"Para ocupar um lugar, divide a tuas tropas. Para expandir teu território, divide benefícios" (A Arte da Guerra, Sun Tzu).

Federated Learning

Paulo Ricardo Lisboa de Almeida

DSBD

UFPR

The problem

We want clients to collaborate with information to train our models, but without sharing their data.

Why?

The problem

We want clients to collaborate with information to train our models, but without sharing their data.

- Data privacy.
- Data minimization.
- Access rights.
- Data Security (avoid leaks).

Uses

- Medical domain.

Sotomaior, Fonseca, Zangari, Britto Jr, Viegas. A Federated Learning Model for Privacy-Preserving and Cross-Domain Kidney Stone Detection in Medical Imaging. SMC 2025.

A Federated Learning Model for Privacy-Preserving and Cross-Domain Kidney Stone Detection in Medical Imaging

Lucas Sotomaior, Luiz F. B. Fonseca, Matheus A. C. Zangari
Rodrigo K. Krebs, Alcen Britto Jr, and Eduardo K. Viegas

***Abstract*— Kidney stones significantly impact healthcare systems, with diagnosis typically requiring time-consuming Computed Tomography (CT) scan consultations between medics and radiologists, often delaying patient care. Achieving a quick and accurate diagnosis is essential to ensure timely and effective treatment, which has motivated the development of Deep Neural Network (DNN)-based approaches for automated kidney stone detection. However, building effective models remains challenging, as it often requires access to large and diverse datasets that are siloed across institutions, and sharing such medical data is rarely feasible due to strict privacy regulations and patient confidentiality concerns. This paper proposes a privacy-preserving Federated Learning (FL) framework that enables multiple medical institutions to collaboratively train a DNN model without sharing sensitive patient data. Each institution trains a local model on its private dataset, and a centralized trusted server securely aggregates model parameters. We evaluate our approach using abdominal CT scan image datasets from two distinct institutions. Experimental results demonstrate that our proposed model achieves high classification accuracy within the same training environment, with an F1-score of up to 0.94. In addition, in cross-dataset evaluations, our approach outperforms traditional centralized baselines, showing significantly lower performance degradation while preserving patient privacy.**

I. Introduction

The growing digitization of healthcare has resulted in the widespread availability of medical imaging data, driving the development of Machine Learning (ML) models to assist in diagnostics and treatment planning [1]. Among various medical challenges, automatically classifying kidney stones using Computed Tomography (CT) images is critical for guiding appropriate treatment strategies [2], specially in emergency room scenarios. In this direction, an ML model can be trained to detect kidney stones by learning patterns from annotated medical images, such as CT scans, where experts previously label the presence and characteristics of stones. During training, the model adjusts its internal parameters to accurately distinguish between images with and without kidney stones, enabling automated detection considering unseen data. Despite the promise of ML solutions in improving diagnostic accuracy, their implementation often requires sensitive patient data across multiple medical institutions. Such a need raises significant concerns regarding privacy, data protection, and compliance with legal frameworks such as the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) [3].

Recent works have demonstrated the effectiveness of Deep Neural Network (DNN) techniques for classifying and diagnosing diseases using medical images [4]. In general, proposed schemes use a centralized learning strategy, aggregating data from multiple sources into a single repository for model training, yielding significantly high detection accuracies [5]. However, despite their success, these approaches rely on the centralization of data, which is often impractical in real-world clinical settings due to privacy, legal, and logistical constraints. Sharing sensitive medical data across institutions can expose patients to privacy breaches, introduce risks of data misuse, and lead to compliance violations with strict regulatory frameworks [6]. Additionally, technical challenges related to data standardization, interoperability, and the secure transfer of large volumes of data further complicate the centralized model training in healthcare environments.

