

Article

Not peer-reviewed version

Robust Client Selection Strategy Using an Improved Federated Random High Local Performance Algorithm to Address High Non-IID Challenges

[Pramote Sittijuk](#) , [Narin Petrot](#) , [Kreangsak Tamee](#) *

Posted Date: 4 February 2025

doi: 10.20944/preprints202502.0125.v1

Keywords: Client Selection; Improved Fed-RHLP algorithm; Non-IID issue; Model Aggregation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Robust Client Selection Strategy Using an Improved Federated Random High Local Performance Algorithm to Address High Non-IID Challenges

Pramote Sittijuk ¹, Narin Petrot ² and Kreangsak Tamee ^{3,*}

¹ Department of Computer Science and Information Technology, Faculty of Science, Naresuan University Phitsanulok 65000, Thailand; pramotes62@nu.ac.th

² Department of Mathematics, Center of Excellence in Nonlinear Analysis and Optimization, Faculty of Science, Naresuan University Phitsanulok 65000, Thailand; narinp@nu.ac.th

³ Department of Computer Science and Information Technology, Center of Excellence in Nonlinear Analysis and Optimization, Faculty of Science, Naresuan University Phitsanulok 65000, Thailand; kreangsakt@nu.ac.th

* Correspondence: kreangsakt@nu.ac.th; Tel.: 065-6380330

Abstract: This paper introduces an improved version of the Federated Random High Local Performance (Fed-RHLP) algorithm, specifically aimed at addressing the difficulties posed by Non-IID (Non-Independent and Identically Distributed) data within the context of federated learning. The refined Fed-RHLP algorithm implements a more targeted client selection approach, emphasizing clients based on the size of their datasets, the diversity of labels, and the performance of their local models. It employs a biased roulette wheel mechanism for selecting clients, which improves the aggregation of the global model. This approach ensures that the global model is primarily influenced by high-performing clients while still permitting contributions from those with lower performance during the model training process. Experimental findings indicate that the improved Fed-RHLP algorithm significantly surpasses existing methodologies, including FedAvg, Power of Choice (PoC), and Fed-Choice, by achieving superior global model accuracy, accelerated convergence rates, and decreased execution times, especially under conditions of high Non-IID data. Furthermore, the improved Fed-RHLP algorithm exhibits resilience even when the number of clients participating in local model updates and aggregation is diminished in each communication round. This characteristic positively influences the conservation of limited communication and computational resources.

Keywords: Client Selection; Improved Fed-RHLP algorithm; Non-IID issue; Model Aggregation

1. Introduction

In the context of Federated Learning (FL), the process of client selection plays a pivotal role in determining the most suitable participants for model training. Effective client selection contributes to accelerated convergence, improved global model accuracy, optimal utilization of heterogeneous data, and enhanced computational capabilities across various devices, ultimately leading to superior model performance [1]. Concurrently, it is essential to ensure communication efficiency and minimize computational resource consumption.

Engaging all clients can present significant challenges, including the potential unavailability of certain clients at specific times and communication bottlenecks arising from network infrastructure limitations. Consequently, an efficient client selection strategy focuses on identifying a smaller subset of clients for participation, which offers multiple advantages, such as reduced communication costs, improved resource efficiency, and enhanced model generalization [2].

A significant challenge in Federated Learning (FL) relates to suboptimal client selection and excessive diversity in model aggregation. Client heterogeneity can arise from statistical disparities, such as imbalanced datasets or substantial variations in data volume across clients, which may lead

to diminished model accuracy and prolonged convergence times [3]. Furthermore, the non-independent and identically distributed (Non-IID) characteristics of the data compound this challenge. Each client possesses training samples from a limited range of classes and demonstrates imbalanced sample distributions during the training process. The Non-IID challenge can be categorized into two distinct levels: high-level and low-level issues. High-level issues emerge when each client is assigned only a few labels for local model training. Conversely, low-level issues occur when clients have access to a broader range of labels, albeit not the complete set.

The challenges associated with heterogeneous client selection and low-quality data are particularly evident in the random client selection mechanism employed by the Federated Averaging algorithm (FedAvg). As previous research has indicated [4], FedAvg may result in suboptimal global model aggregation [5-6], leading to increased delays and diminished model accuracy, especially when addressing Non-IID data across clients [7-9]. In response to the issues arising from random client selection, various strategies have been developed that incorporate bias values to select clients based on their number of training samples and elevated local loss. One such strategy, was highlighted in [10], proposed the Power-of-Choice algorithm (PoC), which prioritizes clients exhibiting higher local loss values to enhance convergence rates. Nevertheless, the PoC approach may deviate from the optimal central loss function, potentially compromising overall model accuracy due to its singular focus on accelerating convergence. To refine the PoC methodology, the FedChoice algorithm was introduced [11]. This algorithm aims to improve global model convergence by favoring clients with higher loss values, thereby increasing their likelihood of participating in local model training and subsequent integration into the global model.

This study presents an improved version of the Fed-RHLP algorithm, building upon the initial iteration introduced in [12]. The primary objective of this improvement is to refine the process of robust client selection by identifying a more suitable and reduced subset of clients, thereby addressing the challenges posed by significant Non-IID issues. The improved Fed-RHLP algorithm prioritizes the selection of clients with three key characteristics: a substantial number of samples, demonstrated label diversity, and high local model performance across both local updates and global aggregation. Concurrently, the algorithm allows for the inclusion of lower-quality clients, enabling them to contribute their local datasets during the training and aggregation phases. The implementation of this algorithm results in accelerated convergence and improved accuracy, particularly in scenarios characterized by high Non-IID conditions. Moreover, it ensures fairness by facilitating proportional participation from the majority of clients.

2. Materials and Methods

2.1. Exploring Issues of Client Selection Methods

In the context of Federated Learning (FL), the selection of diverse clients, combined with limited data samples and suboptimal data quality, can impede convergence rates and contribute to the development of an inaccurate global model. Numerous studies have proposed client selection methodologies designed to address the statistical discrepancies encountered during the global model aggregation process. The challenge of statistical heterogeneity originates from multiple sources, including, variations in data distributions, inconsistent sampling practices, and non-independent and identically distributed (Non-IID) data or imbalanced datasets [13]. To manage this heterogeneity, it is crucial to implement client selection strategies during the training process that ensure convergence across the system, considering the diversity of statistical heterogeneity [14]. Furthermore, additional complexities arise from variations in model architectures, network conditions, and hardware capabilities among participating clients. These factors collectively exacerbate the existing challenges, diminishing the server's ability to perform effective model aggregation and resulting in a global model with compromised precision [15]. A comprehensive review and analysis of the challenges and limitations associated with these algorithms can be conducted as follows:

FedAvg represents a straightforward approach that involves the random selection of clients for local model updates, specifically employing Convolutional Neural Networks (CNNs) to distribute computational tasks among clients and address challenges associated with centralized machine learning. This method applies gradient descent to perform local updates on the models of a

randomly selected subset of clients, which are subsequently aggregated to create a global model. Although FedAvg averages the updates from individual client models, it may encounter significant divergence from the optimal model, particularly when the training data distribution of a single client markedly differs from the heterogeneous and non-IID data distribution [16 - 17].

