Federated Learning Algorithms in Healthcare

Subjects: Computer Science, Artificial Intelligence

Contributor: Malliga Subramanian , Vani Rajasekar , Sathishkumar V. E. , Kogilavani Shanmugavadivel , P. S. Nandhini

Federated Learning (FL), an emerging distributed collaborative artificial intelligence (AI) paradigm, is particularly suitable for smart healthcare by coordinating the training of numerous clients, that is, in healthcare institutes, without the exchange of private data.

federated learning non-IID heterogeneity

1. Introduction

Recent technological improvements have altered intelligent healthcare powered by artificial intelligence (AI). Nevertheless, AI systems necessitate centralized data gathering and processing, which may be impractical in realistic healthcare contexts due to the scalability of modern healthcare networks and the need for data privacy. Federated Learning (FL), an emerging distributed collaborative AI paradigm, is particularly suitable for smart healthcare by coordinating the training of numerous clients, that is, in healthcare institutes, without the exchange of private data [1]. Deep learning, an AI technique, has proven its significant potential in medical image analysis for the early detection of chronic diseases by managing large amounts of health data to facilitate the delivery of healthcare services [2].

Smart healthcare systems have historically relied on cloud-based centralized AI functions for health data analytics. In light of the ever-increasing quantities of health data, this centralized system is inefficient because of the inherent delays in transmitting raw data, and the centralized system cannot accommodate substantial growth in network traffic. Furthermore, depending on a centralized server creates severe privacy risks [3]. This is of paramount importance, particularly in e-healthcare, where patient records are considered extremely confidential and protected by laws.

2. Applications of Federated Learning (FL)

Federated Learning (FL) has numerous business applications, including traffic prediction and monitoring, healthcare, communication, IoT, transportation, autonomous vehicles, medical AI, and more [4]. In all of these applications, FL facilitates the collaborative learning of several devices utilizing a shared model. Thus, utilizing information from a wide variety of sources makes FL a practical tool. Once gathered, these data are used to update the model. These devices will only transmit the information collected from the local model updates to the server.

Thus, FL is used where a secure transmission of private information is necessary. Below, researchers present a few domains where FL finds its application.

FL in healthcare

Because HIPAA and other rules make it difficult to share sensitive health information, healthcare is one of the industries that can gain the most from federated learning. A significant amount of data from diverse healthcare databases and devices can be used to develop AI models that adhere to regulations.

FL in Autonomous Vehicles

Because of its ability to make accurate predictions in real-time, FL is used in the development of autonomous vehicles. Information about current road and traffic circumstances might be added in real-time, facilitating faster learning and better decision-making. This could lead to a safer and more satisfying experience with autonomous vehicles. Hence, the automotive industry has great potential for the application of FL. However, at the present time, nothing but studies are being done in this area. These studies have shown that training time for predicting the angle at which a vehicle's wheels would turn when autonomously driving may be reduced through the use of FL.

FL in IoT:

For the development of smart and privacy-enhanced IoT systems, the notion of FL has recently been proposed [5]. For data training, FL is a distributed collaborative AI approach [6] that coordinates several devices with a central server without sharing actual datasets. To train a neural network, for instance, an intelligent IoT network may employ a group of IoT devices to act as "workers" communicating with a central server.

3. Independent and Identical Distribution (IID) vs. Non-IID Datasets

In general, real-world data are non-independent and identically distributed (non-IID). In contrast, the majority of existing analytical and machine-learning techniques rely on IID data. Therefore, a suitable method is required to manage such types of real-world datasets. When a sample is said to be identically distributed, there are no overarching patterns, no fluctuations in the distribution, and each item is drawn from the same probability distribution. Independent means that the sample items are all independent events. In other words, there is no relationship between them at all. Each subsample differs substantially and fails to accurately represent the dataset's underlying distribution if the data is non-IID.

The performance of the global model on non-IID federated data has been shown to be substantially poorer than that of IID [7][8][9]. Most of the present FL techniques frequently merely ignore the distribution divergence. The performance fairness is also made worse by the distribution divergence of non-IID data [10], which means that the established model produces noticeably varied local performance for various clients. This encourages researchers to use a non-IID FL approach by coordinating the local optimization goals for many clients. FL lets AI models train using private information without compromising privacy and helps to achieve better results in the healthcare field when there is a scarcity of data. It has paved the way for a lot of research [11].