Federated Learning (FL) has emerged as a promising paradigm to address these challenges [7]. In practice, it enables the collaborative training of ML models across multiple decentralized devices or institutions without requiring the exchange of raw data. Instead, only model updates, such as gradients or parameters, are shared, thereby preserving data locality and protecting sensitive information. In the field of medical imaging, FL has been successfully applied to a range of tasks, including brain tumor segmentation [8], histopathological cancer detection [9], and retinal disease classification [10]. These studies often demonstrate that collaborative learning across institutions can produce models with performance comparable to or exceeding that of traditional centralized approaches. Additionally, FL offers further advantages, such as mitigating data silos, improving model generalization by leveraging heterogeneous datasets, and facilitating compliance with stringent data protection regulations.

Unfortunately, despite the progress in applying FL to various medical imaging tasks, its application to kidney stone detection remains largely unexplored. Kidney stone classification presents unique challenges, particularly due to the variability in imaging conditions across different clinical environments [11]. Overcoming domain shifts caused by variations in image acquisition protocols, patient populations, and equipment remains a significant challenge for

[1]Lucas Sotomaior, Luiz F. B. Fonseca, Alcen S. Britto Jr, and Eduardo K. Viegas are with Graduate Program in Computer Science (PPGIa) at Pontifícia Universidade Católica do Paraná (PUCPR), Brazil. Rodrigo K. Krebs and Matheus A. C. Zangari are with Urology Department of the Hospital Nossa Senhora das Graças (HNSG), Brazil. {lucas.sotomaior, luiz.fonseca, krebs.rodrigo, alcen, and eduardo.viegas}@ppgia.pucpr.br

Manuscript 1614 submitted to 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Received May 3, 2025.

Uses

- Medical domain.
- Financial data.

Uses

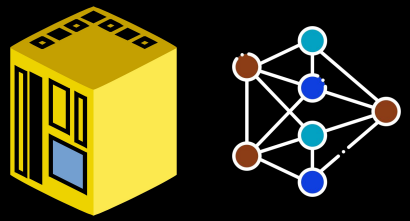
- Medical domain.
- Financial data.
- Domains with large volumes of distributed data.
    - E.g., Smart Cities.

Uses

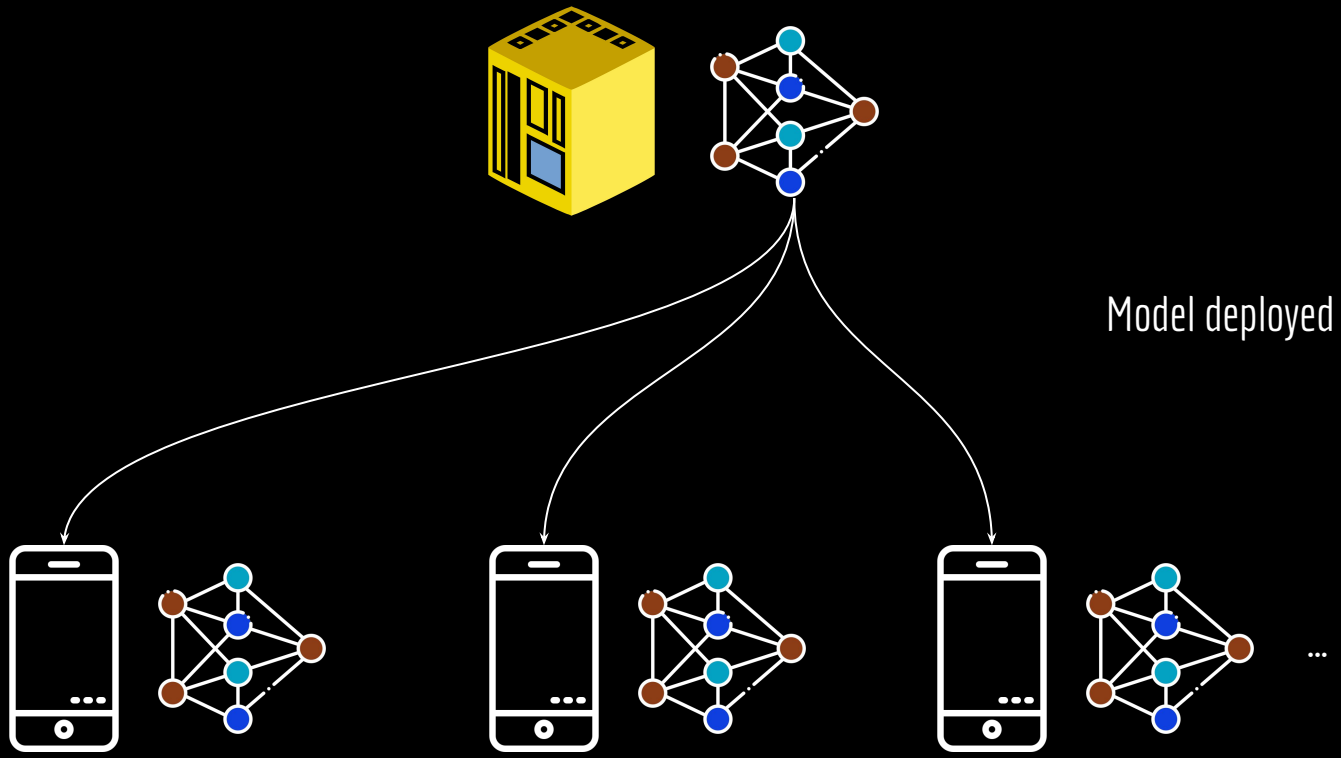
- Medical domain.
- Financial data.
- Domains with large volumes of distributed data.
  - E.g., Smart Cities.
- User data collection.
  - Your smartphone is absolutely not collecting your data.
  - E.g., Google Keyboard

