Federated Learning Technology that Enables Collaboration While Keeping Data Confidential and its Applicability to LLMs

ARAKI Toshinori

Abstract

The advancement of AI based on deep learning is remarkable, and large language models (LLMs) are even capable of natural interaction with humans. However, creating such AI technology requires a large amount of data, making data acquisition a crucial challenge. Federated learning is a technology that helps leverage highly confidential data scattered across multiple organizations, which is part of this challenge. In this paper, we explain three basic types (horizontal, vertical, and transfer) of federated learning and discuss the applicability of federated learning to generative AI, including LLMs, which have garnered significant attention in recent years.

Keywords

federated learning, data coordination, cross-industry collaboration, LLM, fine-tuning

1. Introduction

In recent years, the advancement of generative AI has been remarkable, and the large language model (LLM) GPT-3, which was announced in 2020, enables even natural interaction with humans, rapidly evolving thereafter. To imbue AI with such advanced functionalities, a vast amount of data is required, making data acquisition one of the most critical challenges in AI development.

The federated learning introduced in this paper addresses part of the data acquisition problem, allowing for the effective utilization of data scattered across multiple organizations. Merely centralizing data in one location poses risks of the provided data being used for purposes other than AI learning and potential leakage of personal information, thus compromising data confidentiality and privacy. However, with federated learning, it is possible to utilize all distributed data for AI creation without actually collecting the data itself.

In this paper, we introduce three types of federated learning suited for various scenarios and discuss the applicability of federated learning to generative AI models such as LLMs as advanced topics.

2. What is Federated Learning?

Federated learning is an AI learning technique proposed by Google[1)] in 2017. With this approach, AI learning can be conducted using all data without the need to aggregate data scattered across multiple locations.

Fig. 1 illustrates the basic framework of federated learning, showing three participants who want to train AI using their respective data to create a single global model. It is notable that the global model is created through a central server.

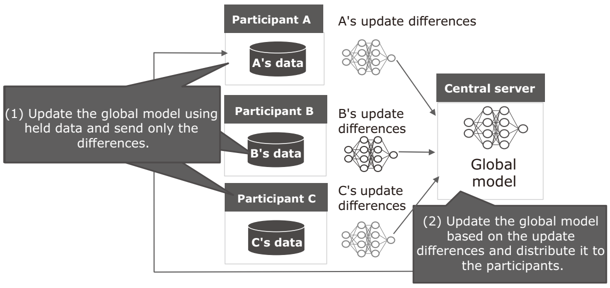
Participant A
A's data
A's update differences
(1) Update the global model using held data and send only the differences.
Participant B
B's data
B's update differences
Central server
Global model
Participant C
C's data
C's update differences
(2) Update the global model based on the update differences and distribute it to the participants.

Fig. 1 Basic framework of federated learning.

Learning is performed by the participants sharing the global model, who repeat the following steps (1) and (2):

(1) Each participant updates the global model using the data he/she holds and sends the update differences to the central server.

(2) The central server updates the global model based on the received update differences and distributes it to the participants. There are various methods for the central server to update the global model. For example, the server updates the global model with the average of the update differences.

In the aforementioned method, it is possible to create a global model that reflects all participants' data without ever collecting the data to the central server. The resulting global model is expected to have higher performance than learning with only the data held by each participant.

In this method, the central server will acquire the global model, but if it is desired to keep it confidential, a technique called "secure computation[2]," which allows computations to be performed on encrypted data, can be effective.

2.1 Horizontal federated learning

The method in which all participants own data of the same format and train AI for the same purpose, as explained at the beginning of section 2, is called "horizontal federated learning." This method is used when participants with similar data train AI with similar functions, such as when financial institutions collaborate to develop AI for fraud detection (**Fig. 2**).

Other federated learning approaches have been developed to accommodate participants who hold different types of data. As representative examples, we will introduce "vertical federated learning" and "transfer federated learning" next.

2.2 Vertical federated learning

In this section, we introduce "vertical federated learning," which is effective when participants have different types of data but share information about the same samples (such as users).

For example, a credit card company holds information such as users' income and assets, while an e-commerce site holds information about users' browsing history and preferences. Comparing information for each user will create rich data for each user, so AI developed using data from both company and site is expected to perform better than if it is developed using data from either the company or site alone.

