2

Pioneering the Hybridization of Federated Learning in Human Activity Recognition

Alfonso Esposito[1], Yasamin Moghbelan[1], Ivan Zyrianoff[1], Leonardo Ciabattini[1], Federico Montori[1,2], and Marco Di Felice[1,2]

[1]University of Bologna, Italy
[2]University of Bologna, Advanced Research Center for Electronic Systems, Italy

Abstract

The Internet of Things (IoT) nowadays greatly benefits from Artificial Intelligence (AI) algorithms implemented in the edge, because of their efficiency and the reduction of costs that they imply. The advent of Federated Learning (FL) has made possible the combination of the advantages of edge-AI, among which the privacy of users, as their data is not shared with the cloud, with the collective intelligence. However, FL is known to have worse performance compared to its centralized counterpart, which may not be tolerable in certain cases. In this paper, we propose a hybrid framework for FL, imagining a number of clients that are willing to share part of their data. We envisioned two types of Hybridization: vertical and horizontal. The goal of this paper is to assess whether a small hybridization can bring advantages to the overall performance of the whole FL procedure in terms of classification accuracy.

Keywords: Internet of Things (IoT), deep learning (DL), federated learning (FL), human activity recognition (HAR), performance evaluation.

2.1 Introduction and Background

The growth of the Internet of Things (IoT) has enabled novel monitoring systems capable of gathering data from heterogeneous and pervasive devices, supporting smart city, healthcare and industrial domains. This data

is processed using advanced Deep Learning (DL) techniques for system state forecasting, contributing to the integrated paradigm of the Artificial Intelligence of Things (AIoT) [1]. However, it is well known that advanced DL techniques, particularly those used in computer vision applications, require large datasets with adequate number of instances for each system state or class to be predicted. In many cases, building reliable DL models requires aggregating data from multiple heterogeneous users or organizations performing decentralized data collection for a common task. This is the case, for instance, of most of Human Activity Recognition (HAR) applications that utilize DL models trained from wearable or camera-based IoT data gathered from multiple volunteers or via crowdsensing techniques [2].

State-of-the-art AIoT systems often rely on cloud-based centralized architectures, which facilitate the aggregation of IoT data from diverse clients. While these solutions enable easy deployment and scalable computational and storage resources, they also raise significant privacy concerns due to the inherent risks associated with data sharing. Federated Learning (FL), first proposed in 2018 [3], has emerged as a privacy-preserving alternative to centralized machine learning, enabling cooperative training in a distributed environment. In a typical FL setup, multiple clients independently train models on their private datasets and only share the trained weights with a central server, which aggregates these weights and sends the updated model back to the clients. This approach ensures that no raw data is exchanged among clients, although concerns about the trustworthiness of the centralized server remain [4]. Moreover, variations in data quality and quantity across clients can create challenges in ensuring fair evaluation of each client's contribution and achieving a high-quality global model [5]. To address the issue of non-i.i.d. (non-independent and identically distributed) data across clients, various solutions have been proposed, such as clustering clients with similar data distributions or adopting weight-based model aggregation techniques [6].

In this paper, we aim to bridge the gap between centralized and Federated Learning (FL) methodologies, in order to address distributed IoT use cases where privacy requirements vary from client to client. For instance, some clients may be willing to share raw data, while others may not, due to differences in client nature (e.g., public vs. private organizations), varying perceptions of privacy—often shaped by social factors [7]—or monetization strategies, where some clients are incentivized to sell their data. We refer to this scenario as *hybrid FL* and describe this *hybridization* approach for data gathering and model building at the server side, which offers a novel interpretation compared to other studies [4]. Specifically, this paper

explores two types of FL hybridization: *vertical* and *horizontal*. In the vertical hybridization, a portion of clients shares raw data with the centralized server, while others only share deep learning model coefficients generated from local training. In the horizontal case, all clients share a variable portion of their raw data (e.g., corresponding to the amount for which they have been compensated) in addition to the DL model coefficients trained on the complementary data. We evaluate the performance of both strategies using two benchmark datasets (related to Human Activity Recognition and image classification) and analyze the impact of different levels of hybridization compared to pure centralized and FL-based approaches. Our results demonstrate that hybridization can be a powerful tool for improving the accuracy of federated systems, even when applied slightly.

To summarize, the key contributions of our paper are the following:

- Introduction of hybrid FL as a new strategy for privacy-adaptive learning in IoT scenarios.
- Proposal of two versions of hybrid FL, respectively horizontal and vertical, based on distinct methods of integrating raw data and deep learning weights at the server.
- Evaluation of proposed strategies across varying degrees of hybridization, using two widely adopted benchmark datasets in the DL community.