The ideal model equation for FedAvg can be decomposed into two primary components: 1) Local Updates: Each client k refines its model w_{t+1}^k through gradient descent, employing a step size determined by η , (the learning rate). This process is influenced by the client's local dataset and the gradient g_k of the loss function and 2) Global aggregation: The locally updated models from the clients are amalgamated to form the global model w_{t+1} , with the aggregation weighted proportionally to the size of each client's dataset n_k relative to the total dataset size n , as shown in equation (1).

$$\forall k, w_{t+1}^k \leftarrow w_t - \eta g_k$$

$$w_{t+1} \leftarrow \sum_{k=1}^K \frac{n_k}{n} w_{t+1}^k \quad (1)$$

The problems with FedAvg can be demonstrated through the following example, as shown in Figure 1.

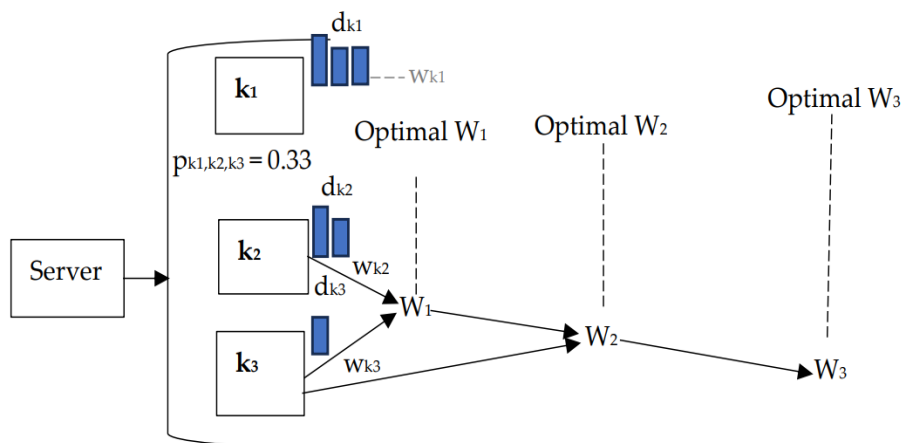


Figure 1. Divergence problem of random client selection in FedAvg.

In Figure 1, it is assumed that three clients ($K = 3$), represented as k_1 , k_2 , and k_3 , each responsible for updating their local models (w_{k1} , w_{k2} , and w_{k3}) with local dataset (d_k) and then contribute to aggregating a global model (W_t) in communication round (t). In the random client selection process, the probability of selecting a client (p_k) for local updates is equally distributed, with a probability value of $p_k = 1/K$ (0.33). We assume that k_1 has 100 samples with 3 labels ($n_k = 100$, $L_k = 3$), k_2 has 50 samples with 1 label ($n_k = 50$, $L_k = 2$), and k_3 has 20 samples with 1 label ($n_k = 20$, $L_k = 1$). If clients with low data quantity and few labels (k_2 and k_3) are frequently selected for updates, this can cause significant divergence.

Upon examining client selection methodologies such as the Power-of-Choice (PoC) algorithm and the FedChoice algorithm, it becomes evident that bias values are allocated to clients based on the volume of training samples and elevated local loss. A significant concern emerges when a client with a substantial number of samples is selected, particularly if these samples are predominantly from a single label. This scenario detrimentally affects the diversity of datasets utilized in model updates and global aggregation [18]. The issue of label scarcity further compounds this challenge, as certain clients may lack adequate labelled data due to resource constraints or insufficient motivation for labelling tasks. Such complications contribute to suboptimal model performance and impede the development of a well-generalized model across varied datasets. Ideally, client selection should encompass a range of labels alongside a considerable number of samples [19 - 20]. This approach ensures that local model training effectively captures diverse patterns, thereby enhancing the global model's generalization [21]. When addressing the bias gap arising from an exclusive reliance on large

sample sizes for client selection in local loss estimation and model up-dates within the PoC and FedChoice frameworks, as well as the conventional practice of utilizing all clients for local model performance assessment in Fed-RHLP, it becomes apparent that these approaches may result in a deficiency of local dataset diversity and significant execution time de-lays. This issue is evident, as shown in Figure 2.

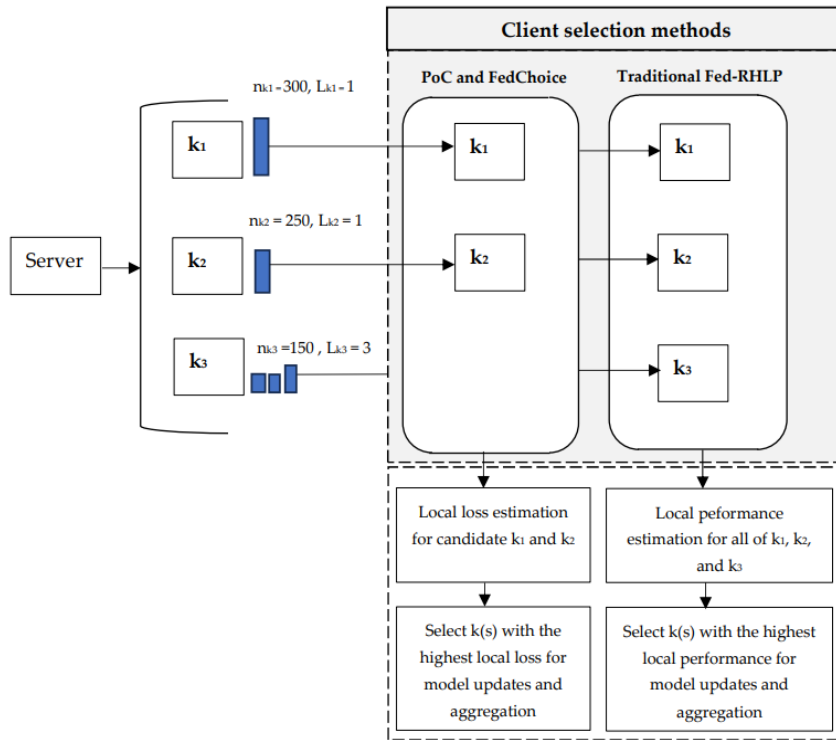


Figure 2. Concepts of client selection methods in PoC, FedChoice, and traditional Fed-RHLP.

In Figure 2, the comparison of client selection methods for local model updates in the PoC, FedChoice, and traditional Fed-RHLP algorithm is shown. In the example, there are three clients (k_1 , k_2 , and k_3).