One characteristic of "learning" in vertical federated learning is that participants divide and hold the AI during the learning process and compute it partially. Each participant inputs his/her respective data to perform part of the AI computation, and the central server, upon receiving partial computation results, calculates the rest (**Fig. 3**, forward direction →). Subsequently, the results are fed back in reverse to update the AI parameters, and this process is also distributed (Fig. 3, reverse direction →).

AI models trained in this manner are shared between the credit card company and the e-commerce site. However, since neither the credit card company nor the e-commerce site can gather the necessary data for predictions independently, they need to cooperate when utilizing AI. Furthermore, meaningful learning cannot occur if there is no correspondence between the data from the credit card company and the e-commerce site. To handle only the necessary correspondences while avoiding unnecessary information sharing, secure computation is useful.

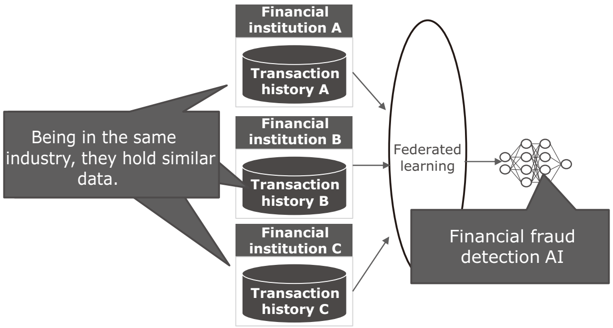

Fig. 2 Example of horizontal federated learning.

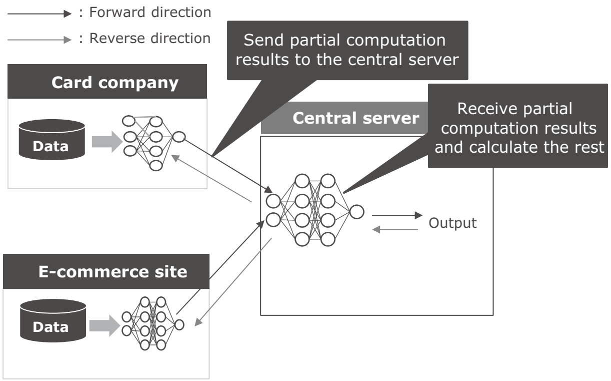
: Forward direction
: Reverse direction
Send partial computation results to the central server
Card company
Data
Central server
Receive partial computation results and calculate the rest
Output
E-commerce site
Data

Fig. 3 Outline of vertical federated learning.

2.3 Transfer federated learning

In this section, we introduce “transfer federated learning.” While vertical federated learning assumes that each participant holds different data for specific users, transfer federated learning is effective in cases where some of the data held by participants overlaps. This technique uses common data as glue to transfer AI created for one domain to another domain. It is considered useful for cross-industry collaboration. For example, in collaboration between an insurance company and a real estate company, the insurance company can discover potential real estate customers from its own customers and utilize them for referrals. Conversely, referrals from the real estate company to the insurance company are also possible (**Fig. 4**).

Learning in transfer federated learning consists of two steps. In the first step, preprocessing is performed to align the data held by each participant. Specifically, AI is collaboratively trained and exchanged to supplement the data held individually from attributes held in common (**Fig. 5**).

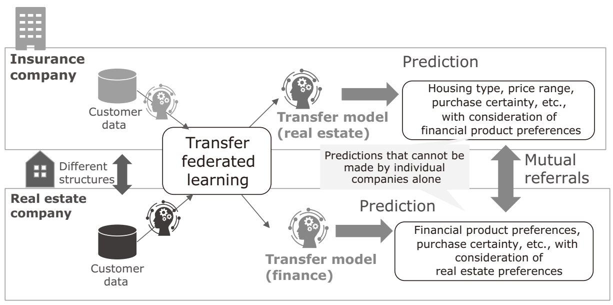
Insurance company
Customer data
Transfer federated learning
Transfer model (real estate)
Prediction
Housing type, price range, purchase certainty, etc., with consideration of financial product preferences
Different structures
Predictions that cannot be made by individual companies alone
Mutual referrals
Real estate company
Customer data
Transfer model (finance)
Prediction
Financial product preferences, purchase certainty, etc., with consideration of real estate preferences

Fig. 4 Illustration of the utilization transfer federated learning.