The rest of the paper is structured as follows. Section 2 introduces the revised FL architecture supporting the vertical and horizontal hybridization. Section 3 describes the evaluation methodology, datasets and metrics. Section 4 presents some evaluation results. Section 5 concludes the work and discusses future research steps.

2.2 Hybrid FL Architecture

We consider the scenario depicted in Figure 2.1, composed of two main actors: N distributed clients and the server. Each client c_i possesses its own dataset D_i gathered through its local sensors, and a DL network topology, denoted as m in the following. In a classical FL application, each client c_i performs local training rounds of model m on D_i and then shares the list of weights W_i with the server: the latter is responsible for aggregating the weights, for instance through the FedAvg [3] algorithm, and sending the updated values back to the clients.

In the proposed hybrid FL architecture, the server performs additional storage and computational tasks, basically behaving as a special client device.

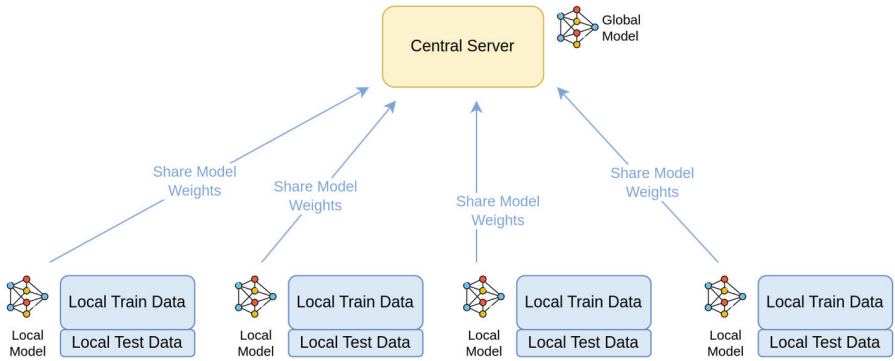

Figure 2.1 Typical Federated Learning Architecture

Indeed, it possesses a local dataset D_s that is built from clients' datasets or portions of them. More formally, we have that $D_s = \bigcup_{i=0}^{n} D_i^*$ where $n \leq$N and $D_i^* \subseteq D_i$. The server trains the m model on D_s getting a local version of weights W_S that is later averaged with the weights W_i received by clients at each round. We denoted this process as **FL hybridization**, distinguishing between two modes for creating the local dataset D_s:

- **Vertical Hybrid FL**. In such case, we have that: $n < N$ and $D_i^* = D_i$. In other words, only a subset of the client nodes shares its own raw data with the server. This setup may model two different use-cases: (*i*) only n clients have been compensated with monetary rewards to share their data and/or (*ii*) n clients do not consider their local datasets privacy-sensitive. We indicate with $r_v = \frac{n}{N}$ the rate of vertical hybridization. Clearly, for $n = N$ and $r_v = 1$, the system is equivalent to centralized learning. Vice versa, for $n = 0$ and $r_v = 0$, the system is equivalent to a pure FL approach. We investigate the impact of varying r_v configurations in the overall DL performance in Section 4. We show in Figure 2.2 an overview on the conceptual architecture of Vertical Hybrid FL.
- **Horizontal Hybrid FL**. In such case, we have that: $n = N$ and $D_i^* \subset D_i$. In other words, all clients share only a portion of their local datasets with the server. This setup may model a practical use-case where each client shares with the server an amount of raw data proportionally to the monetary reward it received. We indicate with $r_h = mean_{0 \leq i < N} \left(\frac{|D_i^*|}{|D_i|} \right)$ the rate of vertical hybridization. Clearly, for $D_i^* = D_i$, the system becomes a centralized learning. Vice versa, if $D_i^* = \emptyset \; \forall c_i$ the system is