The PoC and FedChoice algorithms assume that the server selects only two clients as candidates for local loss estimation in each communication round. These algorithms prioritize clients with larger local datasets for participation. The probability of selecting each client (k) based on its local dataset size is calculated as $P_k = n_k / (\sum_{k=1}^K n_k)$. For k_1 , with a local dataset size of $n_{k1} = 300$, the probability is calculated as $P_{k1} = 300/700 = 0.43$. For k_2 , with $n_{k2} = 250$, $P_{k2} = 250/700 = 0.36$. Finally, for k_3 , with $n_{k3} = 150$, $P_{k3} = 150/700 = 0.21$. As a result, k_1 has the highest probability of being selected, followed by k_2 , which has a higher probability than k_3 . This selection process often leads to clients with a single label, such as k_1 and k_2 , being selected repeatedly for local model training. Consequently, this reduces the diversity of the aggregated global model.

The traditional Fed-RHLP algorithm, proposed in [12], uses all clients (k_1 , k_2 , and k_3) to estimate the clients' local model performance (LP_k) by sending the global model (W) to each client for estimating with the local dataset (d_k). Subsequently, the LP_k of each client is converted into a weight for local model performance (WLP_k), which is added to a biased roulette wheel for randomly selecting model up-dates and aggregation based on clients' local model performance, as shown in Equations (2)–(4).

$$LP_k = f(W; d_k) \quad (2)$$

$$TotalLP = \sum_{k=1}^N LP_k \quad (3)$$

$$WLP_k = \frac{LP_k}{TotalLP} \quad (4)$$

The conventional Fed-RHLP algorithm is marked by significant computational demands, particularly in situations where certain clients are unavailable for processing at specific times. Additionally, the random selection of a large number of clients complicates the identification of those with high-quality datasets. This scenario may lead to missed processing opportunities and impede the accurate identification of clients eligible for participation.

2.2. Developing the Improved Federated RHLP Algorithm for Efficient Client Selection

The Improved Federated RHLP algorithm exhibits enhanced efficiency by increasing the accuracy of global model aggregation and accelerating the convergence rate. This improvement is achieved through its dual-bias approach, consisting of Module 1 and Module 2. These modules incorporate factors such as the number of samples, label diversity, and the high performance of local models from clients into the processes of model updates and aggregation, as shown in Figure 3.

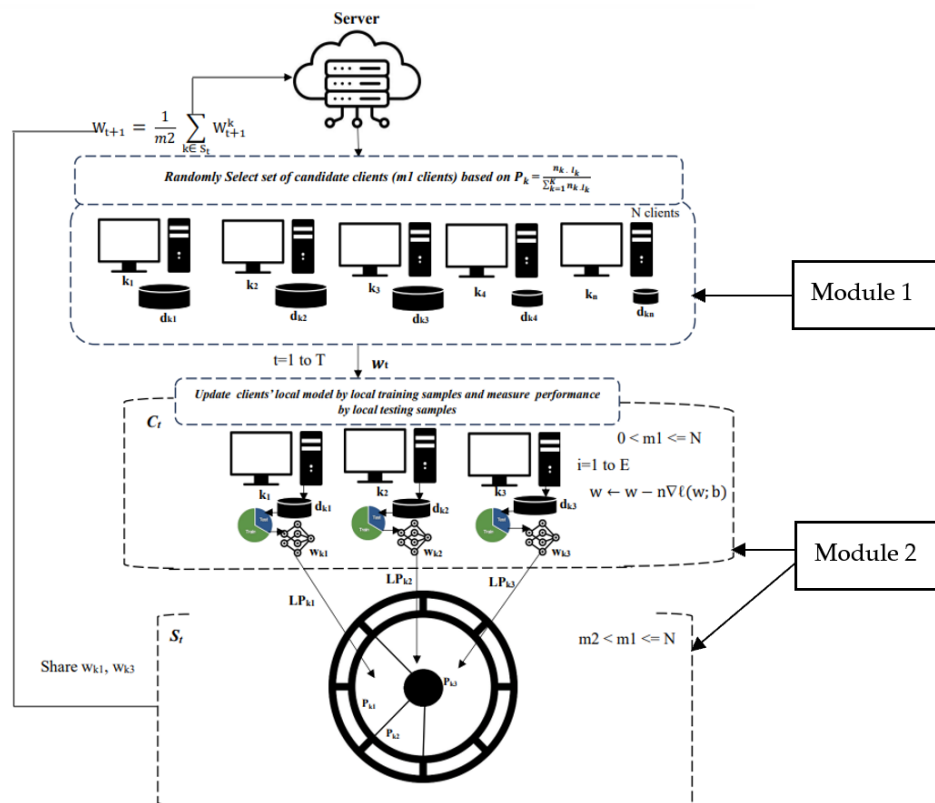


Figure 3. Architecture of Improved RHLP algorithm.

In Figure 3, the architecture of the Improved RHLP algorithm is developed in two modules:
Module 1: Candidate client selection

The client selection process requires improvement to more effectively select a smaller set of candidate clients (C_t) in each communication round, prioritizing based on the number of samples (n_k) and label diversity (L_k). In the example (Figure 4), the probability of selecting each client (k) is calculated as $P_k = \frac{n_k \cdot L_k}{\sum_{k=1}^K n_k \cdot L_k}$. For k_1 , $P_{k1} = (300 \times 1) / 1,000 = 0.3$; for k_2 , $P_{k2} = (250 \times 1) / 1,000 = 0.25$; for k_3 , $P_{k3} = (150 \times 3) / 1,000 = 0.45$. As a result, k_3 has the highest likelihood of being chosen for local model updates and performance estimation, followed by k_1 , which has a higher probability than k_2 . This proposed selection process ensures that clients with more diverse labels and larger datasets are prioritized, leading to a more diverse and robust aggregated global model.

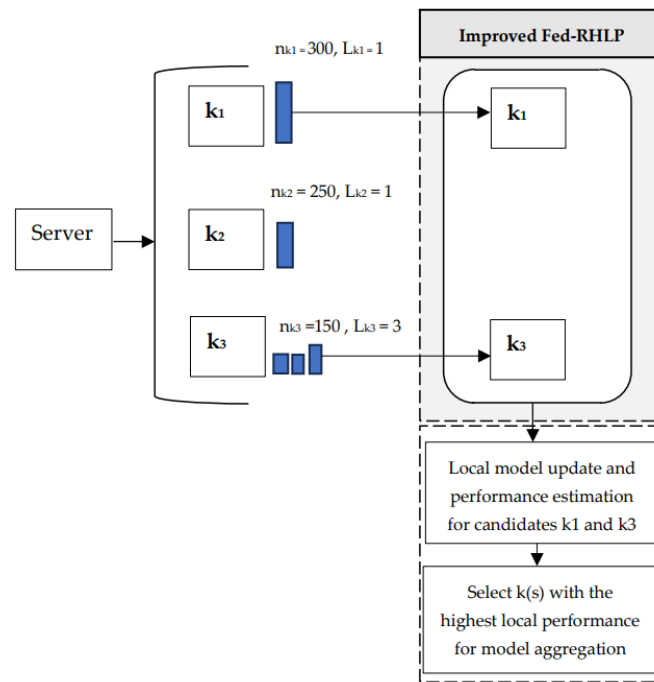


Figure 4. New concept of client selection method in Improved RHL algorithm.