He et al. [5] reconstructed FL as the FedGKT, a group knowledge transfer training algorithm, to train small CNNs on edge nodes and transfer their information through to a central server. FedGKT devises a variation of the alternating minimization technique. The authors used three different datasets (CIFAR-10, CIFAR-100, and CINIC-10) and their non-IID variations to train CNNs based on ResNet-56 and ResNet-110. Their findings indicate that FedGKT can achieve equivalent or perhaps slightly greater precision than FedAvg. A few studies have also attempted to investigate FL's benefits, drawbacks, and potential uses. The benefits and drawbacks of FL in IoT systems, as well as the ways in which it paves the way for a wide range of IoT applications, were explored by Zhang et al. [12]. Specifically, the authors presented seven major obstacles to FL in IoT platforms and discussed some new, promising solutions to these problems. In [2], the authors summarized the recent improvements of FL learning towards enabling FL-powered IoT applications, and they employed a set of measures, including sparsification, robustness, quantization, scalability, security, and privacy, to rigorously evaluate these developments. In addition to outlining various research questions and potential answers, [2] also developed a taxonomy for FL across IoT networks.

For non-IID FL, Zhang et al. [13] presented unified feature learning and an optimization objectives alignment approach (FedUFO). Extensive trials showed that FedUFO is superior to state-of-the-art techniques, including the competitive one data-sharing method. Sahu et al. [14] presented a system called FedProx to handle heterogeneity in federated networks in their study. According to them, FedProx may be thought of as a generalization and reparameterization of FedAvg. The authors have shown that, across several realistic federated datasets, FedProx enables more robust convergence than FedAvg. In order to better understand the effects of various hyperparameter optimization strategies in an FL system, Holly et al. [15] conducted extensive research. In an effort to reduce the overhead, the authors looked into a local hyperparameter optimization strategy that, unlike a global one, provides each client control over its own hyperparameter setting. The algorithms were tested on the MNIST data set with an IID partition and on an Internet of Things sensor-based industrial data set with a non-IID partition to determine how well they performed.

In [16], the authors described Auto-FedRL, a reinforcement learning-based federated hyperparameter optimization approach, where an online agent may dynamically alter the hyperparameters of each client depending on the current training status. Multiple search techniques and agents have been examined through extensive tests. Validation was performed on two real-world medical image segmentation datasets, one for COVID-19 lesion segmentation in chest CT and one for pancreas segmentation in abdominal CT, to ensure the effectiveness of the proposed technique.

Wang et al. [5] paid special attention to image processing applications that ensure the confidentiality of training data. Autonomous diagnosis of COVID-19 using deep learning algorithms has lately been a subject of research by scientists all around the world. It is true that several data-driven deep-learning approaches have been developed to make COVID-19 identification easier, but these data are still scarce due to patient privacy concerns. Under this circumstance, the solution is FL, which allows many organizations to work together to create a powerful deep-learning model without having to share sensitive data.

Zhu et al. [17] analyzed the impact of non-IID data on parametric and non-parametric machine learning models for use in horizontal and vertical FL in great detail. Both the pros and cons of existing methods for dealing with the difficulties posed by non-IID data in federated learning were examined. In addition, several studies on detecting COVID-19 using FL were found throughout the current literature analysis [18][19][20][21][22][23]. These studies proved that an FL approach may successfully detect COVID-19-related CT abnormalities using external validation on patients from various healthcare institutes.

Using differential privacy, Beguiere et al. [24] suggested a method for FL training of genomic cancer prediction models. They showed that it is possible to achieve a good trade-off between accuracy and privacy by training a supervised model for the prediction of breast cancer using genomic data, which is split across two virtual centers while preserving data privacy with regard to model transfer through DP. Choudary et al. [25] presented an FL framework capable of learning a centralized model using decentralized health data stored in many places. There are two tiers of privacy in this framework. For one, it does not involve any data transfer or central server use when training the model. As a second line of defense against privacy breaches, the model employs a differential privacy mechanism. The authors proved that FL architecture might increase privacy while retaining global model's usefulness. A stratified Cox model was created by Hansen et al. [26] and validated at three different locations in different countries using FL. Claiming their models are well-calibrated with high power of discrimination, they demonstrated exceptional overall survival prediction for patients with laryngeal cancer who were treated with curative radiotherapy in three separate cohorts.

