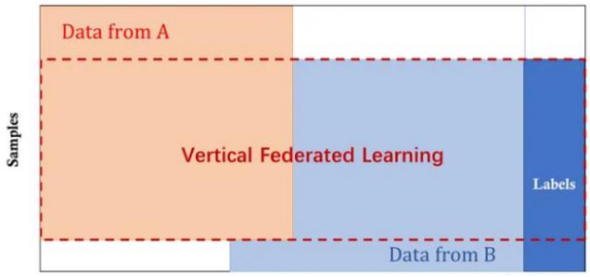

Figure 2. Vertical Federal Learning Framework

Representative algorithms of vertical federated learning include Secure Federated Linear Regression [8] and Secure Federated Boosting Trees [9], as shown in **Figure 2**. These algorithms aim to protect the local data privacy of each participant while achieving the training of machine learning models. In Secure Federated Linear Regression, participants use the gradient descent method to train the linear regression model while utilizing secure methods to calculate the model's loss and gradient. This approach ensures the participants' information is not leaked through encryption and collaboration, thus achieving data privacy protection. Additionally, there are other vertical federated learning algorithms based on homomorphic encryption, such as the method for implementing vertical federated logistic regression under the central federated learning framework proposed by Yang et al. [10]. This method incorporates the concept of homomorphic encryption, encrypting both parties' data and gradients during the training process to protect data privacy.

Specifically, the execution process of the Secure Federated Linear Regression algorithm involves participants initializing their model parameters and encrypting them using homomorphic encryption methods. Then, participants compute the encrypted loss and gradient using their local datasets. Since homomorphic encryption is used, these computations can be performed without exposing the original data. Next, the participants send the encrypted loss and gradient to the server. After receiving these data, the server aggregates them using homomorphic encryption methods to obtain the global loss and gradient. The server sends the global loss and gradient back to the participants, who then update their model parameters using this global information. Finally, this process is repeated until the model converges or reaches a predetermined number of communication rounds.

3.3. Federal Transfer Learning

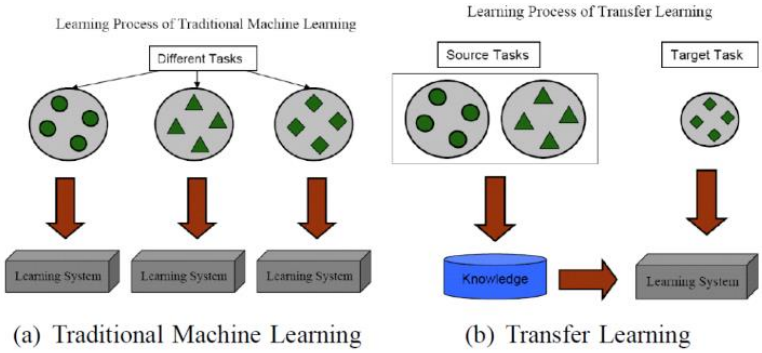

Figure 3. Comparison Chart between Transfer Learning and Traditional Machine Learning

The representative algorithms of federated transfer learning include the Federated Transfer Learning (FTL) algorithm [11]. This algorithm integrates the concepts of federated learning and transfer learning, aiming to address machine learning issues in a distributed environment where data distributions vary across different data sources and there is a need for privacy protection, as shown in **Figure 3**. Additionally, some studies have combined traditional machine learning algorithms with federated learning and transfer learning, proposing federated transfer learning algorithms suitable for specific scenarios. For example, federated transfer learning algorithms based on Support Vector

Machines (SVM) and Random Forest (RF) have been developed. These algorithms protect data privacy while fully utilizing data resources in a distributed environment to enhance model performance and generalization capabilities.

The core idea of the FTL algorithm is to use transfer learning to compensate for differences in data distribution among various data sources, while leveraging the framework of federated learning to protect data privacy. Specifically, the FTL algorithm can be broken down into the following steps:

(1) Model initialization: Choose a global model or initialize the local model parameters of each participant.

(2) Local training and transfer: Each participant trains the model using their local dataset and uses transfer learning techniques to adapt to the distribution differences between data sources. This can be achieved by adjusting model parameters, learning domain-invariant features, or utilizing pretrained models.

(3) Model interaction and update: Participants exchange model parameters or intermediate computational results and update the model based on this information. To protect data privacy, the exchanged information usually needs to be encrypted.

(4) Iterative optimization and convergence: Repeat the above steps for iterative optimization until the model converges or meets preset stopping criteria. It should be noted that the specific implementation of the FTL algorithm may vary according to different application scenarios and requirements. For example, different transfer learning strategies may be chosen based on the characteristics of the data and the type of task; different encryption methods and communication protocols may be selected based on privacy protection requirements.