Module 2: High local model performance updates and aggregation

The local model performance is measured in the algorithm by splitting each candidate's local dataset (d_k) into training samples (d_{traink}), where (d_{traink}) is further split into batches of size B for locally updating the model (w_k) by using the Stochastic Gradient Descent (SGD) Tech-nique, and small test samples (d_{testk}), which are used for testing local model performance of each client (LP_k) with $LP_k = f(w_k; d_{testk})$. High-performing local models, calculated from $TotalLP = \sum_{k \in C_t} LP_k$ and using Equation (4) to determine the weight of local model performance (WLP_k), are then randomly selected using a biased roulette wheel mechanism for global model aggregation. The de-tails of the improved RHL algorithm are presented in the pseudocode in Figure 5.

Algorithm 1: Improved Federated RHLF algorithm

```

1: ServerExecute :
2:  $W_0$  = Initial parameters of the CNN model
3:  $N$  = Total number of clients
4:  $C1$  = Fraction of candidate clients
5:  $C2$  = Fraction of selected clients
6:  $m1$  = Number of candidate clients
7:  $m2$  = Number of selected clients
8:  $n_k$  = Number of samples of client  $k$ 
9:  $l_k$  = Label diversity of client  $k$ 
10:  $d_{train\_proportion}$  = Proportion of local training samples
11:  $d_{test\_proportion}$  = Proportion of local testing samples
12: For  $t = 1$  to  $T$  do
13:    $m1 = \lfloor N \cdot C1 \rfloor$ 
14:    $C_t = \{ \text{random set of candidate (m1) clients based on } P_k = \frac{n_k \cdot l_k}{\sum_{k=1}^K n_k \cdot l_k}, \text{ replace} = \text{False} \}$ 
15:   For  $k \in C_t$  in Parallel do
16:      $w_{t+1}^k, LP_t^k = \text{ClientUpdate\_TestLP}(k, W_t)$ 
17:     TotalLP $_t += LP_t^k$ 
18:   End for
19:   For  $k \in C_t$  do
20:      $WLP_t^k = \frac{LP_t^k}{\text{TotalLP}_t}$ 
21:     ListWLP $_t.append(WLP_t^k)$ 
22:   End for
23:    $S_t = \{ \text{random set of higher local performance (m2) clients based on ListWLP}_t, \text{ replace} = \text{False} \}$ 
24:    $W_{t+1} = \frac{1}{m2} \sum_{k \in S_t} w_{t+1}^k$ 
25: End for
26: ClientUpdate_TestLP(k, w) :
27:  $d_{train}, d_{test} = \text{TrainTestSplit}(d_k, d_{train\_proportion}, d_{test\_proportion})$ 
28:  $B = (\text{Split } d_{train} \text{ into batches})$ 
29: For  $i=1$  to  $E$  do
30:   For  $b \in B$  do
31:      $w \leftarrow w - n \nabla \ell(w; b)$ 
32:   End for
33: End for
34:  $LP = f(w; d_{test})$ 
35: Return  $w, LP$  to the server

```

Figure 5. Pseudocode of Improved RHLF algorithm.

2.3. Experiment Setup

This study employs Convolutional Neural Networks (CNNs), recognized as a highly effective deep learning algorithm suitable for tasks involving image classification [22]. The experimental procedure is organized into three distinct phases, which are outlined in the following sections.

2.3.1. Dataset Preparation

In this study, we address the challenge posed by imbalanced and non-independent identically distributed (Non-IID) datasets that are distributed among clients, a critical concern in the realm of federated learning (FL) [23]. Such imbalances have a profound impact on the efficacy of model training [24]. The effectiveness of the improved Fed-RHLF algorithm was evaluated through comparisons with FedAvg, Power-of-Choice (PoC), and FedChoice, specifically focusing on Non-IID scenarios utilizing the MNIST and Fashion-MNIST (FMNIST) datasets.

The structures of these datasets are described and illustrated in Figure 5.

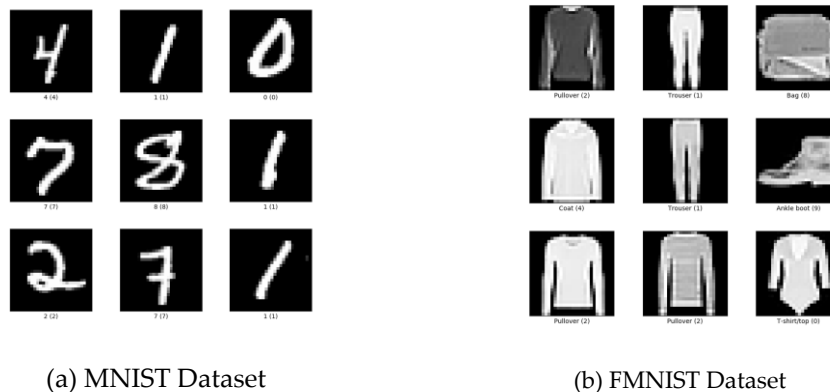


Figure 6. Online datasets in this experiment.

Figure 6 shows the use of two datasets: (a) the MNIST dataset and (b) the FMNIST dataset. Both datasets contain a total of 70,000 images, each being a 28×28 grayscale image. The MNIST dataset consists of handwritten digit images (ranging from 0 to 9) [25], while the FMNIST dataset contains fashion item images categorized into 10 classes [26]. Both datasets are split into 60,000 training images and 10,000 test images.

To demonstrate the resilience of the improved Federated RHL algorithm, we constructed Non-IID challenges to assess the algorithm's robustness at two distinct levels: 1) High Non-IID Challenge: Each client was assigned 1–2 classes for local model updates, and 2) Low Non-IID Challenge: Each client was assigned 5–6 classes for local model updates. The datasets were distributed among 100 clients, with each client receiving a varying proportion of data samples, ranging from 10% to 30%, to support the training of their local models.

2.3.2. Setting Up Algorithms, CNN Model Structures, and Hardware Specifications

Setting Up Algorithms

The algorithms will be run for 200 communication rounds to aggregate the global model. The SGD optimization algorithm is used in CNN for local model updates over five local epochs (E). The batch size is set to 64, and the learning rate for SGD is 0.01 [27 - 30].

In each communication iteration, the Improved Federated RHL, PoC, and FedChoice algorithms identify a subset of candidate clients (C_t), comprising 10% ($C=0.1$) of the total pool of 100 clients. Subsequently, a smaller group of clients (S_t) is selected in varying proportions [31] of 2%, 4%, 6%, 8%, and 10%. When the selection for S_t reaches 10%, the number of candidate clients is increased by 15% (adding 15 additional candidates) to ensure a sufficiently random selection process. Additionally, in the Improved Federated RHL approach, each client's local dataset is divided into local testing samples, randomly ranging between 3% and 5%, with the remaining data allocated for training in local model updates and performance evaluation. In contrast, the FedAvg approach does not designate predefined candidate clients but directly selects fractions of 2%, 4%, 6%, 8%, and 10% of clients for model training.

