Federated Learning: Evaluating Popular Frameworks and Developing a Sever to Mimic Real Scenarios

Jose Maria Rico Leal

A Thesis Submitted in Partial Fulfilment of the requirements for the

Degree of

Master of Science in Data Analytics



September 2024 Supervisor: Sam Weiss

# Abstract

Accurate hlhafa lhafkahf lk lahrlkah lkhlkahf alk khalfha lh lkfah lfa lahflakh jfa lahflakh falh lhalfalh lahfalh fa.

**GitHub: Wordcount: 15,000**

[Abstract 2](#_Toc172495928)

[Abbreviations 5](#_Toc172495929)

[Figures 6](#_Toc172495930)

[Tables 7](#_Toc172495931)

[1. Introduction 8](#_Toc172495932)

[1.1. Motivation 8](#_Toc172495933)

[1.2. Research Objectives 9](#_Toc172495934)

[1.3. Thesis Overview 9](#_Toc172495935)

[2. Background 9](#_Toc172495936)

[2.1. What is Federated Learning 9](#_Toc172495937)

[2.2. Federated Learning Classification Based on Client Nature 10](#_Toc172495938)

[2.2.1. Cross-device 10](#_Toc172495939)

[2.2.2. Cross-silo 11](#_Toc172495940)

[2.3. Categorization of Federated Learning 11](#_Toc172495941)

[2.3.1. Horizontal Federated Learning 11](#_Toc172495942)

[2.3.2. Vertical Federated Learning 12](#_Toc172495943)

[2.3.3. Federated Transfer Learning 12](#_Toc172495944)

[2.4. Federated Learning vs Distributed Machine Learning 13](#_Toc172495945)

[3. Literature Review 14](#_Toc172495946)

[3.1. Federated Learning Frameworks 14](#_Toc172495947)

[3.1.1. TensorFlow Federated, Flower, EasyFL, IBM Federated Learning and FLGo 14](#_Toc172495948)

[3.1.2. OpenFL, PySyft, MLbase, FATE-LLM, SecureBoost, Personalised Federated Learning and Flint 16](#_Toc172495949)

[3.1.3. Federated Learning Algorithms 16](#_Toc172495950)

[3.2. Distributed Machine Learning 16](#_Toc172495951)

[3.2.1. Theme BB 16](#_Toc172495952)

[3.3. Federated Learning Server Implementation 17](#_Toc172495953)

[3.4. Federated Learning Datasets 17](#_Toc172495954)

[3.4.1. Theme Theme DD 17](#_Toc172495955)

[3.5. Real World Federated Learning Scenarios 17](#_Toc172495956)

[3.6. Conclusion 17](#_Toc172495957)

[5. State of the Art 17](#_Toc172495958)

[6. Research Methodology 17](#_Toc172495959)

[6.1. Methodology Frameworks 17](#_Toc172495960)

[6.2. Project Management Framework 17](#_Toc172495961)

[6.3. Business Requirements 18](#_Toc172495962)

[6.4. Data Synthetic and Medical 18](#_Toc172495963)

[6.5. Data Preparation 18](#_Toc172495964)

[6.6. Limitations and Ethical Considerations 18](#_Toc172495965)

[7. Data Analysis 18](#_Toc172495966)

[7.1. Data Collection 18](#_Toc172495967)

[7.2. Initial Data Extraction and Exploration 18](#_Toc172495968)

[7.2.1. Data Elaboration 18](#_Toc172495969)

[7.2.2. Data Elaboration 2 18](#_Toc172495970)

[7.3. Conclusion 18](#_Toc172495971)

[8. Data Preparation 18](#_Toc172495972)

[8.1. Data Cleaning 18](#_Toc172495973)

[8.1.1. AAAAA 18](#_Toc172495974)

[9. Experimentation 18](#_Toc172495975)

[9.1. BBBB Option 19](#_Toc172495976)