In summary, the **Table 1** shows the design concepts of different federated learning methods and the suitable scenarios for different federated learning algorithms.

Table 1. Application scenario description for different federate learning methods

Division basis	Classification results	Overview of applicable scenarios
data distribution	Horizontal federal learning	More features overlap and less sample overlap
	Vertical Federal Learning	Samples overlap more and have less overlapping features
	Federal transfer learning	Both the characteristics and the sample overlap were less common

3.4. Security Research in Federated Learning

Federated learning is a distributed machine learning model, with security research primarily focusing on the model's availability and integrity, as well as data confidentiality and privacy. Privacy protection is at the core of federated learning security research. To protect participants' data privacy, federated learning employs a variety of privacy-preserving techniques such as differential privacy, homomorphic encryption, secure multi-party computation, and secret sharing. These techniques help protect data confidentiality during joint training processes and prevent attackers from inferring other participants' data samples through auxiliary datasets. Additionally, federated learning utilizes asymmetric encryption to ensure the security of data transmission, preventing data leakage. To counter potential attacks, federated learning implements a series of defensive measures, including identifying malicious users or samples, and defending against poisoning attacks through anomaly detection and adversarial training. Furthermore, by leveraging the decentralized, secure, transparent, and tamper-proof characteristics of blockchain, federated learning can also audit the computing process and monitor malicious behavior during model training, thus enhancing privacy protection.

Although these methods strengthen privacy protection, there are still potential vulnerabilities originating from communication protocols, data sharing, and model updates. Attackers might exploit these vulnerabilities to steal sensitive data or disrupt the training process. Therefore, analyzing the security vulnerabilities of federated learning models and implementing appropriate defensive measures is crucial.

Differential privacy is a robust privacy-preserving technology designed to prevent differential attacks, i.e., inferring sensitive information of specific individuals by analyzing changes in datasets. Its core idea is to add noise during the data querying or computing process, making the probabilities of obtaining the same results through model inference very close between two datasets differing by only one record. Measures of differential privacy primarily include the differential privacy budget and perturbation, reflecting the upper limit of how the output probability can change when one record is added or removed from the dataset. The smaller the value, the stricter the differential privacy conditions. The probability that model behavior can change arbitrarily is typically set to a small constant. There are various methods to implement differential privacy, one major method being the introduction of randomness into query results, such as using Laplace noise to create mechanisms that comply with differential privacy. The key lies in designing appropriate noise levels based on the sensitivity of the query (i.e., the extent to which the query results vary with data changes). Differential privacy can be categorized into local differential privacy, distributed differential privacy, centralized differential privacy, and hybrid differential privacy. Local differential privacy primarily implements training and privacy protection on the client side, possessing the potential for decentralization; distributed differential privacy involves perturbation through a trusted intermediary node; centralized differential privacy is carried out by servers; and hybrid differential privacy combines the aforementioned methods or more. Differential privacy has many advantages. Firstly, the noise added by differential privacy is insensitive to data scale. Even for large datasets, a small amount of noise can significantly alter data distribution, effectively protecting privacy. Secondly, differential privacy enables quantitative estimation of privacy risks. However, differential privacy also faces some challenges, such as balancing privacy protection with model accuracy and performance, and designing appropriate differential privacy mechanisms to address different types of queries and sensitivities. Additionally, implementing differential privacy may require additional computational resources and time, which are also considerations in practical applications. Overall, as a robust privacy-preserving technology, differential privacy not only protects individual privacy but also maintains data usability and model performance as much as possible. With ongoing technological advancements and improvements, differential privacy will be applied in more fields, providing a more solid guarantee for personal privacy protection.

Table 2. Comparison of classification performance of different methods on CIFAR100 dataset

Dataset	FedAVG	FedSAM	FedASAM	FedSAM+SWA	FedASAM+SWA
CIFAR-100(alpha=0, 5clients per round）	30.25	31.04	36.04	39.3	42.01
CIFAR-100(alpha=0, 10clients per round）	36.74	36.93	39.51	39.76	42.64

4. Experiment

4.1. Dataset

The common experimental datasets used in federated learning research include FEMNIST, Shakespeare, Twitter, CelebA, Synthetic Dataset, Reddit, MNIST, Fashion MNIST, CIFAR10, and CIFAR100, among others. The Shakespeare dataset is compiled from the complete works of Shakespeare and contains data from 1,129 users (characters from the works) with a total of 422,615 samples. CIFAR-100 contains images of 100 different categories. For each category, CIFAR-100 collects 500 images for the training set and 100 images for the test set. Each image has a resolution of 32x32 pixels.

