IgMin Research
STEM A Multidisciplinary Open Access Journal

TECHNOLOGY SUBJECT
Data Science | Machine Learning | Data Security TOPIC(S)

Federated Learning- Hope and Scope

Lhamu Sherpa* and Nandan Banerji

Article Information

Department of Computer Science and Engineering, Sikkim Manipal Institute of Technology, Sikkim, India

Correspondence: Lhamu Sherpa, Department of Computer Science and Engineering, Sikkim Manipal Institute of Technology, Sikkim, India, Email: lhamusherpa1206@gmail.com

Submitted: October 10, 2023
Approved: November 10, 2023
Published: November 16, 2023

How to cite this article: Sherpa L, Banerji N. Federated Learning- Hope and Scope. IgMin Res. Nov 16, 2023; 1(1): 022-024. IgMIn ID: TEC063A112; DOI: 10.61927/igmin112; Available at: www.igminresearch.com/articles/pdf/TEC063A112.pdf

Keywords: Conventional Machine Learning (CML); Data obesity; Federated learning; Fintech services; Machine Learning (ML); Smart-Health

Abstract

People are suffering from" data obesity" as a result of the expansion and quick development of various Artificial Intelligence (AI) technologies and machine learning fields. The management of the current techniques is becoming more challenging due to the data created in the Smart-Health and Fintech service sectors. To provide stable and reliable methods for processing the data, several Machine Learning (ML) techniques were applied. Due to privacy-related issues with the aforementioned two providers, ML cannot fully use the data, which becomes difficult since it might not give the results that were expected. When the misuse and exploitation of personal data were gaining attention on a global scale and traditional machine learning (CML) was facing difficulties, Google introduced the concept of Federated Learning (FL). In order to enable the cooperative training of machine learning models among several organizations under privacy requirements, federated learning has been a popular research area. The expectation and potential of federated learning in terms of smart-health and fintech services are the main topics of this research.

Introduction

Federated Learning, additionally known as collaborative learning, is a ML setting where a group of client devices (such as smart devices, handheld devices, or fairly capable accessible systems) come together, form a team, and collaborate to build their model in a decentralized manner as shown in Figure 1.

Only the updates are shared among the devices in federated learning; the data is not. Model training and model inference are the two steps of federated learning [1]. Information, but not data, can be transferred between parties throughout the model

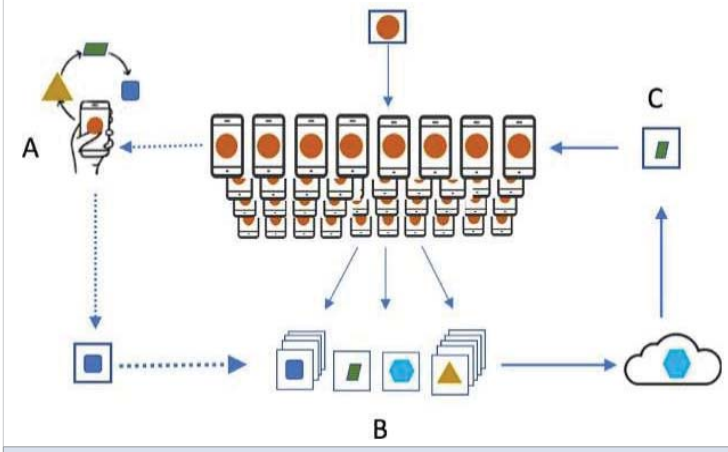

Figure 1: Federated Learning [4].

DOI: 10.61927/igmin112

training process.

No secure, private data at either site is exposed as a result of the transaction. The practiced model may be kept by one party or distributed among several parties [2]. With FL, distributed ML can be successfully executed by a number of users or computer nodes while safeguarding enormous amounts of data, exchanging expertise, ensuring the confidentiality of mobile data, and abiding by the law. If user activities do not provide results, (for example, by guessing the user's next word when typing), developers can't expect users to put in the extra work to label the machine learning model's training data. Language modelling is an example of an unsupervised learning system in which federated learning is helpful [3]. FL allows for the faster, more efficient, and less power-intensive development and testing of better models. The primary objective of FL's creators is to create models of machine learning utilizing data sets dispersed across several devices while preventing data loss. FL can be characterized as follows: Horizontal Federated Learning (HFL), Vertical Federated Learning (VFL) and Federated Transfer Learning (FTL) [2].

The datasets used in HFL are all comparable; for instance, take the case of two banks that are situated in separate cities (X and Y). The users can access data from users in bank X while using data from users in bank Y since the banks are in the same industry and keep user data with the same characteristics. The schematic diagram of HFL is shown in Figure 2.

