TP1: Introduction to Flower Framework Report

Mahmoud Mayaleh

May 24, 2025

1 Introduction

This report presents a systematic study of how key hyperparameters affect the performance and convergence of Federated Learning (FL) on the FashionMNIST dataset using the Flower framework. The experiments analyze the effects of the number of rounds, epochs, clients, batch size, learning rate, and Dirichlet distribution parameter (α).

2 Experimental Setup

- **Dataset:** FashionMNIST
- **Model:** [Simple CNN image classification model]
- **Framework:** Flower
- **Default Hyperparameters:**
  - SEED = 42
  - NUM_ROUNDS = 15
  - NUM_CLIENTS = 10
  - EPOCHS = 1
  - ALPHA_DIRICHLET = 1
  - BATCH_SIZE = 32
  - LEARNING_RATE = 0.01

3 Results and Analysis

3.1 Federated Learning Results for 30 Rounds

We conducted a federated learning experiment with 30 communication rounds using the FashionMNIST dataset. The results demonstrate a steady improvement in model performance as training progresses. As shown in Figure 1, the global model's accuracy increases from approximately 39% in the first round to over 86% by round 30, indicating effective convergence. The detailed accuracy values for each round are presented in Table 2. These results confirm that the chosen hyperparameters and federated averaging strategy enable the model to learn efficiently in a distributed setting.

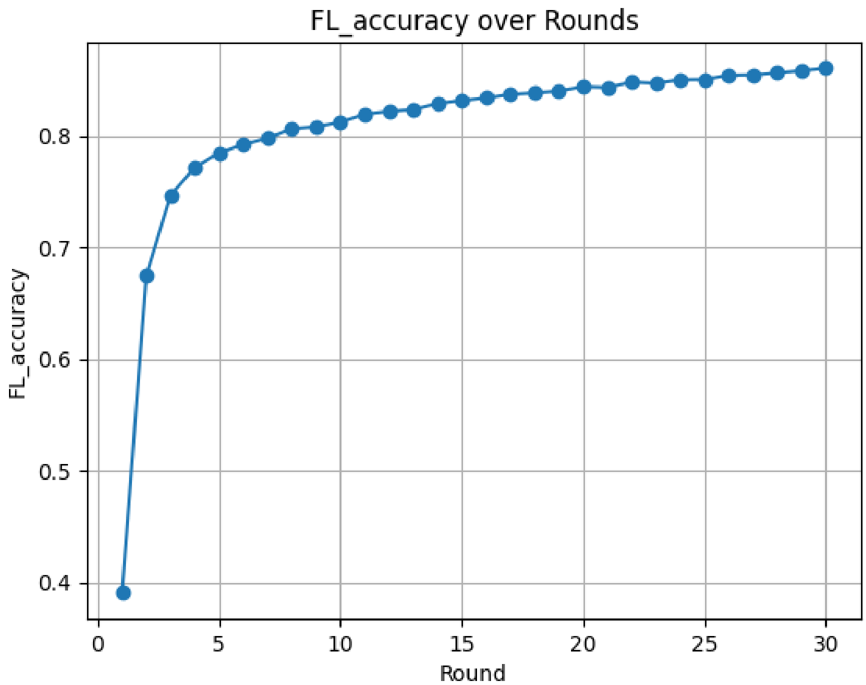

Figure 1: FL Accuracy over 30 Rounds

	Round	FL_loss	FL_accuracy
0	1	1.819	0.391
1	2	0.998	0.675
2	3	0.717	0.747
3	4	0.64	0.772
4	5	0.604	0.784
5	6	0.576	0.792
6	7	0.557	0.798
7	8	0.539	0.806
8	9	0.527	0.808
9	10	0.515	0.812
10	11	0.501	0.819
11	12	0.494	0.822
12	13	0.484	0.823
13	14	0.475	0.829
14	15	0.467	0.831
15	16	0.459	0.834
16	17	0.453	0.837
17	18	0.447	0.838
18	19	0.439	0.84
19	20	0.434	0.844
20	21	0.429	0.843
21	22	0.423	0.848
22	23	0.418	0.847
23	24	0.411	0.85
24	25	0.407	0.85
25	26	0.402	0.854
26	27	0.401	0.854
27	28	0.394	0.856
28	29	0.39	0.858
29	30	0.387	0.86

Figure 2: FL Accuracy Table for 30 Rounds

3.2 Effect of Number of Epochs

Tested values: 1, 2, 4, 10

Increasing the number of epochs per round generally led to faster convergence and higher final accuracy, but with diminishing returns and sometimes overfitting for very high values.