Furthermore, researchers also reviewed a few recent attempts based on FL for health cases. A new federated learning framework was proposed [27] that can develop predictive models through peer-to-peer collaboration without raw data exchanges, and this work uses a distributed algorithm to solve a binary supervised classification issue to predict hospitalizations for cardiac events. Kumar and Singla [28] demonstrated how well FL approaches may be adapted to the healthcare system and how they function, particularly in the areas of medication discovery, clinical diagnostics, digital health monitoring, and various illness forecasts and detection systems. This work [28] also includes comparative analysis, which compares various FL algorithms for various health sectors using metrics including accuracy, the area under the curve, precision, recall, and F-score. In an attempt by [29], the way in which FL can facilitate the growth of an open health ecosystem with the use of AI was described. This work also covers current issues with FL and potential remedies. Antunes et al. [30] presented a systematic review of the research on FL in the context of an electronic health record for healthcare applications. In their analysis, they highlighted the most important research topics, proposed solutions, case studies, and ML methods. The research also presented a general architecture for applying FL to healthcare data.

To summarize, from the literature survey, it is understood that only very few attempts have been carried out using non-IID data in an FL environment, and these works have focused on COVID-19 detection only. These works update the global models irrespective of the local model result. Thus, in the current research attempt, researchers use three types of cancer that are distributed across the clients in a non-IID fashion and implement FL algorithms. In the proposed work, a client updates the weights to the global model if the current local model parameters provide improved performance.

References

1. Sheller, M.J.; Edwards, B.; Reina, G.A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.R. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Sci. Rep. 2020, 10, 12598.
2. Khan, L.U.; Saad, W.; Han, Z.; Hossain, E.; Hong, C.S. Federated learning for internet of things: Recent advances, taxonomy, and open challenges. IEEE Commun. Surv. Tutor. 2021, 23, 1759–1799.
3. Zhu, Y.; Shen, B.; Pei, X.; Liu, H.; Li, G. CT, MRI, and PET imaging features in cervical cancer staging and lymph node metastasis. Am. J. Transl. Res. 2021, 13, 10536.
4. Shaheen, M.; Farooq, M.S.; Umer, T.; Kim, B.-S. Applications of federated learning; Taxonomy, challenges, and research trends. Electronics 2022, 11, 670.
5. He, C.; Annavaram, M.; Avestimehr, S. Group knowledge transfer: Federated learning of large cnns at the edge. Adv. Neural Inf. Process. Syst. 2020, 33, 14068–14080.
6. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Poor, H.V. Federated learning for internet of things: A comprehensive survey. IEEE Commun. Surv. Tutor. 2021, 23, 1622–1658.
7. Jeong, E.; Oh, S.; Kim, H.; Park, J.; Bennis, M.; Kim, S.-L. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. arXiv 2018, arXiv:1811.11479.
8. Shen, T.; Zhang, J.; Jia, X.; Zhang, F.; Huang, G.; Zhou, P.; Kuang, K.; Wu, F.; Wu, C. Federated mutual learning. arXiv 2020, arXiv:2006.16765.
9. Liu, L.; Zhang, J.; Song, S.; Letaief, K.B. Client-edge-cloud hierarchical federated learning. In Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6.
10. Li, T.; Sanjabi, M.; Beirami, A.; Smith, V. Fair resource allocation in federated learning. arXiv 2019, arXiv:1905.10497.
11. Mohammadi, R.; Shokatian, I.; Salehi, M.; Arabi, H.; Shiri, I.; Zaidi, H. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. Radiother. Oncol. 2021, 159, 231–240.