Setting Up CNN Model Structures

The structure of the CNN model is configured for training on the MNIST and FMNIST datasets, as outlined below:

1) MNIST Dataset:

The CMM model architecture comprises two convolutional layers. In the first layer, 10 feature maps are produced using 5×5 filters. The second layer increases the number of feature maps to 20, also employing 5×5 filters. To reduce the risk of overfitting, a dropout layer with a rate of 0.5 is applied after the second convolutional layer. The network architecture further includes two fully connected layers: the first layer reduces 320 features to 50 neurons, and the second layer maps these 50 neurons to 10 output classes, representing the digits 0 to 9.

2) FMNIST Dataset:

The CNN model architecture comprises two convolutional layers. The initial layer generates 32 feature maps using 3×3 filters with a padding of 1. A max pooling layer with a 2×2 kernel and a stride of 2 follows this first layer. The subsequent convolutional layer increases the number of feature maps to 64, again employing a 3×3 filter size. This is followed by another max pooling layer, utilizing the same 2×2 kernel and stride of 2. The model further incorporates three fully connected layers: The first layer transforms 2,304 features into 600 neurons. A dropout layer with a rate of 0.25 is introduced after this layer to mitigate the risk of over-fitting. The second fully connected layer reduces the 600 features to 120 neurons. The final layer maps the 120 features to 10 output classes, representing various fashion items.

Setting Up Hardware Component Specifications

This experiment prepares the hardware specifications for setting up and executing the algorithms, as outlined in Table 1.

Table 1. Hardware specifications for executing the algorithms.

Device	Specification
--------	---------------

Central Processing Unit	11th Gen Intel(R) Core(TM) i5-11400F 2.59 GHz
Graphics Processing Unit	Radeon (TM) RX 480 Graphics
Random Access Memory	32.0 GB
Operating System	Windows 11
Software Environment	Python 3.7.2 with the Pytorch framework

2.3.3. Algorithm Performance Evaluation

The global model of the algorithms was evaluated three times over 200 communication rounds, using 10,000 global test samples from the MNIST and FMNIST datasets. The formula for evaluating the Improved RHLP algorithm, compared with other benchmark algorithms, involves measuring prediction accuracy, con-vergence speed, and reduced execution time [11]. The calculation methods for these three-performance metrics are shown in Equations (5) – (7).

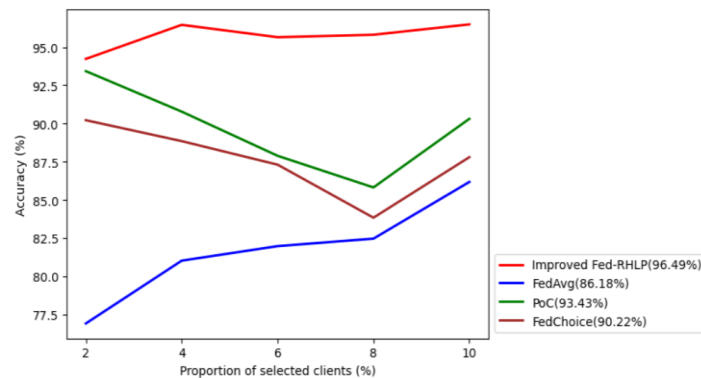
$$\text{Accuracy (\%)} = \frac{(\text{Total number of test samples} - \text{Number of incorrectly predicted samples}) \times 100}{\text{Total number of test samples}} \quad (5)$$

$$\text{Convergence speed (\%)} = \frac{(\text{Higher round} - \text{Lower round}) \times 100}{\text{Higher round}} \quad (6)$$

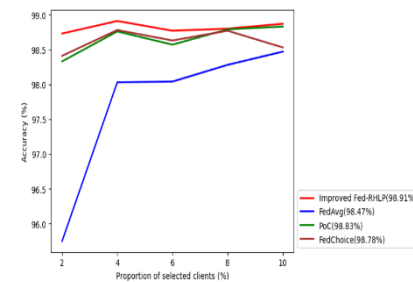
$$\text{Reduced execution time (\%)} = \frac{(\text{Higher execution time} - \text{Lower execution time}) \times 100}{\text{Higher execution time}} \quad (7)$$

3. Results

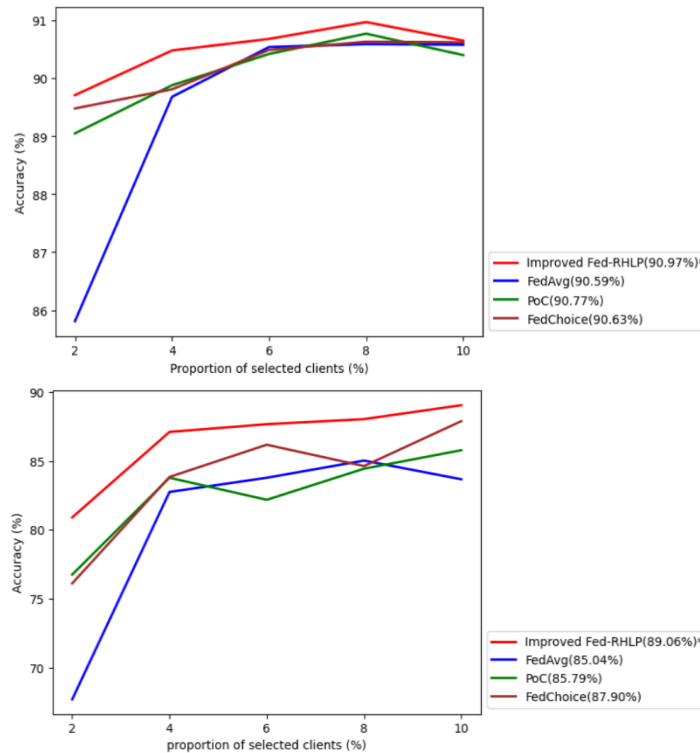
The Improved Fed-RHLP was assessed for its overall model accuracy and benchmarked against various algorithms across different ratios of selected clients. This evaluation was conducted using the MNIST and FMNIST datasets, which respectively represent high and low degrees of Non-IID challenges. The results are shown below



(a) Proportion of client selection in high level Non-IID issue in MNIST



(b) Proportion of client selection in low level Non-IID issue in FMNIST



(c) Proportion of client selection in high level selection in low level

Non-IID issue in FMNIST

FMNIST

(d) Proportion of client selection in low level

Non-IID issue in

Figure 7. Accuracy of the global model of the algorithms under varying proportions of client selection.