The HFL may be summed up as [1]:

$$X_i = X_j = Y_i = Y_j, I_i, I_j, D_i, D_j, i, j \quad (1)$$

The datasets in vertical federated learning are all complimen- tary, i.e., the bank and the hospital are situated nearby. Al- though they have information on local customers, banks only record financial information, whereas hospitals only record medical data.

The schematic diagram of VFL is shown in Figure 3.

The VFL may be summed up as [1]:

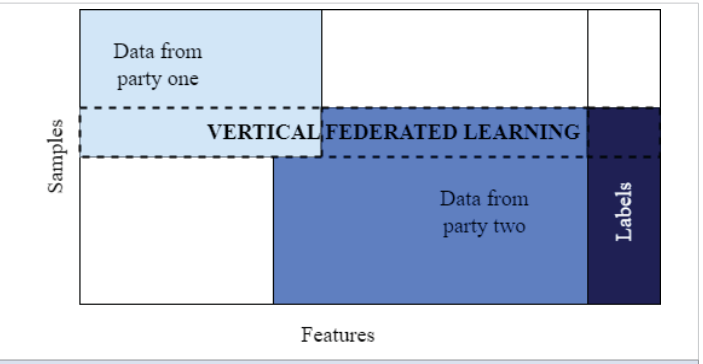

Figure 3: Vertical Federated Learning.

$$X_i \neq X_j, Y_i \neq Y_j, I_i = I_j \forall D_i, D_j, i \neq j \quad (2)$$

In FTL, we take lessons from a trained model. We can identify the consumers who save money at location A and receive treatment at location B using the FTL example given above. The schematic diagram of FTL is shown in Figure 4.

The FTL may be summed up as [1]:

$$X_i \neq X_j, Y_i \neq Y_j, I_i \neq I_j \forall D_i, D_j, i \neq j \quad (3)$$

The installation of necessary components in a system, such as edge devices, cloud servers, or mobile devices, is known as federated learning deployment. The procedure entails creating the architecture of the system, choosing appropriate models and algorithms, putting the system into practice, testing and verifying it, and then releasing it to users. One difficulty is protecting user data privacy and security, which may be done via sophisticated cryptographic methods.

According to the authors in [5], some of the most prevalent frameworks used in FL are as follows:

- ***TensorFlow Federated (TFF)*:** TFF is an open-source TensorFlow module that provides an API for FL. TFF offers several optimizations for effective model training and enables developers to easily construct and test federated models.
- ***PySyft:*** A secure and decentralized machine learning application may be created using the open-source Python module *PySyft*. FL, differential privacy, and

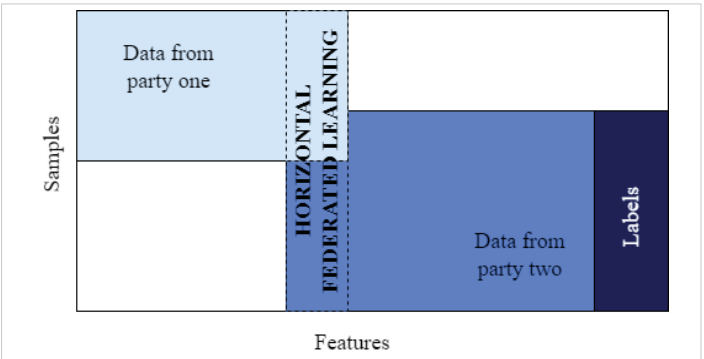

Figure 2: Horizontal Federated Learning (HFL).

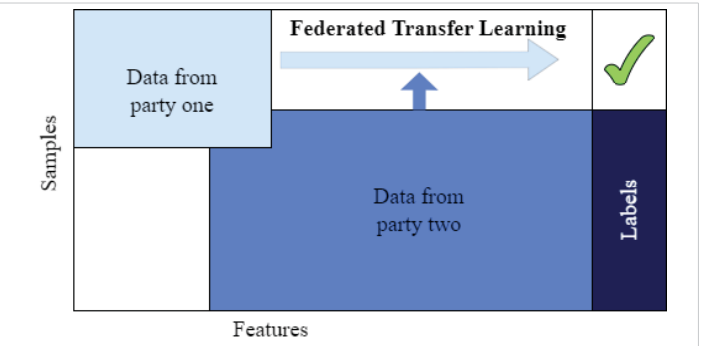

Figure 4: Federated Transfer Learning.