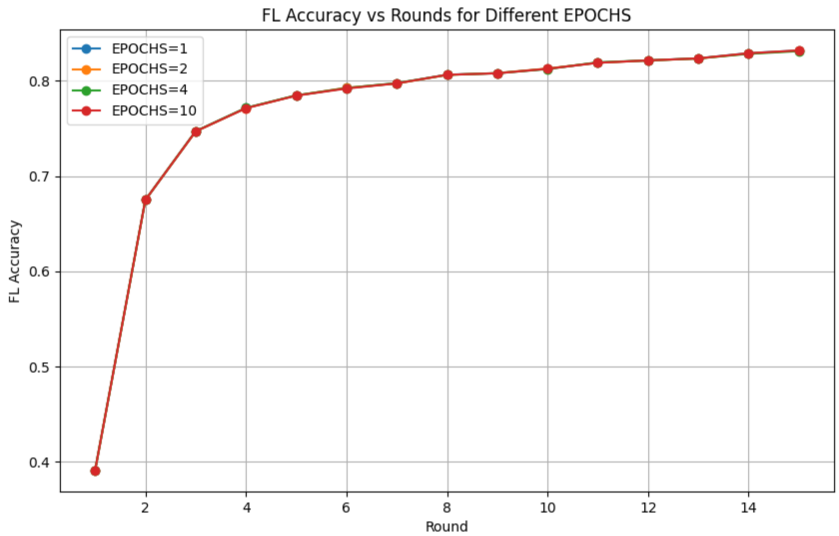

Figure 3: FL Accuracy vs Rounds for Different Epochs

	Round	Acc (EPOCHS=1)	Acc (EPOCHS=2)	Acc (EPOCHS=4)	Acc (EPOCHS=10)
0	1	0.391	0.391	0.391	0.391
1	2	0.675	0.675	0.675	0.675
2	3	0.747	0.747	0.747	0.747
3	4	0.772	0.771	0.772	0.771
4	5	0.785	0.785	0.785	0.784
5	6	0.793	0.793	0.792	0.792
6	7	0.797	0.797	0.798	0.797
7	8	0.806	0.806	0.806	0.806
8	9	0.808	0.808	0.808	0.808
9	10	0.813	0.813	0.812	0.813
10	11	0.819	0.819	0.819	0.819
11	12	0.821	0.821	0.821	0.821
12	13	0.824	0.823	0.824	0.824
13	14	0.829	0.829	0.828	0.829
14	15	0.832	0.832	0.831	0.832

Figure 4: Accuracy Table for Different Epochs

3.3 Effect of Number of Clients

Tested values: 2, 5, 10

More clients increased data diversity and generalization, but too many clients could slow convergence due to increased heterogeneity.