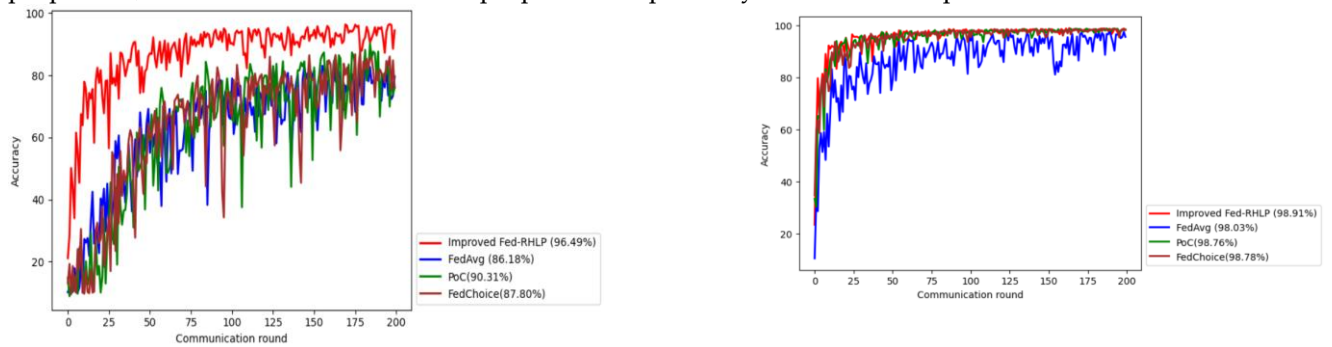
The data depicted in Figure 7 demonstrates that the improved Fed-RHLP exhibits superior global model accuracy compared to alternative client selection algorithms, irrespective of the severity of Non-IID challenges, as substantiated by the results derived from the MNIST and FMNIST datasets. In scenarios characterized by high Non-IID issues, the improved Fed-RHLP significantly outperforms its competitors across all client selection proportions. Conversely, under conditions of low Non-IID challenges, the algorithm provides a modest advantage over alternative approaches.

In scenarios characterized by high Non-IID challenges within the MNIST dataset, the improved Fed-RHLP achieves a peak global model accuracy of 96.49% when employing a selection proportion of 10%. This performance demonstrates notable improvements: 10.31% higher than FedAvg (highest accuracy: 86.18% at a 10% selection proportion) 3.06% higher than PoC (highest accuracy: 93.43% at a 2% selection proportion) and 6.27% higher than FedChoice (highest accuracy: 90.22% at a 2% selection proportion). Similarly, in the context of high Non-IID issues within the FMNIST dataset, the improved Fed-RHLP attains a global model accuracy of 89.06% at a selection proportion of 10%. This represents performance improvements of: 4.02% higher than FedAvg (highest accuracy: 85.04% at a 8% selection proportion), 3.27% higher than PoC (highest accuracy: 85.79% at a 10% selection proportion) and 1.16% higher than FedChoice (highest accuracy: 87.90% at a 10% selection proportion).

In scenarios characterized by low Non-IID challenges within the MNIST dataset, the improved Fed-RHLP achieves a peak global model accuracy of 98.91% with a selection proportion of 4%. This performance demonstrates marginal improvements: 0.44% higher than FedAvg (highest accuracy: 98.47% at a 10% selection proportion), 0.08% higher than PoC (highest accuracy: 98.83% at a 10% selection proportion), and 0.14% higher than FedChoice (highest accuracy: 98.78% with a 4% selection proportion). Similarly, in the context of low Non-IID issues in the FMNIST dataset, the improved Fed-RHLP attains a maximum global model accuracy of 90.97% with a selection proportion of 8%. This represents incremental performance improvements of: 0.38% higher than FedAvg (highest

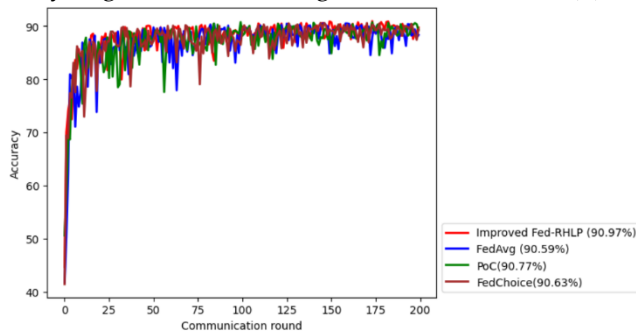
accuracy: 90.59% at a 8% selection proportion), 0.20% above PoC (highest accuracy: 90.77% at a 8% selection proportion), and 0.34% greater than FedChoice (highest accuracy: 90.63% at a 8% selection proportion).

Additionally, the improved Fed-RHLP underwent assessment for its convergence speed and execution time, benchmarked against alternative algorithms across both high and low levels of Non-IID challenges. The evaluation focused on achieving the highest global model accuracy under the following parameters: High Non-IID Conditions: MNIST: 10% client selection proportion FMNIST: 10% client selection proportion and in Low Non-IID Conditions: MNIST: 4% client selection proportion, FMNIST: 8% client selection proportion respectively. The results are presented as follows.



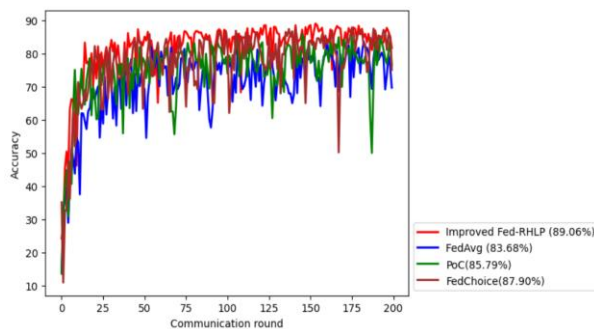
(a) Accuracy of global model in high level

(b) Accuracy of global model in low level



Non-IID issue in MNIST

Non-IID issue in MNIST



(c) Accuracy of global model in high level
level Non-IID issue in FMNIST

(d) Accuracy of global model in low level
Non-IID issue in FMNIST

Figure 8. Accuracy of the global model based on the proportion of selected clients that achieve the highest global model accuracy.

Figures 8a–d show that the improved Fed-RHLP achieves higher global model accuracy, faster convergence speed, and reduced execution time compared to FedAvg, PoC, and FedChoice when tested on MNIST and FMNIST under both high and low levels of Non-IID issues. The insights from the results are presented in Tables 2–3.

Table 2. Performance metrics for the global model of the algorithms on the Non-IID MNIST dataset.

Algorithm	Issue level	Accuracy of Global Model	Convergence Rounds and Time (Seconds) with Different Accuracy							
			60%		70%		80%		90%	
			Rounds	Time	Rounds	Time	Rounds	Time	Rounds	Time
improved Fed-RHLP	High	96.49%	6	633.6	11	1161.6	19	2006.4	41	4329.6
	Low	98.91%	3	486	3	486	6	972	10	1620
Fed-Avg	High	86.18%	32	1747.2	74	4040.4	126	6879.6	-	-
	Low	98.03%	9	680.4	12	907.2	12	907.2	52	3931.2
PoC	High	90.31%	43	3947.4	57	5232.6	85	7803	185	16983
	Low	98.76%	3	451.8	5	753	8	1204.8	15	2259
Fed-Choice	High	87.80%	40	3816	60	5724	92	8776.8	-	-
	Low	98.78%	3	464.4	5	774	8	1238.4	14	2167.2

Table 3. Performance metrics for the global model on the Non-IID FMNIST dataset.