Modeling

Central server trains the initial model.

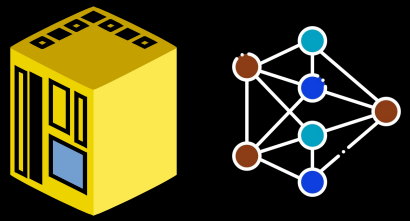


Modeling

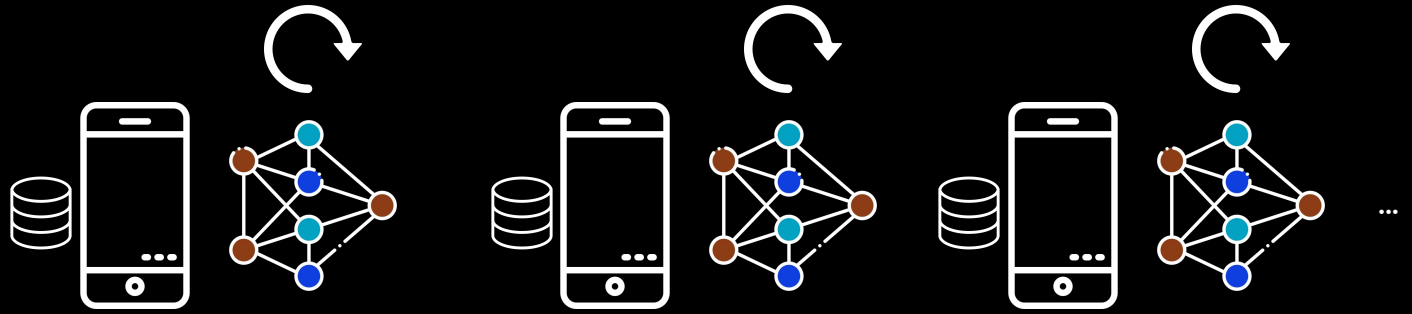


Model deployed in the clients.

