# 연합 학습에서 모델 하이퍼 매개변수의 효과 분석

쿤드루 마지드, 김태홍
충북대학교 정보통신공학부, {kundroomajid, taehongkim}@cbnu.ac.kr

# Effects of Model Hyper-parameters in Federated Learning: A Short Analysis

Majid Kundroo, Kim Taehong
*Chungbuk National University, School of Information and Communication Engineering*

**Abstract**

This study emphasizes the importance of proper hyper-parameter selection, specifically learning rates and epochs, in improving the accuracy and effectiveness of a federated learning model. This study compares the effects of different learning rates and epoch sizes on the performance of a federated learning model to gain insights into their effects on convergence.

## 1. Introduction

Federated learning (FL) is becoming increasingly popular due to its privacy-preserving nature. However, current studies often employ the same hyper-parameters for all clients, disregarding their heterogeneity. This sub-optimal approach impedes model training as common hyper-parameters may not suit all clients, leading to uneven model performance and slower convergence. The participants in FL are inherently unique, so using the same hyper-parameters for all distinct clients is not a wise choice.

This study examines the impact of model hyper-parameters on the performance of FL models, focusing on learning rate and epoch size, to gain insights into their effects and role in convergence.

## 2. Methodology

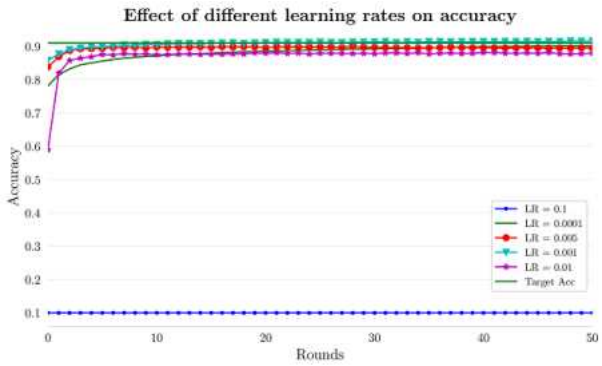
Our objective is to analyze the impact of model hyper-parameters in a FL. The learning rate determines the step size in optimizing the loss function, while the epoch refers to a complete iteration of model training.

We designed a FL system with 20 clients and one server to conduct our analysis. During each round, 10 randomly selected clients train their local models on their private data. After local training, all the clients send updated model weights to the server, where they are aggregated and transmitted back to all the participating clients. We utilize the FashionMNIST [1] data set, dividing it into 20 disjoint partitions to ensure independent and identically distributed (IID) data. A two-layer convolutional neural network serves as our machine learning model, with a target accuracy set at 91%.

Two sets of experiments were conducted: one varying the learning rates (0.1, 0.01, 0.001, 0.005, 0.0001) while keeping other hyper-parameters constant, and another with different epoch sizes (1, 2, 3, 4, 5, 10, 15, 20) while maintaining consistent hyper-parameters and conditions. The subsequent sections discuss the results obtained, their implications, conclusion and future scope.

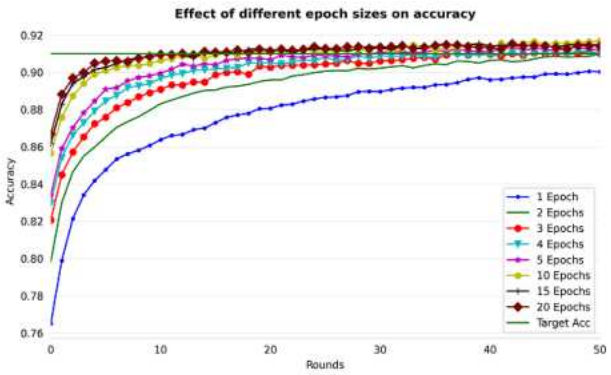
## 3. Results and Discussions

Figure 1 shows the influence of different learning rates on accuracy of a FL model. Higher learning rate values, such as 0.1, result in low accuracy and should be avoided. Lower learning rates, such as 0.0001, increase accuracy, but also increase training time and communication cost. Intermediate values between high and low learning rates, like 0.1, 0.005, and 0.0001, can be used to trade off an acceptable accuracy level and the number of rounds. To improve overall performance, it is critical to select an optimal learning rate in FL.




(Figure 1) Accuracies using different learning rates over rounds.

Figure 2 depicts the influence of different epochs on model performance. We can see that having fewer epochs on each round enhances the accuracy of the model at the cost of having more rounds. The needed number of rounds is lowered when the number of epochs per round is increased. The findings also indicate that some models may benefit from employing fewer epochs per round, particularly in the later phases of model training. Figure 2 shows that by using smaller values like 1 epoch per round, the model cannot reach higher accuracies quickly, while as using higher epoch values per round results in quicker convergence, thus by decreasing the convergence time and reduces the communication cost. But if we use larger epochs per round, each round will take more time to finish. As a result, it is preferable to customize the number of epochs per round for each client for optimal performance in FL.




(Figure 2) Accuracies using different epoch sizes over rounds.

## 4. Conclusion

In this study, we explored the effect of learning rates and epochs on the performance of FL models in this work. Our findings highlight the importance of selecting appropiate hyper-parameters in FL.

## Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2022R1I1A3072355).

## References

[1] Xiao, H., Rasul, K. and Vollgraf, R., 2017. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. arXiv preprint arXiv:1708.07747.