[9.2. CCCC Option 19](#_Toc172495977)

[9.3. Aaaaaa Option 19](#_Toc172495978)

[9.4. BBBB Option 19](#_Toc172495979)

[9.4.1. CCCCC Option 19](#_Toc172495980)

[10. Validation of the Experimentation’s Results 19](#_Toc172495981)

[10.1. Validation BBBB 19](#_Toc172495982)

[11. Research Conclusion: Summary, Limitations, and Recommendations 19](#_Toc172495983)

[References: 19](#_Toc172495984)

# Abbreviations

|  |  |
| --- | --- |
| FL | Example of Abbreviation |
| AEA | Another Example Abbreviation |

# Figures

Figure 1.1. Illustration of a single FL communication round in FL for mobile keyboard prediction.

Figure 2.2.1. Cross-device scenario.

Figure 2.2.2. Cross-silo scenario.

Figure 2.3.1. Horizontal Federated Learning.

Figure 2.3.2. Vertical Federated Learning.

Figure 2.3.3. Federated Transfer Learning.

# Tables

Table 2.3.3. Differences between Federated Learning and Distributed Machine Learning

.

# Introduction

## Motivation

Within the domain of Data Analytics (DA), there is an important field known as Machine Learning (ML), which is embedded in everyday people's lives. A significant topic within this field is Federated Learning (FL). FL occurs when different devices collaborate to build a common model without exchanging their data; instead, the data remains on the devices, and only model updates are sent to a central server where aggregation occurs.

A good example for FL in everyday people’s lives is when phone users are typing a message, and the keyboard predicts words to complete the sentence, a model developed by Google engineers (Hard *et al.*, 2019).

A diagram of a phone system

Description automatically generated  
Figure 1.1. Illustration of a single FL communication round in FL for mobile keyboard prediction.

A different example where FL is helping society move forward is in the healthcare sector. Patients are using wearable devices to track their movements and help doctors monitor their health conditions (Arikumar *et al.*, 2022).

FL is growing rapidly and is helping the technological and medical sectors build robust machine learning models. This significant development motivates the focus of this thesis.

## Research Objectives

The primary purpose of this project is to experiment with FL frameworks to evaluate their implementability and develop a functional FL server. Therefore, the research objectives are:

* **To evaluate the implementability of existing FL frameworks.** This section experiments with popular FL frameworks, such as *TensorFlow Federated* (TFF)*, Flower, EasyFL, IBM Federated* (IBMF)and *FLGo*, by examining their architecture and their applicability to real-world FL scenarios. This evaluation serves as the starting point for the primary research.
* **To develop a FL server**. A practical example of a web Flask FL server will be built, featuring two scenarios: technological and pharmaceutical. Each scenario will run separately, connecting five clients. Each scenario will have its own data, synthetic data for the technological scenario and images for the pharmaceutical scenario. These configurations will be trained, validated, and tested using Neural Networks (NN) that will classify binary outputs. This experiment aims to bridge the gap between popular FL frameworks and real-world FL applications.
* **To compare FL frameworks and the FL server.** This includes discussing the pros and cons each method brings to key industries such as the technological and pharmaceutical sectors. The comparison will focus on the following aspects ease of implementation, scalability, data privacy, and model performance.

## Thesis Overview

# Background

## What is Federated Learning

This concept was introduced in 2016 by Google engineers (McMahan et al., 2016). FL is a shared model that is trained across multiple devices, often referred to as clients. Each client trains its own local model and sends the updates to a central server, where the updates are aggregated to improve the global model. It is important to note that clients do not exchange their data; the data remains private for each client. This principle drove the design of FL, following the concepts of *focused collection* or *data minimization*, which were introduced by the White House in 2013. The intent is to prevent personal data from being sent over the network and potentially being stolen or manipulated by malicious third parties.

## Federated Learning Classification Based on Client Nature

Depending on the nature of client FL can be classified in two types cross-device and cross-silo (Yang *et al.*, 2021).