Modeling

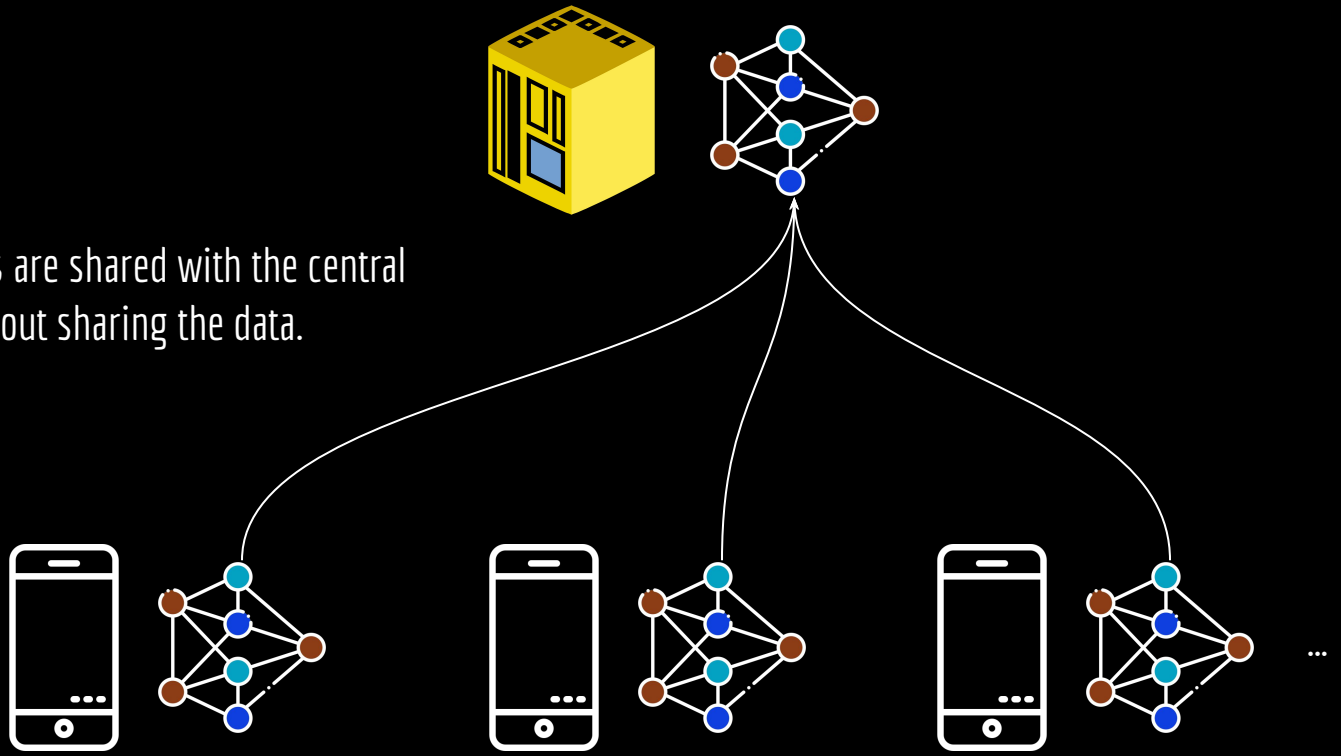


Clients update their models using their private data.