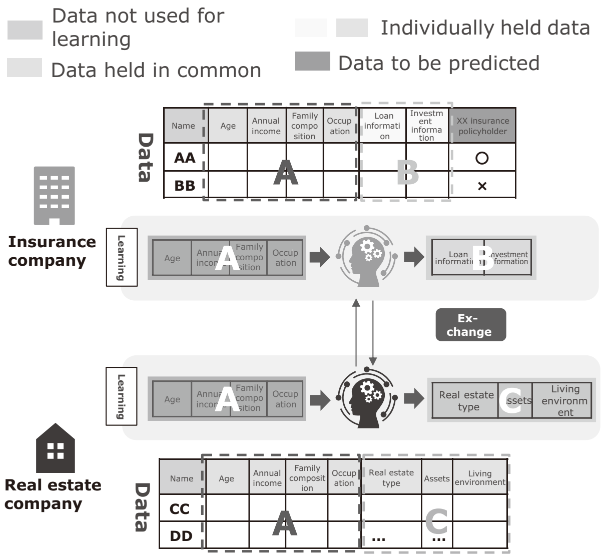

Fig. 5 Preprocessing to align held data.

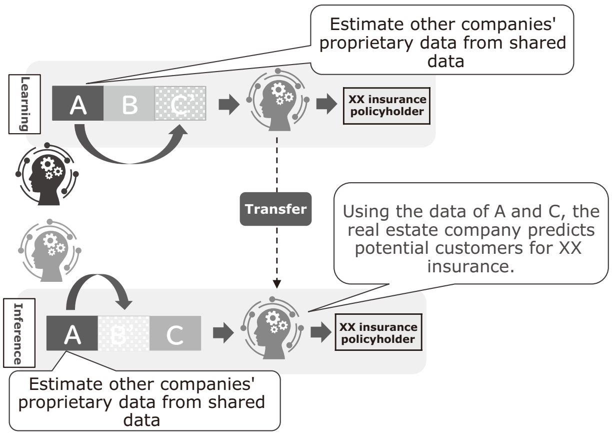
Estimate other companies' proprietary data from shared data
Learning
A
B
C
XX insurance policyholder
Transfer
Using the data of A and C, the real estate company predicts potential customers for XX insurance.
Inference
A
B
C
XX insurance policyholder
Estimate other companies' proprietary data from shared data

Fig. 6 AI learning and prediction using aligned data.

Next, each participant fills in the data he/she does not hold based on AI predictions, using the exchanged AI. Subsequently, the side holding the data to be predicted (XX insurance policyholders), i.e., the insurance company, trains the AI and transfers it to the real estate company. The real estate company evaluates the interest in XX insurance among its customers, using the transferred AI. Based on the evaluation results, contacting potential customers using direct mail or other means enables referrals to the insurance company to be realized (**Fig. 6**).

In the example mentioned earlier, it is assumed that both positive instances (policyholders) and negative instances (non-policyholders) of the prediction target—enrollment status for XX insurance—are available. However, for instance, if the goal is to gauge interest in the insurance product itself, the data held by the insurance company will at minimum consist of people who have inquired or visited, and can be considered positive instances (data on people with some interest). When no negative instances are available, AI learning becomes more challenging, but NEC has developed an effective transfer federated learning method[3] that works even when only positive instances exist.

3. Applicability of Federated Learning to Generative AI

In section 3, we discuss the usability and challenges of federated learning in generative AI, which has rapidly become widely used in many scenarios to address future challenges.

3.1 Applicability of horizontal federated learning to generative AI

Large language models (LLMs), a type of generative AI, are increasingly being used in cases where users adjust them based on the data they hold (referred to as fine-tuning), leading to the proliferation of fine-tuned LLMs in various locations (**Fig. 7**). While combining multiple scattered LLMs may potentially yield even higher-performance LLMs, fine-tuned LLMs and additional training data may constitute trade secrets that cannot be easily shared. Leveraging insights from federated learning, it is believed that integrated LLMs can be generated while keeping fine-tuned LLMs and their training data confidential. There seem to be technical challenges regarding how to effectively integrate a large amount of information into an LLM and how efficiently (in terms of communication and computation amounts) this integration can be done. However, horizontal federated learning can contribute to further expanding the scope of LLM usage.