### Cross-device

The clients for cross-device can be mobile devices, edge devices, Internet of Things (IoT) devices, smartphones, tablets, wearables, etc. Figure 2.2.1. illustrates this scenario.  
The characteristics are, the high number of participants it can be thousands to millions of devices, it may have limited processing power and battery life, datasets tend to be small and network bandwidth may be limited. Devices may also connect and disconnect intermittently.

A diagram of a computer network

Description automatically generated  
Figure 2.2.1. Cross-device scenario.

### Cross-silo

In this scenario, clients can be organizations or institutions such as hospitals, banks, and companies using large data centres. Figure 2.2.2 illustrates a cross-silo ecosystem. Some differences compared to the cross-device scenario include, clients are no longer small devices, there are fewer clients, clients have high computational power and large datasets, and the network is reliable with stable communication.

A diagram of a server

Description automatically generated  
Figure 2.2.2. Cross-silo scenario.

## Categorization of Federated Learning

FL can be categorized according to the distribution of the data held by the clients participating in the modelling. These categories help to understand the different methodologies and use cases for FL (Yang *et al.*, 2019).

### Horizontal Federated Learning

Horizontal Federated Learning (HFL) or sample-based FL, occurs when different clients have datasets that share the same feature space but differ in the samples they hold (see Figure 2.3.1). A practical example of HFL is when two hospitals in different regions each have patient records with the same features (e.g., age, height, weight, diagnosis) but for different patients. These hospitals can collaborate to train a model to predict disease outcomes without sharing patient data.

A diagram of a dataset

Description automatically generated with medium confidence  
Figure 2.3.1. Horizontal Federated Learning.

### Vertical Federated Learning

Vertical Federated Learning (VFL), or feature-based FL, occurs when different clients have datasets that share the same sample IDs but differ in the feature space. An example of this is a bank and an e-commerce company that have data on the same set of customers. The bank has financial information such as credit scores and loan histories, while the e-commerce company has purchase behaviours and browsing records. By combining their data, they can build a model to predict customer credit without sharing raw data.

A diagram of a dataset

Description automatically generated with medium confidence  
Figure 2.3.2. Vertical Federated Learning.

### Federated Transfer Learning

Federated Transfer Learning (FTL) is applicable when the datasets of different clients differ in both samples and features, and the overlap between both is minimal. To clarify, a good example is a European pharmaceutical company and a healthcare research institution in China collaborating using FTL. The pharmaceutical company has drug efficacy data, while the healthcare institution has patient health records. Even though they have different types of data, it is possible to train a global model to predict drug effectiveness on certain health conditions.

A diagram of dataset and dataset

Description automatically generated  
Figure 2.3.3. Federated Transfer Learning.

## Federated Learning vs Distributed Machine Learning

Terms like FL and Distributed Machine Learning (DML) can create confusion due to their similarities. The main difference lies in the training process: in FL, there is a central server that aggregates updates sent by the clients, whereas in DML, there is no central server; instead, data is spread across different nodes and computations are shared among these nodes (Li *et al.*, 2020). Table 2.4 will help clarify the intricacies of each concept:

A table of information

Description automatically generated  
Table 2.3.3. Differences between Federated Learning and Distributed Machine Learning

# Literature Review

This literature review comprises thirty sources, structured using methodology as the organizing principle. Throughout the research, five themes have been identified as follows: Federated Learning frameworks, Distributed Machine Learning, Federated Learning implementation (grey materials), commonly used datasets, and real-world Federated Learning scenarios.

The literature review is key to this proposal as it sheds light on many aspects of the FL. Starting with the frameworks, a total of fourteen sources have been reviewed. The criteria for selecting these sources were based on publication date and whether the framework is widely used. All included sources are recent, and the frameworks are extensively used by both industry and researchers. Some of the aspects considered include how user-friendly these frameworks are, which machine learning algorithms they use, and how the FL is carried out (centralized, decentralized, or vertical).