Algorithm	Issue level	Accuracy of Global Model	Convergence Rounds and Time (Seconds) with Different Accuracy							
			60%		70%		80%		90%	
			Rounds	Time	Rounds	Time	Rounds	Time	Rounds	Time
improved Fed-RHLP	High	89.06%	6	1288.8	11	2362.8	15	3222	-	-
	Low	90.97%	1	328.2	2	656.4	7	2297.4	48	15753.6
Fed-Avg	High	83.68%	13	2386.8	23	4222.8	55	10098	-	-
	Low	90.59%	3	747	3	747	3	747	83	20667
PoC	High	85.79%	8	1929.6	9	2170.8	39	9406.8	-	-
	Low	90.77%	3	1072.8	5	1788	6	2145.6	85	30396
Fed-Choice	High	87.90%	8	1934.4	12	2901.6	21	5077.8	-	-
	Low	90.63%	2	714	4	1428	6	2142	59	21063

From Tables 2 and 3, the testing results on the MNIST and FMNIST datasets demonstrate that the improved Fed-RHLP exhibits significant advantages in addressing high Non-IID levels compared to other algorithms.

In the context of the MNIST dataset characterized by high Non-IID challenges, the improved Fed-RHLP demonstrates exceptional performance, achieving an accuracy of 96.49%. This result surpasses that of FedAvg, PoC, and FedChoice by margins of 10.31%, 6.18%, and 8.69%, respectively. At an accuracy level of 80%, Fed-RHLP exhibits both faster convergence and reduced execution time compared to all other algorithms. Most notably, it reaches 90% accuracy within merely 41 communication rounds, which represents a 77.84% improvement over PoC, accompanied by a 74.51% reduction in execution time. In scenarios with low Non-IID levels, Fed-RHLP significantly outperforms FedAvg, achieving an 80.77% enhancement in convergence speed and a 58.79% decrease

in execution time. Additionally, the algorithm marginally exceeds the performance of PoC and FedChoice.

In the context of the FMNIST dataset characterized by high Non-IID challenges, the improved Fed-RHLP demonstrates a peak accuracy of 89.06%. This performance exceeds that of FedAvg, PoC, and FedChoice by margins of 5.38%, 3.27%, and 1.16%, respectively. When achieving an accuracy of 80%, Fed-RHLP exhibits remarkable convergence characteristics: 72.73% more rapid convergence than FedAvg, 61.54% faster than PoC and 28.57% quicker than FedChoice. Correspondingly, the algorithm achieves significant reductions in execution time: 68.09% reduction compared to FedAvg, 65.75% reduction compared to PoC, and 36.55% reduction compared to FedChoice. In scenarios with low levels of Non-IID, Fed-RHLP demonstrates a marginal advantage over FedAvg, PoC, and FedChoice.

In summary, the improved Fed-RHLP exhibits greater efficiency and precision in managing elevated Non-IID levels across both datasets, showing notable advancements in convergence speed and execution duration when compared to alternative algorithms.

4. Discussion and Conclusions

This paper presents an enhanced version of the Fed-RHLP algorithm designed to address challenges posed by Non-IID data distributions. It selects a focused and diverse group of clients for training using a dual-biased strategy, considering factors such as dataset size, label variety, and local model performance. The goal is to optimize both data diversity and the quality of the global model, aligning with the approach in [32]. The Oort algorithm also tackles the participant selection challenge by considering data and system diversity to choose suitable participants.

A key feature of the enhanced version of the Fed-RHLP algorithm is the prioritization of high-performing clients while still allowing lower-performing clients to contribute. Focusing on high-performing clients during model aggregation improves global model accuracy, while including lower-performing clients ensures fairness and maintains data diversity. This balance between high and low-performing clients accelerates convergence and helps the model generalize well across diverse data distributions.

Additionally, this method reduces the computational cost of federated learning by selectively choosing clients based on dataset characteristics and local model performance. By minimizing the number of clients required for each aggregation round, execution time is reduced, which is especially beneficial in resource-constrained or large-scale scenarios. As outlined in [33], the approach improves computational efficiency and addresses communication overhead by dynamically selecting the best client models to share. This reduces unnecessary communication, lowers costs, and accelerates aggregation, making it especially valuable in federated settings with limited network bandwidth. The result is a more efficient FL framework with reduced computational load and faster model updates.

In conclusion, the enhanced Fed-RHLP algorithm improves both accuracy and convergence, especially in Non-IID environments with significant data heterogeneity. By balancing the involvement of high- and low-performing clients, the algorithm ensures fairness and increases the diversity of contributions to the global model. This balance accelerates convergence and reduces computational costs, making it an effective solution for federated learning in diverse data environments.

5. Future Works

In future work, we aim to improve the Fed-RHLP algorithm by incorporating hierarchical federated learning (FL) structures. As the client pool expands, the central server faces heightened computational demands. To alleviate this, we propose dividing clients into smaller subgroups for localized aggregation before transmitting the results to the central server. This strategy will help reduce the computational load and enhance scalability, as well as decrease communication overhead and speed up the aggregation process. Additionally, we plan to adapt this hierarchical framework to handle various data distributions and model types, ensuring its effectiveness in diverse environments. Furthermore, we aim to dynamically adjust client participation in each round, considering factors such as fluctuating client participation ratios and local training epochs. This

adjustment will be based on the influence of global model tuning and the contributions of client groups in each communication round, optimizing the efficiency of the global model.

Author Contributions: Conceptualization, P.S., N.P., and K.T.; methodology, P.S. and N.P.; software, P.S.; validation, P.S., N.P., and K.T.; formal analysis, P.S., N.P., and K.T.; investigation, P.S., N.P., and K.T.; resources, P.S., N.P., and K.T.; data curation, P.S., N.P., and K.T.; writing—original draft preparation, P.S.; writing—review and editing, P.S., N.P., and K.T.; visualization, P.S. and K.T.; supervision, P.S. and K.T.; project administration, P.S. and K.T.; funding acquisition, K.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Naresuan University (NU) and National Science, Research and Innovation Fund (NSRF), grant number 180613.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We thank Olalekan Israel Aiiikulola, Lecturer (Special Knowledge and Abilities), Faculty of Medical Science, Naresuan University, Thailand for carefully checking the general format of the manuscript.