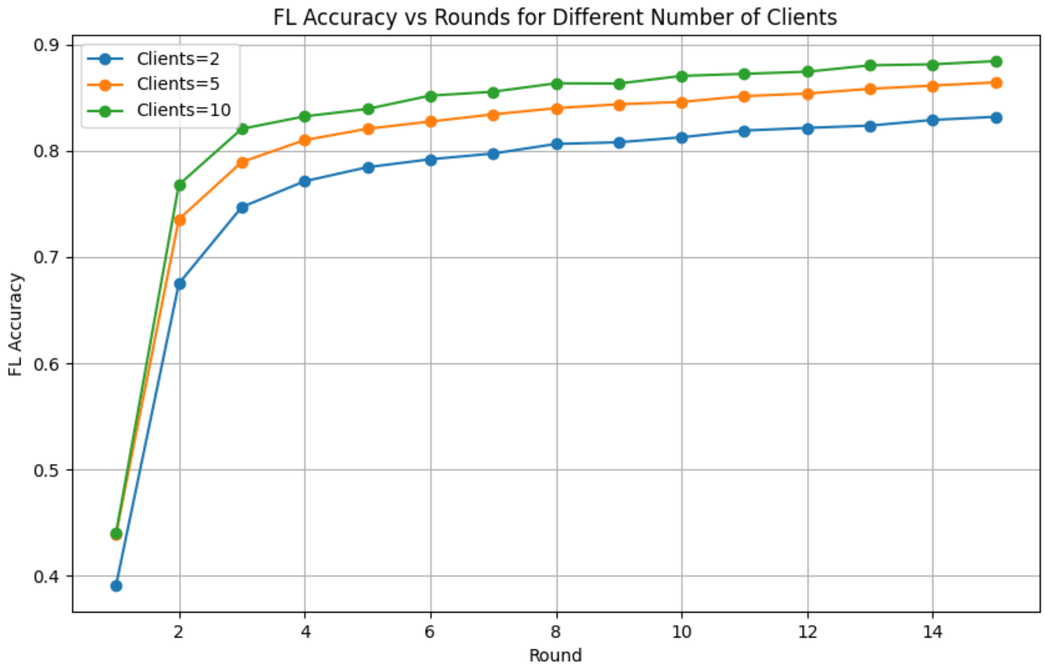

Figure 5: FL Accuracy vs Rounds for Different Number of Clients

	Round	Acc (Clients=2)	Acc (Clients=5)	Acc (Clients=10)
0	1	0.391	0.440	0.440
1	2	0.675	0.735	0.768
2	3	0.747	0.789	0.821
3	4	0.771	0.810	0.832
4	5	0.784	0.821	0.839
5	6	0.792	0.827	0.852
6	7	0.797	0.834	0.855
7	8	0.806	0.840	0.863
8	9	0.808	0.844	0.863
9	10	0.813	0.846	0.870
10	11	0.819	0.851	0.872
11	12	0.821	0.854	0.874
12	13	0.824	0.858	0.880
13	14	0.829	0.861	0.881
14	15	0.832	0.864	0.884

Figure 6: Accuracy Table for Different Number of Clients

3.4 Effect of Learning Rate

Tested values: 0.001, 0.01, 0.05, 0.1

A moderate learning rate (e.g., 0.01) provided stable and fast convergence. Too high a learning rate caused instability, while too low slowed learning.

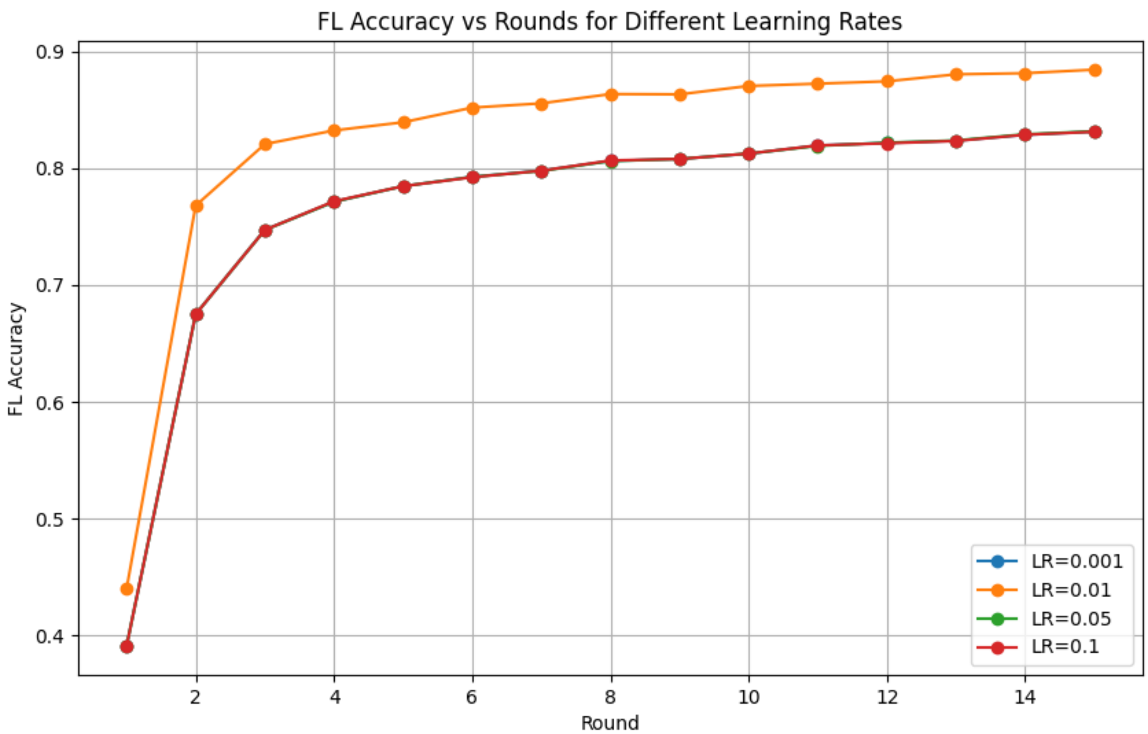

Figure 7: FL Accuracy vs Rounds for Different Learning Rates

	Round	Acc (LR=0.001)	Acc (LR=0.01)	Acc (LR=0.05)	Acc (LR=0.1)
0	1	0.391	0.440	0.391	0.391
1	2	0.675	0.768	0.675	0.675
2	3	0.747	0.821	0.747	0.747
3	4	0.771	0.832	0.771	0.772
4	5	0.785	0.839	0.785	0.785
5	6	0.793	0.852	0.793	0.792
6	7	0.798	0.855	0.798	0.798
7	8	0.806	0.863	0.806	0.807
8	9	0.808	0.863	0.808	0.808
9	10	0.812	0.870	0.813	0.813
10	11	0.819	0.872	0.819	0.820
11	12	0.822	0.874	0.822	0.821
12	13	0.823	0.880	0.824	0.823
13	14	0.829	0.881	0.829	0.829
14	15	0.831	0.884	0.832	0.831