For the second section, a total of five sources have been reviewed. This section aligns with the second research objective and introduces the idea that a potential FL system can be built within a distributed file system (Hadoop). Datasets stored across different directories will emulate the clients. Using MapReduce or Spark as a central server, datasets can be trained and aggregated into a global model.

In the third section, two YouTube videos serve as grey materials. These demonstrate in a straightforward manner how to deploy a FL system, maintaining client privacy with ease while training the global model.

There are many FL datasets primarily used for research purposes, while those used within the industry are kept private for obvious reasons. The purpose of this section is to identify the most popular FL datasets based on their usage in FL frameworks. These datasets will also serve as part of the experimentation in this project. A total of four sources have been reviewed in this section.

The last section, which has reviewed a total of five sources, has been instrumental to the third research objective. It enhances understanding of how real-world industries deploy their FL systems.

## Federated Learning Frameworks

### TensorFlow Federated, Flower, EasyFL, IBM Federated Learning and FLGo

Solanki et, al., (2022) delve into how TFF, an open-source framework, is utilized for machine learning on decentralized data. It has been designed for research and experimentation. Some of the key features are TFF enables FL through low-latency models with less power consumption. The framework uses two layers, the FL learning Application Programming Interface (API) and the federated core (FC) API. The FL API allows developers to implement training and evaluation on existing TensorFlow models through a high-level interface. The FC API integrates TensorFlow with distributed communication operators focusing on computations across distributed systems like mobile phones, tablets, and sensors. Comparing TFF to other frameworks, it offers a unique well integrated structure others do not provide this level of integration. TFF allows experimenting with new algorithms is not tied to predefined algorithms. A different study that showcases a fairness-aware federated learning algorithm designed to group fairness while computing (Salazar et al., 2023). It incorporates a fairness-aware momentum to compute the global model by considering client model fairness level. Some of the key points are fairness-aware aggregation, momentum term, group fairness and real-world experiments. Compared to existing FL frameworks like *FedAvg* and *FedMom*, *FAIR-FATE* utilizes a momentum-based approach to address fairness specifically. Unlike some methods that require local debiasing strategies for each client, FAIR-FATE operates without that requirement. Other frameworks have attempted to aggregate models based on fairness, but FAIR-FATE seems the one achieving this. As a conclusion, FAIR-FATE is an effective approach for increasing fairness in federated learning models, showing significant improvements in various heterogeneous scenarios.

Also, Beutel et al., (2021) presented a user-friendly framework, *Flower*, an open-source framework designed to enhance the capabilities and practical implementation of FL, across different environments. *Flower* addresses the challenges associated with the deployment and scalability of FL. Flower architecture allows simulations in real-world scenarios, making it a great tool for federated learning scenarios. The framework is agnostic supporting different machine frameworks like *TensorFlow* and *PyTorch*. Comparing Flower to other federating learning frameworks such as TFF, PySyft, FedScale, and *LEAF*. Flower stands out by supporting actual deployment on real devices rather than just simulation, unlike *TFF* and *LEAF*, which mainly focus on simulated environments.

Zhuang et al., (2022) developed *EasyFL* designed a low code platform to assist beginners and researchers to experiment and prototype FL artefacts. It offers practical functionalities such as handling heterogeneity, simulation, comprehensive tracking, optimization of distributed training, and seamless deployment. While numerous FL platforms have been developed by institutions and companies, these are difficult to implement. TFF is deployable but does not optimize distributed training. On the other hand, *FATE* supports deployment but is not user-friendly, presenting high entry barriers. In contrast, *EasyFL* is user-friendly and facilitates efficient experimentation along with seamless deployment. It also supports diverse training methods, including standalone, distributed, and remote.