DOI: 10.61927/igmin112

other methods for handling delicate data are supported by *PySyft*.

- ***Flower:*** Another open-source FL frame- work is Flower. Developers may simply construct and manage federated models using Flower, which offers an abstraction for federated learning. Various methods are supported by Flower to enhance model training.
- ***IBM Federated Learning*:** IBM FL is a platform offered by the company which enables developers to build and train models on distributed data while also providing tools for regulating data security and privacy.

The decision relies on the particular requirements of the project because these frameworks vary in several ways, such as support for various models and optimizations. However, they are certainly an ideal place to begin using federated learning.

Study on smart health and fintech services

Federated learning offers a wide range of applications, including sales, smart health, financial services, and many more, since it can train a single model on data from various sources while maintaining data privacy and security. However, this paper focuses on the applications of FL over *Smart Health* and *Fintech services* because of a variety of criteria such as *privacy protection*, *integrity*, *data security*, and others, the data acquired from these services cannot be aggregated for the training of ML models.

Without disclosing the data, a multi-party database querying technology called FL may be employed. In the case of fintech services, multiparty borrowing poses a significant risk to the financial services industry. This happens when certain people fraudulently borrow through one bank and return the loan at another, putting financial stability at risk. A large number of such illicit transactions might bring the entire banking system down. Therefore, a federated learning architecture may be utilized to identify multiparty borrowing across banks X and Y without disclosing the user list to one another. A vertical federated learning framework-like action is one in which the user lists at each party are encrypted via federated learning's encryption mechanism, and then the encrypted lists are connected inside the federation. The list of multi-party borrowers is revealed upon decryption of the final result, but the other party is not made aware of the identities of the other moral users. Smart healthcare is another area that will gain a lot from the expansion of federated learning practices. Clinical trial data, medical reports, patient identification, and other types of medical data are extremely cryptic and private, but medical data are difficult to get since they are kept in inaccessible medical facilities and hospitals. Due to restricted access to data and a lack of labels, the machine learning models function poorly, becoming the bottleneck of existing smart healthcare [6]. We believe that if the entire medical community joined forces and shared their data, the accuracy and efficiency of ML model trained on a big medical dataset would be significantly improved and the basic strategy for fulfilling this goal is using a combination of transfer learning and federated learning. Transfer learning might be used to add the labels that are missing, increasing the amount of data that is accessible and enhancing the performance of a trained model. FTL would thus be critical in advancing the blooming of smart healthcare and could even raise the bar for patient care [1].

Conclusion

In recent years, data segregation and a focus on data privacy have emerged as the future artificial intelligence difficulties, while FL has provided us renewed hope. The word "federated learning" relates to both a corporate approach and a technological standard. It may give a uniform paradigm for various enterprises while securing local data, allowing organizations to prosper together on the cornerstone of data security.

In general, this paper discusses the core principles and methodologies of federated learning and examines its potential applications in the sectors of smart health and financial services. The application of FL in smart-health and fintech services is still in its infancy, but it will rapidly evolve in the coming years for the provision of privacy-enhancing and smart health services. FL is expected to play a significant influence in the development of large-scale and collaborative systems in order to migrate from centralized data analytics to distributed operations with privacy awareness.

References

1. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: Concept and applications. 2019.
2. Yang Q, Liu Y, Cheng Y, Kang Y, Chen T, Yu H. Federated Learning, ser. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2019. https://books.google.co.in/books?id=JdPGDwAAQBAJ
3. Long G, Tan Y, Jiang J, Zhang C. Federated learning for openbanking. 2021.
4. Hussain GKJ, Manoj G. Federated learning: A survey of a new approach to machine learning. In 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT). 2022; 1-8.
5. Stano M, Hluchy L, Boba´k M, Krammer P, Tran V. Federated learning methods for analytics of big and sensitive distributed data and survey. In *2023* IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI). 2023; 000 705–000 710.
6. Dasaradharami Reddy K, Gadekallu TR. A Comprehensive Survey on Federated Learning Techniques for Healthcare Informatics. Comput Intell Neurosci. 2023 Mar 1;2023:8393990. doi: 10.1155/2023/8393990. PMID: 36909974; PMCID: PMC9995203.

How to cite this article: Sherpa L, Banerji N. Federated Learning- Hope and Scope. IgMin Res. Nov 16, 2023; 1(1): 022-024. IgMIn ID: TEC063A112; DOI: 10.61927/igmin112; Available at: www.igminresearch.com/articles/pdf/TEC063A112.pdf