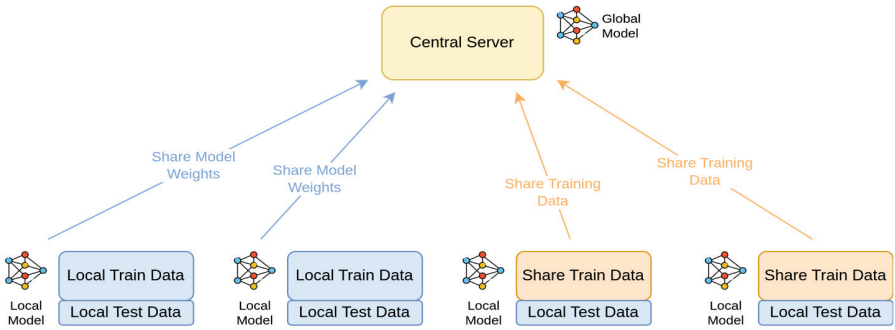

Figure 2.2 Vertical Hybrid Federated Learning Architecture

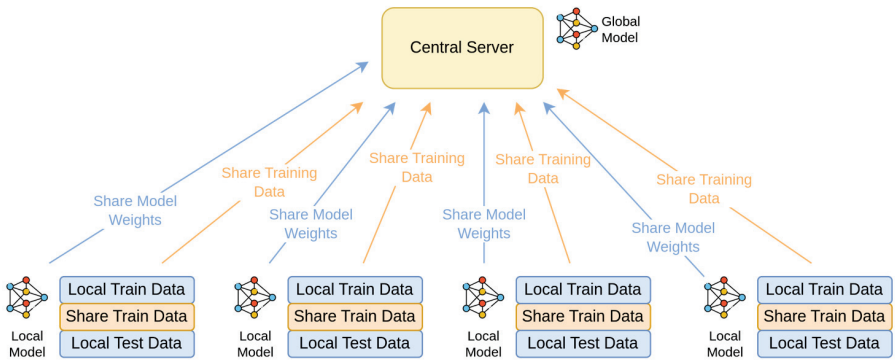

Figure 2.3 Horizontal Hybrid Federated Learning Architecture

equivalent to a pure FL approach. We investigate the impact of varying r_h configurations in the overall DL performance in Section 4. We show in Figure 2.3 an overview on the conceptual architecture of Vertical Hybrid FL.

We further highlight that the hybridization rates determine the amount of required privacy preservation. The maximum value (1) corresponds to when all clients share their local data with the server. Vice versa, the minimum value (0) forbids any transmission of raw data outside the clients' devices.

2.3 Evaluation Methodology and Metrics

In this section, we describe the experiments that we performed to investigate the performances of the two hybrid approaches explained in the previous

section. First, we will describe the methodology used to test the effectiveness of HFL across two datasets: FEMNIST and UCI HAR.

The widely recognized University of California Irvine (UCI) HAR dataset [8] was created using data from smartphone accelerometer and gyroscope sensors, which were used to classify six types of human activities. It was gathered from 30 volunteers, each carrying a smartphone while performing six distinct activities: walking, walking upstairs, walking downstairs, sitting, standing, and lying down. The dataset consists of time-series data captured across the three axes of both sensors, along with the corresponding activity labels. It has been extensively used in research for developing and assessing HAR models using machine learning techniques. Each record in the dataset contains a vector of 561 features, derived from both time and frequency domain calculations.

The FEMNIST dataset [9] is an adaptation of the extended version of the MNIST dataset that has been modified to be suitable for FL tasks. The MNIST dataset contains 28 by 28 pixels images of handwritten digits and characters (62 classes in total), and the goal of the DL model is to guess the actual character represented. The FEMNIST dataset groups the elements on top of the user that actually performed the handwriting, producing a number of sub-datasets each of them with a different style of writing. The number of users of the FEMNIST dataset is 3500, however, for the purpose of our experiment, we considered only 30 users, in order to make experiments comparable between the two datasets.

For UCI HAR we employed, as a base local model, a simple feed-forward neural network, while for FEMNIST we adopted a convolutional network. Each of the local sub-datasets is split into training and test set using a stratified split with a 70%-30% ratio.

We performed federated classification experiments by employing 6 epochs and 20 rounds of federation, recording the accuracy score at the end of the last round.

We tested both vertical and horizontal hybridization, by setting alternately r_v and r_h to values spanning from 0% to 100% with a 10% step.

The experiments were implemented in Python using the Flower framework (https://flower.readthedocs.io/en/latest/). Since Flower does not support the implementation of hybridization, we adopted the two following methods to simulate the two hybridization methods (as Figure 2.4 suggests):

- Vertical Hybridization was simulated by aggregating all the clients that share their whole dataset instead of the weights into a single client.