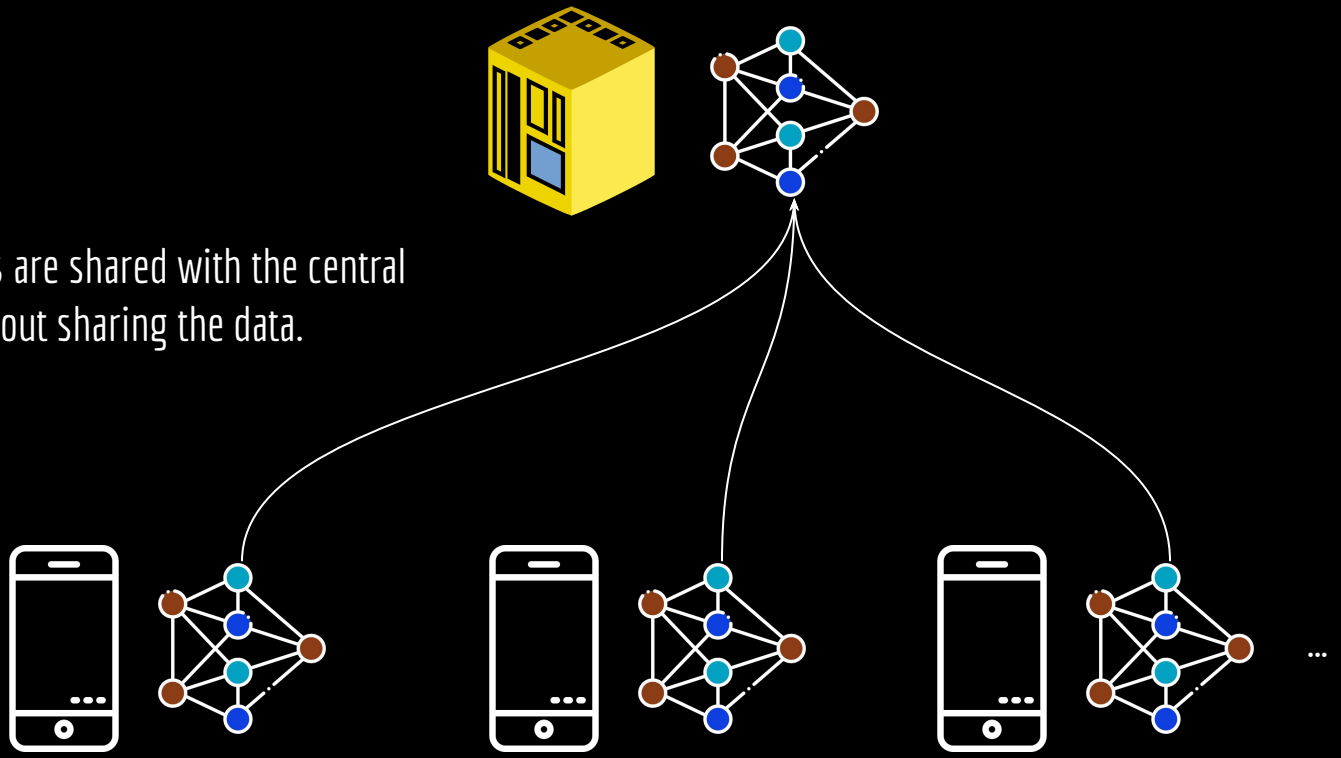


Modeling

The updates are shared with the central model, without sharing the data.



