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](#_Toc172303619)

[1. Introduction 8](#_Toc172303620)

[1.1. Motivation 8](#_Toc172303621)

[1.2. Research Objectives 9](#_Toc172303622)

[1.3. Thesis Overview 9](#_Toc172303623)

[2. Background 9](#_Toc172303624)

[2.1. What is Federated Learning 9](#_Toc172303625)

[2.2. History of Federated Learning 9](#_Toc172303626)

[2.3. Types of Federated Learning 9](#_Toc172303627)

[2.4. Federated Learning vs Distributed training 9](#_Toc172303628)

[3. Literature Review 9](#_Toc172303629)

[3.1. Theme A Federated Learning Frameworks 9](#_Toc172303630)

[3.1.1. Theme AA 9](#_Toc172303631)

[3.2. Theme B Theme 9](#_Toc172303632)

[3.2.1. Theme BB 9](#_Toc172303633)

[3.3. Theme Theme Theme Cc 10](#_Toc172303634)

[3.4. Theme Theme D 10](#_Toc172303635)

[3.4.1. Theme Theme DD 10](#_Toc172303636)

[3.5. Theme Ed 10](#_Toc172303637)

[3.6. Conclusion 10](#_Toc172303638)

[5. State of the Art 10](#_Toc172303639)

[6. Research Methodology 10](#_Toc172303640)

[6.1. Methodology Frameworks 10](#_Toc172303641)

[6.2. Project Management Framework 10](#_Toc172303642)

[6.3. Business Requirements 11](#_Toc172303643)

[6.4. Data Synthetic and Medical 11](#_Toc172303644)

[6.5. Data Preparation 11](#_Toc172303645)

[6.6. Limitations and Ethical Considerations 11](#_Toc172303646)

[7. Data Analysis 11](#_Toc172303647)

[7.1. Data Collection 11](#_Toc172303648)

[7.2. Initial Data Extraction and Exploration 11](#_Toc172303649)

[7.2.1. Data Elaboration 11](#_Toc172303650)

[7.2.2. Data Elaboration 2 11](#_Toc172303651)

[7.3. Conclusion 11](#_Toc172303652)

[8. Data Preparation 11](#_Toc172303653)

[8.1. Data Cleaning 11](#_Toc172303654)

[8.1.1. AAAAA 11](#_Toc172303655)

[9. Experimentation 11](#_Toc172303656)

[9.1. BBBB Option 12](#_Toc172303657)

[9.2. CCCC Option 12](#_Toc172303658)

[9.3. Aaaaaa Option 12](#_Toc172303659)

[9.4. BBBB Option 12](#_Toc172303660)

[9.4.1. CCCCC Option 12](#_Toc172303661)

[10. Validation of the Experimentation’s Results 12](#_Toc172303662)

[10.1. Validation BBBB 12](#_Toc172303663)

[11. Research Conclusion: Summary, Limitations, and Recommendations 12](#_Toc172303664)

[References: 12](#_Toc172303665)

**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.

**Tables**

Table 1.1. XXXXXX.

Table 5.1. XXXXXX.

Table 5.2. XXXXXX.

.

# 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

## History of Federated Learning

## Types of Federated Learning

## Federated Learning vs Distributed training

# Literature Review

## Theme A Federated Learning Frameworks

### Theme AA

## Theme B Theme

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.

## Theme Theme Theme Cc

## Theme Theme D

### Theme Theme DD

## Theme Ed

## 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).