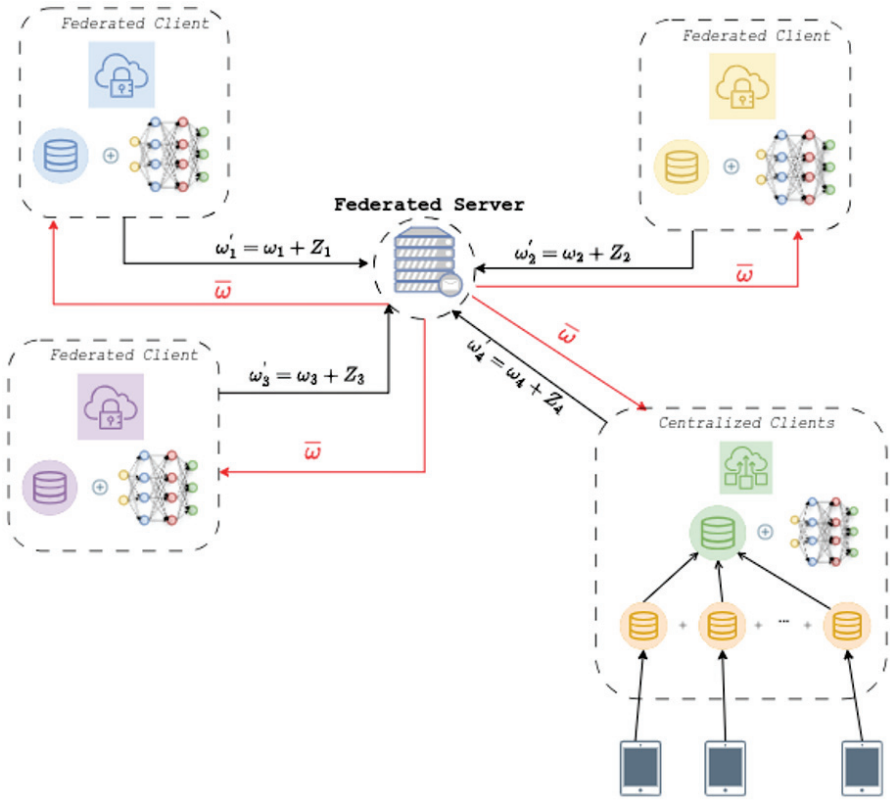

Figure 2.4 Implementation of the Vertical Hybridization in Flower

- Horizontal Hybridization was simulated by extracting from each client the portion of the training dataset that they aim to share and assigning it to a new "sink" client.

Each experiment was then repeated 20 times, by randomizing the clients or the dataset portions to be shared. This ensures scientific rigor and smooths out certain corner situations that may arise.

2.4 Evaluation Results

This section presents the outcome of the experiments presented in the previous section. The results are aimed at evaluating the HFL approaches on the datasets. We investigate how different levels of data sharing in vertical

and horizontal hybridization affect model performance. We first discuss the results for the UCI HAR dataset, followed by FEMNIST, to uncover any dataset-specific trends and performance differences. A general overview of the results shows, as expected, an overall improvement in model performance as the degree of hybridization increases, which is notable for both hybrid approaches. The UCI HAR dataset performed best. Even with the relatively small model size, it consistently achieved excellent results, maintaining accuracy above 90% and reaching almost 95% in experiments with higher levels of hybridisation. This is shown in Figures 2.5 and 2.6, where the increase in model performance as the level of hybridisation increases is evident, with an increase of 2 percentage points already at a low level of horizontal hybridization (20%).

The FEMNIST is the dataset where the performance improvements from sharing data is most noticeable. As we can observe in the Figures 2.7 and 2.8, the sharing of a small number of data points could lead to a significant improvement in performance. Specifically, sharing 10% of the data led to an improvement of approximately 5% in accuracy, while sharing 20% led to an additional improvement of approximately 2/3, resulting in a final performance of 88%. This is comparable to the 90% accuracy obtained through centralized training. Beyond a data sharing rate of 30%, the improvements