Modeling

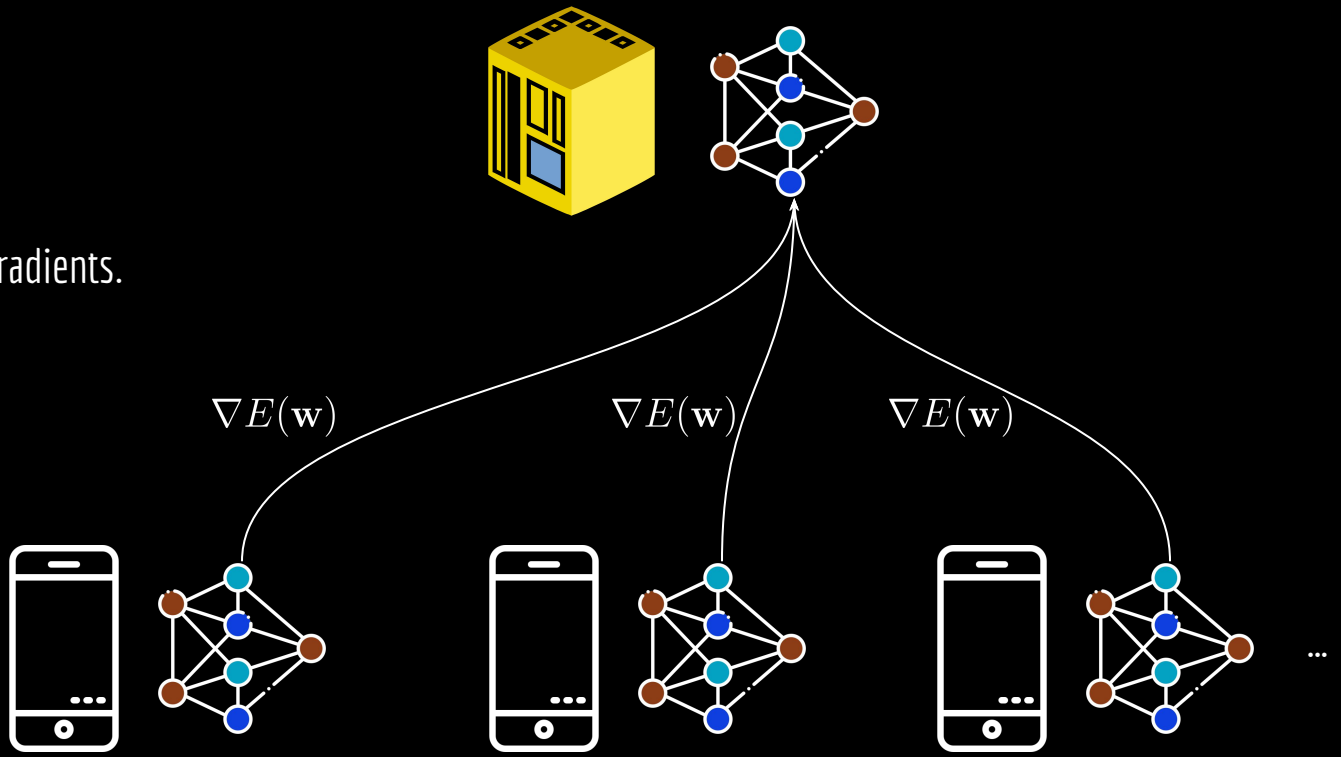


The updates are shared with the central model, without sharing the data.

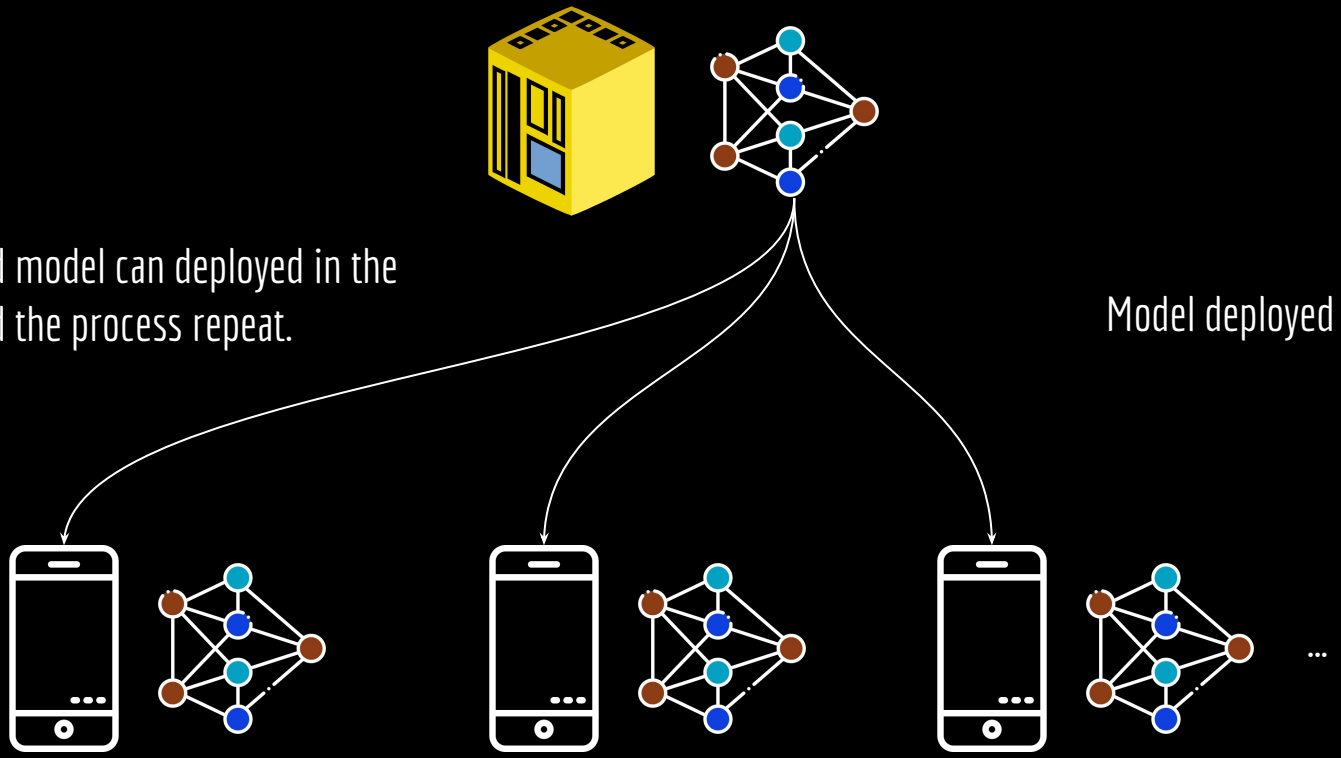
How?