A different FL framework developed by the industry is IBM FL (Ludwig et al., 2020) This framework is designed to facilitate the implementation of federated learning across diverse enterprise environments. It helps users to model without centralizing training data, addressing the key issues of privacy. It supports the integration of Deep Neural Networks and traditional machine learning methods. It also provides tools for design and deployment of federated jobs minimizing the learning curve. IBM FL is different from other existing FL frameworks by its focus on enterprise needs, including secure deployment, failure tolerance, and rapid model specification. Also, when compared to other existing frameworks IBM’s solution is tailored for multi-cloud or hybrid cloud environments where data privacy is critical. It supports both federated learning systems using a central aggregator and more decentralized models. In conclusion, it provides an effective bridge between traditional centralized data processing and the emerging needs of decentralized, privacy preserving machine learning applications

Wang et al. (2023) developed *FLGo* a platform designed to streamline the process of cross-application FL research and enhance shareability among developers. It is a lightweight FL framework aiming to be a customizable solution to suit different applications and data heterogeneity. *FLGo* addresses the gap that exists in current FL frameworks which often make the FL deployment very complex. Some of the key Features, are benchmarks and algorithms, customization, experimental tools, and high degree of shareability. Compared to other frameworks it stands out in, system heterogeneity, high-level API, multi-architecture support, asynchronous operations and customization and flexibility. As a conclusion *FLGo* has been developed with the intention of making FL more accessible to a broader range of developers by simplifying customization and enhancing its shareability. It also aims to bridge the existing gap with conventional machine learning and FL.

### OpenFL, PySyft, MLbase, FATE-LLM, SecureBoost, Personalised Federated Learning and Flint

### Federated Learning Algorithms

## Distributed Machine Learning

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

### Theme BB

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Federated Learning Server Implementation

## Federated Learning Datasets

### Theme Theme DD

## Real World Federated Learning Scenarios

## Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Identification of gaps

# State of the Art

# Research Methodology

## Methodology Frameworks

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Project Management Framework

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Business Requirements

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

## Data Synthetic and Medical

## Data Preparation

## Limitations and Ethical Considerations

# Data Analysis

## Data Collection

## Initial Data Extraction and Exploration

### Data Elaboration

### Data Elaboration 2

## Conclusion

# Data Preparation

## Data Cleaning

### AAAAA

# Experimentation

## BBBB Option

## CCCC Option

## Aaaaaa Option

## BBBB Option

### CCCCC Option

# Validation of the Experimentation’s Results

## Validation BBBB

# Research Conclusion: Summary, Limitations, and Recommendations

# References:

Arikumar, K.S., Prathiba, S.B., Alazab, M., Gadekallu, T.R., Pandya, S., Khan, J.M. and Moorthy, R.S. (2022) ‘FL-PMI: Federated Learning-Based Person Movement Identification through Wearable Devices in Smart Healthcare Systems’, *Sensors*, 22(4), p. 1377. Available at: https://doi.org/10.3390/s22041377.

Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C. and Ramage, D. (2019) ‘Federated Learning for Mobile Keyboard Prediction’. arXiv. Available at: http://arxiv.org/abs/1811.03604 (Accessed: 18 July 2024).

Li, T., Sahu, A.K., Talwalkar, A. and Smith, V. (2020) ‘Federated Learning: Challenges, Methods, and Future Directions’, *IEEE Signal Processing Magazine*, 37(3), pp. 50–60. Available at: https://doi.org/10.1109/MSP.2020.2975749.

Yang, Q., Liu, Y., Chen, T. and Tong, Y. (2019) ‘Federated Machine Learning: Concept and Applications’. arXiv. Available at: http://arxiv.org/abs/1902.04885 (Accessed: 20 July 2024).

Yang, X., Feng, Y., Fang, W., Shao, J., Tang, X., Xia, S.-T. and Lu, R. (2021) ‘An Accuracy-Lossless Perturbation Method for Defending Privacy Attacks in Federated Learning’. arXiv. Available at: http://arxiv.org/abs/2002.09843 (Accessed: 20 July 2024).