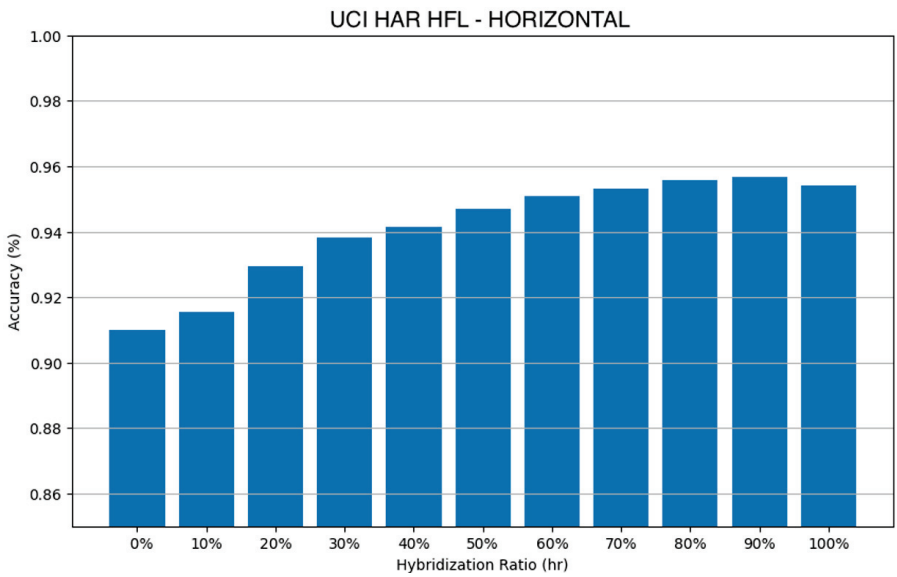

Figure 2.5 Horizontal Hybridization Results for UCI HAR

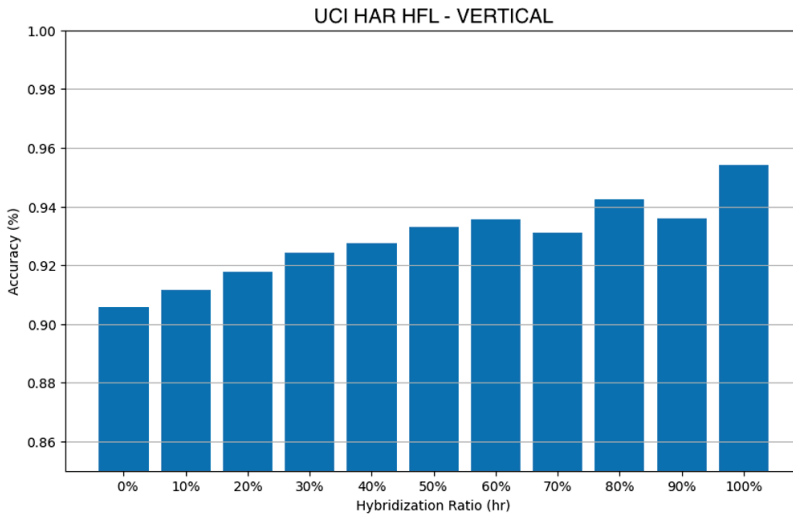

Figure 2.6 Vertical Hybridization Results for UCI HAR

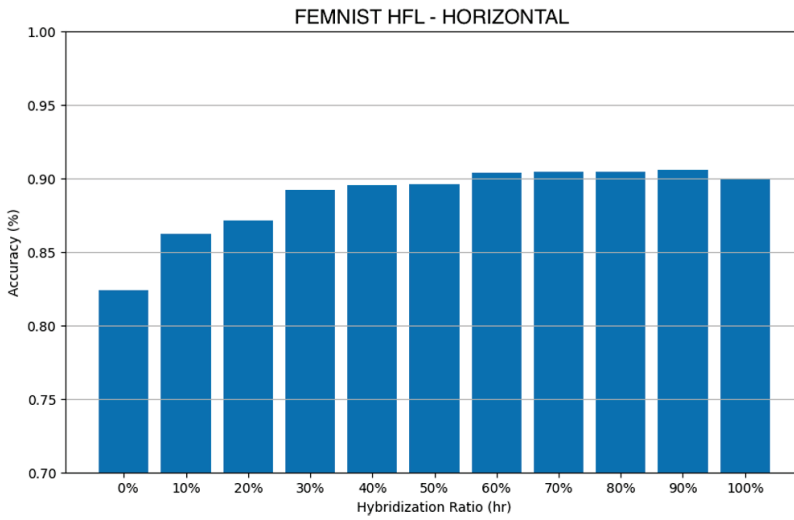

Figure 2.7 Horizontal Hybridization Results for FEMNIST

obtained are increasingly marginal, with a maximum of 1%. This suggests that a data sharing rate of 20% represents an optimal balance between performance and data sharing.