References

1. Tahir M, Ali MI. Multi-Criterion Client Selection for Efficient Federated Learning. In Proceedings of the AAAI Symposium Series 2024 May 20 (Vol. 3, No. 1, pp. 318-322).
2. de Souza AM, Maciel F, da Costa JB, Bittencourt LF, Cerqueira E, Loureiro AA, Villas LA. Adaptive client selection with personalization for communication efficient Federated Learning. *Ad Hoc Networks*. 2024 Apr 15;157:103462.
3. Zhou T, Lin Z, Zhang J, Tsang DH. Understanding and improving model averaging in federated learning on heterogeneous data. *IEEE Transactions on Mobile Computing*. 2024 May 28.
4. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 20–22 April 2017; PMLR: Birmingham, UK, 2017; pp. 1273–1282.
5. Putra MA, Putri AR, Zainudin A, Kim DS, Lee JM. Accuracy-based client selection mechanism for federated industrial iot. *Internet of Things*. 2023 Apr 1;21:100657.
6. Schoinas, I., Triantafyllou, A., Ioannidis, D., Tzovaras, D., Drosou, A., Votis, K., Lagkas, T., Argyriou, V. and Sarigiannidis, P., 2024. Federated learning: challenges, SoTA, performance improvements and application domains. *IEEE Open Journal of the Communications Society*.
7. Zhang SQ, Lin J, Zhang Q. A multi-agent reinforcement learning approach for efficient client selection in federated learning. In Proceedings of the AAAI Conference on Artificial Intelligence 2022 Jun 28 (Vol. 36, No. 8, pp. 9091-9099).
8. Zhou H, Lan T, Venkataramani G, Ding W. On the Convergence of Heterogeneous Federated Learning with Arbitrary Adaptive Online Model Pruning. *arXiv preprint arXiv:2201.11803*. 2022 Jan 27.
9. Mu X, Shen Y, Cheng K, Geng X, Fu J, Zhang T, Zhang Z. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Generation Computer Systems*. 2023 Jun 1;143:93-104.
10. Cho YJ, Wang J, Joshi G. Client selection in federated learning: Convergence analysis and power-of-choice selection strategies. *arXiv preprint arXiv:2010.01243*. 2020 Oct 3.
11. Zeng et al. (2023). A Client Selection Method Based on Loss Function Optimization for Federated Learning. *Computer Modeling in Engineering & Sciences* 2023, 137(1), 1047-1064. <https://doi.org/10.32604/cmescs.2023.027226>.
12. Sittijuk P, Tamee K. Fed-RHLP: Enhancing Federated Learning with Random High-Local Performance Client Selection for Improved Convergence and Accuracy. *Symmetry*. 2024;16(9):1181. doi: 10.3390/sym16091181.
13. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, Bonawitz K, Charles Z, Cormode G, Cummings R, et al. Advances and open problems in federated learning. *Found Trends Mach Learn*. 2021;14(1–2):1–210.

14. Le DD, Tran AK, Dao MS, Nguyen-Ly KC, Le HS, Nguyen-Thi XD, Pham TQ, Nguyen VL, Nguyen-Thi BY. Insights into multi-model federated learning: An advanced approach for air quality index forecasting. *Algorithms*. 2022 Nov 17;15(11):434.
15. Yang H, Li J, Hao M, Zhang W, He H, Sangaiah AK. An efficient personalized federated learning approach in heterogeneous environments: a reinforcement learning perspective. *Scientific Reports*. 2024 Nov 21;14(1):28877.
16. Iyer VN. A review on different techniques used to combat the non-IID and heterogeneous nature of data in FL. *arXiv preprint arXiv:2401.00809*. 2024 Jan 1.
17. Hu M, Yue Z, Xie X, Chen C, Huang Y, Wei X, Lian X, Liu Y, Chen M. Is aggregation the only choice? federated learning via layer-wise model recombination. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2024 Aug 25* (pp. 1096-1107).
18. Zhang L, Fu L, Liu C, Yang Z, Yang J, Zheng Z, Chen C. Towards Few-Label Vertical Federated Learning. *ACM Transactions on Knowledge Discovery from Data*. 2024.
19. Xu Y, Liao Y, Wang L, Xu H, Jiang Z, Zhang W. Overcoming Noisy Labels and Non-IID Data in Edge Federated Learning. *IEEE Transactions on Mobile Computing*. 2024 May 9.
20. Li, A., Zhang, L., Tan, J., et al. (2021). Sample-level data selection for federated learning. *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 1–10. <https://doi.org/10.1109/INFOCOM42981.2021.9488723>.
21. Yan R, et al. Label-Efficient Self-Supervised Federated Learning for Tackling Data Heterogeneity in Medical Imaging. *IEEE Trans Med Imaging*. 2023 Jul;42(7):1932-43. doi: 10.1109/TMI.2022.3233574.
22. Tirumalapudi R, Sirisha J. Onward and Autonomously: Expanding the Horizon of Image Segmentation for Self-Driving Cars through Machine Learning. *Scalable Computing: Practice and Experience*. 2024 Jun 16;25(4):3163-71.
23. Singh G, Sood K, Rajalakshmi P, Nguyen DD, Xiang Y. Evaluating Federated Learning Based Intrusion Detection Scheme for Next Generation Networks. *IEEE Transactions on Network and Service Management*. 2024 Apr 4.
24. Shi X, Zhang W, Wu M, Liu G, Wen Z, He S, Shah T, Ranjan R. Dataset Distillation-based Hybrid Federated Learning on Non-IID Data. *arXiv preprint arXiv:2409.17517*. 2024 Sep 26.
25. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* 1998, 86, 2278–2324.
26. Xiao, H.; Rasul, K.; Vollgraf, R. Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms. *arXiv 2017*, arXiv:1708.07747.
27. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. *Proc. Mach. Learn. Syst.* 2020, 2, 429–450.
28. Ren, H.; Deng, J.; Xie, X. GRNN: Generative regression neural network — A data leakage attack for federated learning. *ACM Trans. Intell. Syst. Technol. (TIST)* 2022, 13, 1–24.
29. Wu, Y.; Liu, L.; Bae, J.; Chow, K.H.; Iyengar, A.; Pu, C.; Wei, W.; Yu, L.; Zhang, Q. Demystifying learning rate policies for high accuracy training of deep neural networks. In *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, 9–12 December 2019; IEEE: New York, NY, USA, 2019; pp. 1971–1980.
30. Zhai R, Jin H, Gong W, Lu K, Liu Y, Song Y, Yu J. Adaptive client selection and model aggregation for heterogeneous federated learning. *Multimedia Systems*. 2024 Aug;30(4):211.
31. Casella B, Esposito R, Sciarappa A, Cavazzoni C, Aldinucci M. Experimenting with normalization layers in federated learning on non-iid scenarios. *IEEE Access*. 2024 Apr 1.
32. Juan PH, Wu JL. Enhancing Communication Efficiency and Training Time Uniformity in Federated Learning through Multi-Branch Networks and the Oort Algorithm. *Algorithms*. 2024 Jan 23;17(2):52.
33. Al-Betar MA, Abasi AK, Alyasseri ZA, Fraihat S, Mohammed RF. A Communication-Efficient Federated Learning Framework for Sustainable Development Using Lemurs Optimizer. *Algorithms*. 2024 Apr 15;17(4):160.