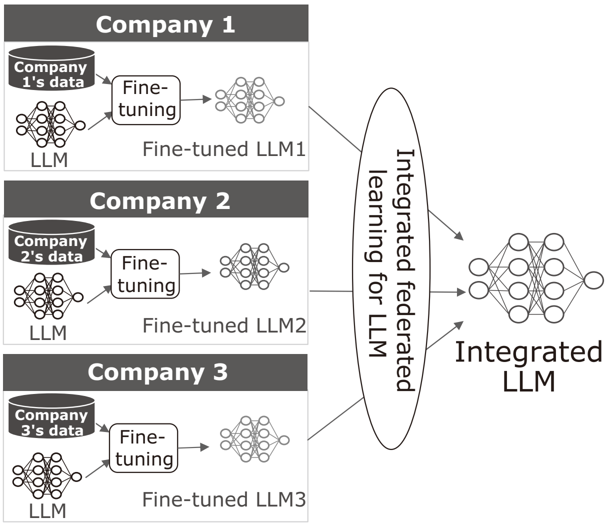

Fig. 7 Integration of Fine-tuned LLMs.

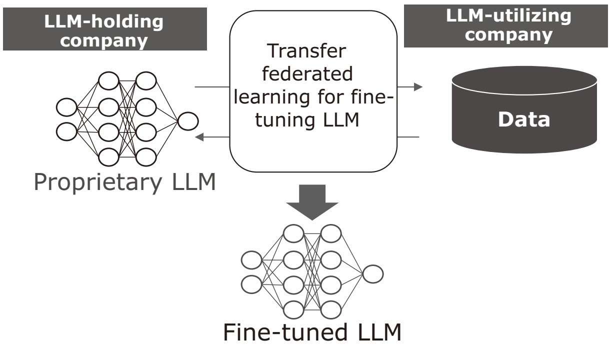

Fig. 8 Transfer federated learning for fine-tuning.

3.2 Applicability of transfer federated learning to generative AI

Alongside fine-tuning for LLMs, there is another challenge to consider. Let us consider a situation where a company providing LLM services holds proprietary LLMs, and user companies that want to fine-tune LLMs first and then utilize them hold additional training data. In this case, the LLM service provider may want to keep their proprietary LLMs confidential, while user companies may also want to keep sensitive information contained in the additional training data confidential (**Fig. 8**).

The following cases are conceivable: Proprietary LLMs that contribute to a company's competitiveness cannot be disclosed to other companies; and user companies want to keep raw sensitive information confidential from other companies. In such settings, the insights from transfer federated learning can be utilized for user companies to store (transfer) the training data they hold to LLMs. There are technical challenges regarding how much training data can be kept confidential and how efficiently (in terms of communication volume and required computational resources such as GPUs) fine-tuning can be performed.

4. Conclusion

In this paper, we have explained the overview and characteristics of the main federated learning approaches (horizontal, vertical, and transfer) and discussed the applicability of federated learning to the rapidly evolving field of generative AI.

* Google is a trademark of Google LLC.

* All other company names and product names that appear in this paper are trademarks or registered trademarks of their respective companies.

References

1) H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas: Communication-Efficient Learning of Deep Networks from Decentralized Data, PMLR 54, pp.1273-1282, 2017
https://arxiv.org/abs/1602.05629

2) NEC: Highly secure federated learning protects personal privacy and confidential information while leveraging AI, September 2021 (Japanese)
https://jpn.nec.com/rd/special/202103/index.html

3) Junki Mori, Ryo Furukawa, Isamu Teranishi and Jun Sakuma: Heterogeneous Domain Adaptation with Positive and Unlabeled Data, IEEE Big Data, 2023
https://arxiv.org/abs/2304.07955

Author’s Profile

ARAKI Toshinori
Director
Secure System Platform Research Laboratories

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.17 No.2 Special Issue on Revolutionizing Business Practices with Generative AI
– Advancing the Societal Adoption of AI with the Support of Generative AI Technologies

Orchestrating a brighter world NEC
NEC Technical Journal
Vol.17 No.2
June 2024
Special Issue on Revolutionizing Business Practices with Generative AI
Advancing the Societal Adoption of AI with the Support of Generative AI Technologies

Vol.17 No.2
June 2024

Special Issue TOP

Papers for Special Issue

Market Application of Rapidly Spreading Generative AI

Fundamental Technologies that Enhance the Potential of Generative AI

For AI Technology to Penetrate Society

NEC Information