Modeling

Share the gradients.



Modeling


The updated model can deployed in the devices, and the process repeat.
Model deployed in the clients.
...

This is Just One Possible Approach

This is defined as "Horizontal federated learning" in Zhang et al. 2025.

Check the paper.

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775.


Knowledge-Based Systems 216 (2021) 106775

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A survey on federated learning

Chen Zhang [a], Yu Xie [b,*], Hang Bai [a], Bin Yu [a], Weihong Li [a], Yuan Gao [c]

[a] School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi Province 710071, China
[b] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China
[c] School of Electronic Engineering, Xidian University, Xi'an, Shaanxi Province 710071, China

ARTICLE INFO

Article history:
Received 3 September 2020
Received in revised form 12 December 2020
Accepted 11 January 2021
Available online 21 January 2021

Keywords:
Federated learning
Privacy protection
Machine learning

ABSTRACT

Federated learning is a set-up in which multiple clients collaborate to solve machine learning problems, which is under the coordination of a central aggregator. This setting also allows the training data decentralized to ensure the data privacy of each device. Federated learning adheres to two major ideas: local computing and model transmission, which reduces some systematic privacy risks and costs brought by traditional centralized machine learning methods. The original data of the client is stored locally and cannot be exchanged or migrated. With the application of federated learning, each device uses local data for local training, then uploads the model to the server for aggregation, and finally the server sends the model update to the participants to achieve the learning goal. To provide a comprehensive survey and facilitate the potential research of this area, we systematically introduce the existing works of federated learning from five aspects: data partitioning, privacy mechanism, machine learning model, communication architecture and systems heterogeneity. Then, we sort out the current challenges and future research directions of federated learning. Finally, we summarize the characteristics of existing federated learning, and analyze the current practical application of federated learning.

1. Introduction

1.1. Backgrounds of federated learning

With the development of big data, the amount of data is no longer the focus of our attention. The urgent problem that needs to be solved is the privacy and security of data. The leakage of data is never a small problem, and recently the public pay growing attention to data security [1–3]. Not only individuals, collectives and society are also strengthening the protection of data security and privacy. Taking the General data Protection Regulations implemented by the European Union on May 25th, 2018 as an example, GDPR [4] aims to protect users' personal privacy and data security. It requires operators to clearly express the user agreements and cannot deceive or induce users to give up the privacy requirements. In addition, operators are prohibited from training model without the user's permission. At the same time, it allows users to delete their private data. Similarly, China's Cyber Security Law of the People's Republic of China [5] and the General principles of the Civil Law of the People's Republic of China [6], which have been implemented since 2017, also point out that network operators shall not disclose, tamper with or destroy the personal information they collect. When conducting data transactions with the third, it is necessary to ensure that the proposed contract clearly specifies the scope of the data to be traded and the obligations of data protection. The establishment of these laws and regulations poses new challenges to the traditional data processing mode of artificial intelligence to varying degrees.