The results demonstrate that the sharing of a portion of the dataset has a significant positive impact on the model's performance. This effect was

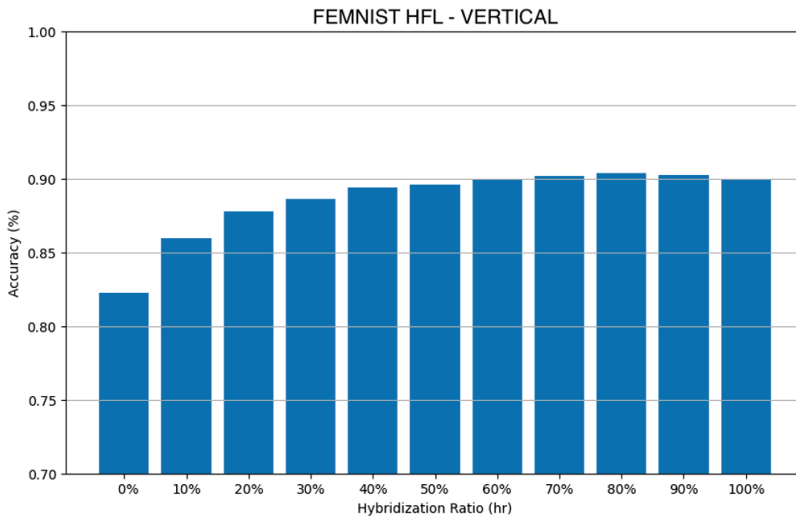

Figure 2.8 Vertical Hybridization Results for FEMNIST

observed across two distinct datasets, UCI HAR and FEMNIST, indicating that the improvement is not specific to any problem or model. The performance enhancement was especially evident in the case of the FEMNIST dataset.

2.5 Conclusion and Future Works

In this paper we examined the effects of hybridization for Federated Learning scenarios. We specifically directed our research towards Human Activity Recognition, imagining scenarios in which certain clients would be willing to share (part of) their data with the central server for a reward, penalizing their privacy to an extent. Results showed that a minimal amount of hybridization does provide an increase in the performance. The extent to which the privacy of the users is compromised by this is a future work. We aim to study how to select carefully data in a way in which the privacy is minimally affected, as well as to blend the two hybridization techniques, to select the best configuration.

Acknowledgements

This research is funded by the “Progetto Casa delle Tecnologie Emergenti” - Comune di Bologna - PSC MISE 2014-2020.

References

[1] K. S. Awaisi, Q. Ye and S. Sampalli, "A Survey of Industrial AIoT: Opportunities, Challenges, and Directions," in IEEE Access, vol. 12, pp. 96946-96996, 2024. https://doi.org/10.1109/ACCESS.2024.3426279

[2] S. Ankalaki, "Simple to Complex, Single to Concurrent Sensor-Based Human Activity Recognition: Perception and Open Challenges," in *IEEE Access*, vol. 12, pp. 93450-93486, 2024. https://doi.org/10.1109/ACCESS.2024.3422831

[3] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, B. Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", arXvi, https://arxiv.org/abs/1602.05629

[4] J. Cui, H. Zhu, H. Deng, Z. Chen, and D. Liu, "Fearh: Federated machine learning with anonymous random hybridization on electronic medical records," Journal of Biomedical Informatics, vol. 117, p. 103735, 2021. https://doi.org/10.1016/j.jbi.2021.103735

[5] M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, "Reviewing federated learning aggregation algorithms; strategies, contributions, limitations and future perspectives," Electronics, vol. 12, no. 10, p. 2287, May 2023. https://doi.org/10.3390/electronics12102287

[6] H. Lee and D. Seo, "FedLC: Optimizing Federated Learning in Non-IID Data via Label-Wise Clustering," in *IEEE Access*, vol. 11, pp. 42082-42095, 2023. https://doi.org/10.1109/ACCESS.2023.3271517

[7] D. Ibdah, N. Lachtar, S. M. Raparthi and A. Bacha, "Why Should I Read the Privacy Policy, I Just Need the Service": A Study on Attitudes and Perceptions Toward Privacy Policies," in IEEE Access, vol. 9, pp. 166465-166487, 2021. https://doi.org/10.1109/ACCESS.2021.3130086

[8] Anguita, Davide, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge Luis Reyes-Ortiz. "A public domain dataset for human activity recognition using smartphones." In Esann, vol. 3, p. 3. 2013. https://www.esann.org/sites/default/files/proceedings/legacy/es2013-84.pdf

[9] Caldas, Sebastian, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečný, H. Brendan McMahan, Virginia Smith, and Ameet Talwalkar. "Leaf: A benchmark for federated settings." arXiv preprint arXiv:1812.01097. https://arxiv.org/abs/1812.01097