4.2. Performance Comparison

To further analyze the strengths and weaknesses of different algorithms, this paper quantitatively compares the image recognition accuracy of various representative methods on the CIFAR-100 dataset, as shown in **Table 2**. For the classical federated averaging algorithm FedAVG, when using 5 sub-servers, it achieves an average classification accuracy of 30.25 on CIFAR-100. As the number of sub-servers increases to 10, the accuracy of FedAVG increases by 6.49%. Similar trends in results are also evident with the FedSAM and FedASAM methods. These results indicate that an increase in the number of sub-servers can to some extent enhance the model's classification performance. The intuitive explanation is that as sub-servers learn local data and build models, an increase in the number of sub-servers is beneficial for the global model to accommodate more variations in data distribution, significantly enhancing the model's generalization capability. Furthermore, an increase in the number of sub-servers often also means an increase in training data. Under a data-driven strategy, massive amounts of data can enable the model to learn more knowledge, thereby improving the final classification performance.

4.3. Discussion

To address issues like data silos and privacy protection in traditional machine learning research, federated learning implements distributed model training among multiple data sources that each possess local data, thereby balancing data privacy protection and shared computation. Despite significant progress in the research of federated learning, several challenges still urgently need to be addressed:

(1) Quantifying sub-server participation. Federated learning integrates knowledge learned by different sub-servers to build the final global model. Due to variations in data distribution, local computational capabilities, and sub-server network structures, the contributions of different sub-servers to the final model are clearly not uniform. Intuitively, during the aggregation of sub-server models, those with greater relevance to the global model need more participation, while the information from less relevant sub-servers should be weighted less. Therefore, precisely quantifying the value of different sub-servers is an important research issue.

(2) Optimizing reward mechanisms. For different sub-servers, based on quantitatively assessing their contribution levels, further encouraging those with higher contributions while restraining those with lesser contributions is crucial for more effective gradient updates and parameter learning in the model. Thus, how to optimize existing reward mechanisms is a research direction worth exploring deeply.

(3) Integrating horizontal and vertical federations. Current research typically focuses on either horizontal or vertical scenarios, but in practice, scenarios that mix both are common. Therefore, how to integrate existing horizontal and vertical federated learning algorithms to achieve hybrid federated learning also warrants attention.

5. Summary

As an emerging distributed machine learning paradigm, federated learning demonstrates a bright application prospect across various fields. This article explores the latest research advancements in federated learning from three perspectives: horizontal, vertical, and transfer, tailored to different application scenarios. It includes the design concepts and basic procedures of representative methods. Additionally, the article quantitatively compares the performance of these methods to analyze their respective advantages and disadvantages. Further, it discusses the progress in research on security issues within the field of federated learning, along with the challenges and future directions in this area.

References

[1] C Xu, W Lang, R Xin, K Mao, H Jiang.Generative detect for occlusion object based on occlusion generation and feature completing. Journal of Visual Communication and Image Representation, 2021, 78:103189.

[2] H Sun, W Lang, C Xu, N Liu, H Zhou. Graph-based discriminative features learning for fine-grained image retrieval. Signal Processing: Image Communication, 2023, 110:116885.

[3] C Xu, H Jiang, Peter Y., Khan Z. A., Y Chen. MHW-PD: A robust rice panicles counting algorithm based on deep learning and multi-scale hybrid window. Computers and Electronics in Agriculture, 2020, 173:105375.

[4] Wang J, Wang Y, Liu L, Yin H, Ye N, Xu C. Weakly Supervised Forest Fire Segmentation in UAV Imagery Based on Foreground-Aware Pooling and Context-Aware Loss. Remote Sensing, 2023, 15, no. 14: 3606.

[5] T. Y u, E. Bagdasaryan, and V . Shmatikov, "Salvaging federated learning by local adaptation," arXiv preprint arXiv:2002.04758, 2020.

[6] V . Smith, C.-K. Chiang, M. Sanjabi, and A. S. Talwalkar, "Federated multi-task learning," in Advances in Neural Information Processing Systems, 2017, pp. 4424-4434.

[7] H. Brendan McMahan, Eider M., Daniel R., Seth H., Blaise A. y. A. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics 2017.

[8] Reza N., Reihaneh T., Jan B., David B. Blumenthal. On the Privacy of Federated Pipelines. In SIGIR'21: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. July 2021, 1975-1979.

[9] K Cheng, T Fan, Y Jin, Y Liu, T Chen, Papadopoulos D., Q Yang, 2021. Secureboost: A lossless federated learning framework. IEEE Intelligent Systems, 2021, 36(6):87-98.

[10] Yang H, Dasdan A, Hsiao R L, et al. Map-reduce-merge: simplified relational data processing on large clusters. Proceedings of the 2007 ACM SIGMOD international conference on Management of data. 2007: 1029-1040.

[11] Liu, Y., Kang, Y., Xing, C., Chen, T. and Yang, Q. A secure federated transfer learning framework. IEEE Intelligent Systems, 2020, 35(4), pp.70-82.