In the field of artificial intelligence, data is the foundation, therefore model training cannot be performed without data. However, data often exists in the form of data islands. The direct solution to data islands is to process the data in a centralized manner. The popular data processing method is through centralized collection, unified processing, cleaning and modeling. In most cases, data is leaked during collection and processing. With the improvement of regulations, user's private information is well protected, but it is getting harder to collect data to train models. How to legally solve the problem of data islands has attracted a lot of attention and thinking of artificial intelligence. To solve the dilemma of data silos, traditional data statistics methods are already stretched in the face of various regulations. Federated learning shifts the focus of research to the problem of data islands. The traditional machine learning mostly uses the centralized method to train the machine learning model, which requires the training data to be concentrated in the same server.

* Corresponding author.
E-mail address: sxlljcxy@gmail.com (Y. Xie).

https://doi.org/10.1016/j.knosys.2021.106775

Limitations

What are the limitations?

Limitations

- The model structure must be the same in the server and in the clients.

Limitations

- The model structure must be the same in the server and in the clients.
- Even the gradients may be used to discover information from the client's data.

Limitations

- The model structure must be the same in the server and in the clients.
- Even the gradients may be used to discover information from the client's data.
- Synchronicity.

  When will the gradients be sent? What happens when we are close to a local minima, and many clients send the same gradient? When to update the client's models?...

Limitations

- The model structure must be the same in the server and in the clients.
- Even the gradients may be used to discover information from the client's data.
- Synchronicity.

  When will the gradients be sent? What happens when we are close to a local minima, and many clients send the same gradient? When to update the client's models?...
- Can we trust the clients?

Limitations

- The model structure must be the same in the server and in the clients.
- Even the gradients may be used to discover information from the client's data.
- Synchronicity.

  When will the gradients be sent? What happens when we are close to a local minima, and many clients send the same gradient? When to update the client's models?...
- Can we trust the clients?

These limitations have mitigations in the state-of-the-art. Check it out!

E.g., The gradients can be shared using a secure communication channel.

Exercises

1. Implement a simulation of a federated learning environment.
    a. Train a small model on some data.
    b. Copy this model in at least two “clients”.
    c. Use data that is private for these client to update theis models.
        i. Share the gradients with the central model.
    d. What happened? How the central model quality compares with a model that had all the data available?

References

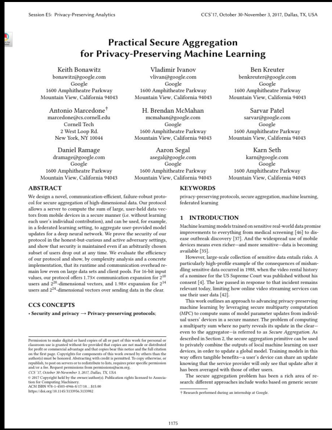
Sotomaior, Fonseca, Zangari, Britto Jr, Viegas. A Federated Learning Model for Privacy-Preserving and Cross-Domain Kidney Stone Detection in Medical Imaging. SMC 2025.

A Federated Learning Model for Privacy-Preserving and Cross-Domain Kidney Stone Detection in Medical Imaging

Lucas Sotomaior, Luiz F. B. Fonseca, Matheus A. C. Zangari
Rodrigo K. Krebs, Alceu Britto Jr, and Eduardo K. Viegas

I. Introduction

Manuscript 1614 submitted to 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Received May 3, 2025.



Session E5: Privacy-Preserving Analytics — CCS'17, October 30-November 3, 2017, Dallas, TX, USA

Practical Secure Aggregation for Privacy-Preserving Machine Learning

Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, Karn Seth

ABSTRACT

KEYWORDS

1 INTRODUCTION

CCS CONCEPTS

1175

Bonawitz, Ivanov, Kreuter, Marcedone, McMahan, Patel, … & Seth. Practical secure aggregation for privacy-preserving machine learning. ACM SIGSAC 2017.

Knowledge-Based Systems 216 (2021) 106775

Knowledge-Based Systems

A survey on federated learning

Chen Zhang, Yu Xie, Hang Bai, Bin Yu, Weihong Li, Yuan Gao

ARTICLE INFO

ABSTRACT

1. Introduction

1.1. Backgrounds of federated learning

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775.

Licence