12. Zhang, T.; Gao, L.; He, C.; Zhang, M.; Krishnamachari, B.; Avestimehr, A.S. Federated Learning for the Internet of Things: Applications, Challenges, and Opportunities. IEEE Internet Things Mag. 2022, 5, 24–29.
13. Zhang, L.; Luo, Y.; Bai, Y.; Du, B.; Duan, L.-Y. Federated learning for non-iid data via unified feature learning and optimization objective alignment. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021; pp. 4420–4428.
14. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. Proc. Mach. Learn. Syst. 2020, 2, 429–450.
15. Holly, S.; Hiessl, T.; Lakani, S.R.; Schall, D.; Heitzinger, C.; Kemnitz, J. Evaluation of hyperparameter-optimization approaches in an industrial federated learning system. In Data Science–Analytics and Applications; Springer: Berlin/Heidelberg, Germany, 2022; pp. 6–13.
16. Guo, P.; Yang, D.; Hatamizadeh, A.; Xu, A.; Xu, Z.; Li, W.; Zhao, C.; Xu, D.; Harmon, S.; Turkbey, E. Auto-FedRL: Federated Hyperparameter Optimization for Multi-institutional Medical Image Segmentation. arXiv 2022, arXiv:2203.06338.
17. Zhu, H.; Xu, J.; Liu, S.; Jin, Y. Federated learning on non-IID data: A survey. Neurocomputing 2021, 465, 371–390.
18. Kumar, R.; Wang, W.; Yuan, C.; Kumar, J.; Qing, H.; Yang, T.; Khan, A.A. Blockchain based privacy-preserved federated learning for medical images: A case study of COVID-19 CT scans. arXiv 2021, arXiv:2104.10903.
19. Dou, Q.; So, T.Y.; Jiang, M.; Liu, Q.; Vardhanabhuti, V.; Kaissis, G.; Li, Z.; Si, W.; Lee, H.H.; Yu, K. Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. NPJ Digit. Med. 2021, 4, 1–11.
20. Kumar, R.; Khan, A.A.; Kumar, J.; Golilarz, N.A.; Zhang, S.; Ting, Y.; Zheng, C.; Wang, W. Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging. IEEE Sens. J. 2021, 21, 16301–16314.
21. Yang, Q.; Zhang, J.; Hao, W.; Spell, G.P.; Carin, L. Flop: Federated learning on medical datasets using partial networks. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Singapore, 14–18 August 2021; pp. 3845–3853.
22. Salam, M.A.; Taha, S.; Ramadan, M. COVID-19 detection using federated machine learning. PLoS ONE 2021, 16, e0252573.
23. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Zomaya, A.Y. Federated learning for COVID-19 detection with generative adversarial networks in edge cloud computing. IEEE Internet Things J. 2022, 9, 10257–10271.

24. Beguier, C.; Du Terrail, J.O.; Meah, I.; Andreux, M.; Tramel, E.W. Differentially private federated learning for cancer prediction. arXiv 2021, arXiv:2101.02997.

25. Choudhury, O.; Gkoulalas-Divanis, A.; Salonidis, T.; Sylla, I.; Park, Y.; Hsu, G.; Das, A. Differential privacy-enabled federated learning for sensitive health data. arXiv 2019, arXiv:1910.02578.

26. Hansen, C.R.; Price, G.; Field, M.; Sarup, N.; Zukauskaite, R.; Johansen, J.; Eriksen, J.G.; Aly, F.; McPartlin, A.; Holloway, L. Larynx cancer survival model developed through open-source federated learning. Radiother. Oncol. 2022, 176, 179–186.

27. Brisimi, T.S.; Chen, R.; Mela, T.; Olshevsky, A.; Paschalidis, I.C.; Shi, W. Federated learning of predictive models from federated electronic health records. Int. J. Med. Inform. 2018, 112, 59–67.

28. Kumar, Y.; Singla, R. Federated learning systems for healthcare: Perspective and recent progress. In Federated Learning Systems; Springer: Berlin/Heidelberg, Germany, 2021; pp. 141–156.

29. Long, G.; Shen, T.; Tan, Y.; Gerrard, L.; Clarke, A.; Jiang, J. Federated learning for privacy-preserving open innovation future on digital health. In Humanity Driven AI; Springer: Berlin/Heidelberg, Germany, 2022; pp. 113–133.

30. Antunes, R.S.; Da Costa, C.A.; Küderle, A.; Yari, I.A.; Eskofier, B. Federated Learning for Healthcare: Systematic Review and Architecture Proposal. ACM Trans. Intell. Syst. Technol. (TIST) 2022, 13, 1–23.