Figure 8: Accuracy Table for Different Learning Rates

3.5 Effect of Batch Size

Tested values: 16, 32, 64, 128

Smaller batch sizes led to noisier but sometimes faster convergence. Larger batch sizes were more stable but could slow down learning.

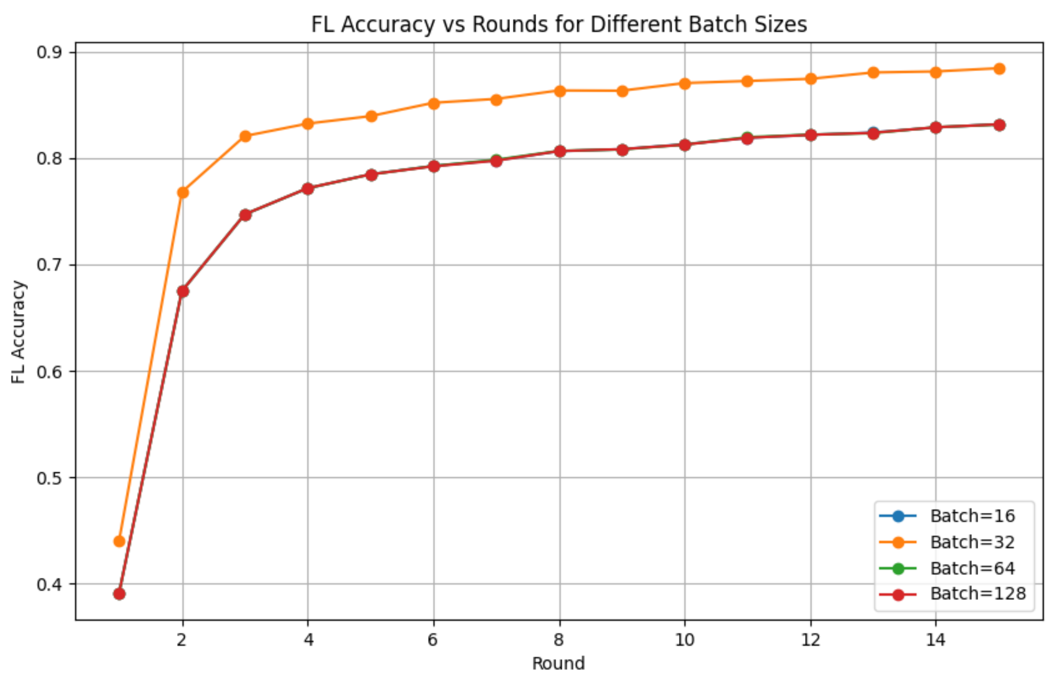

Figure 9: FL Accuracy vs Rounds for Different Batch Sizes

	Round	Acc (Batch=16)	Acc (Batch=32)	Acc (Batch=64)	Acc (Batch=128)
0	1	0.391	0.440	0.391	0.391
1	2	0.675	0.768	0.675	0.675
2	3	0.747	0.821	0.747	0.747
3	4	0.771	0.832	0.772	0.772
4	5	0.785	0.839	0.785	0.785
5	6	0.792	0.852	0.792	0.792
6	7	0.798	0.855	0.798	0.797
7	8	0.806	0.863	0.807	0.806
8	9	0.808	0.863	0.808	0.808
9	10	0.813	0.870	0.813	0.813
10	11	0.819	0.872	0.820	0.819
11	12	0.822	0.874	0.822	0.822
12	13	0.824	0.880	0.823	0.824
13	14	0.829	0.881	0.829	0.829
14	15	0.832	0.884	0.831	0.832

Figure 10: Accuracy Table for Different Batch Sizes

3.6 Effect of Dirichlet α (Data Heterogeneity)

Tested values: 0.1, 0.5, 1, 10

Lower α increases data heterogeneity (clients have less overlap in data), making FL harder and slowing convergence. Higher α makes data more homogeneous, improving convergence.

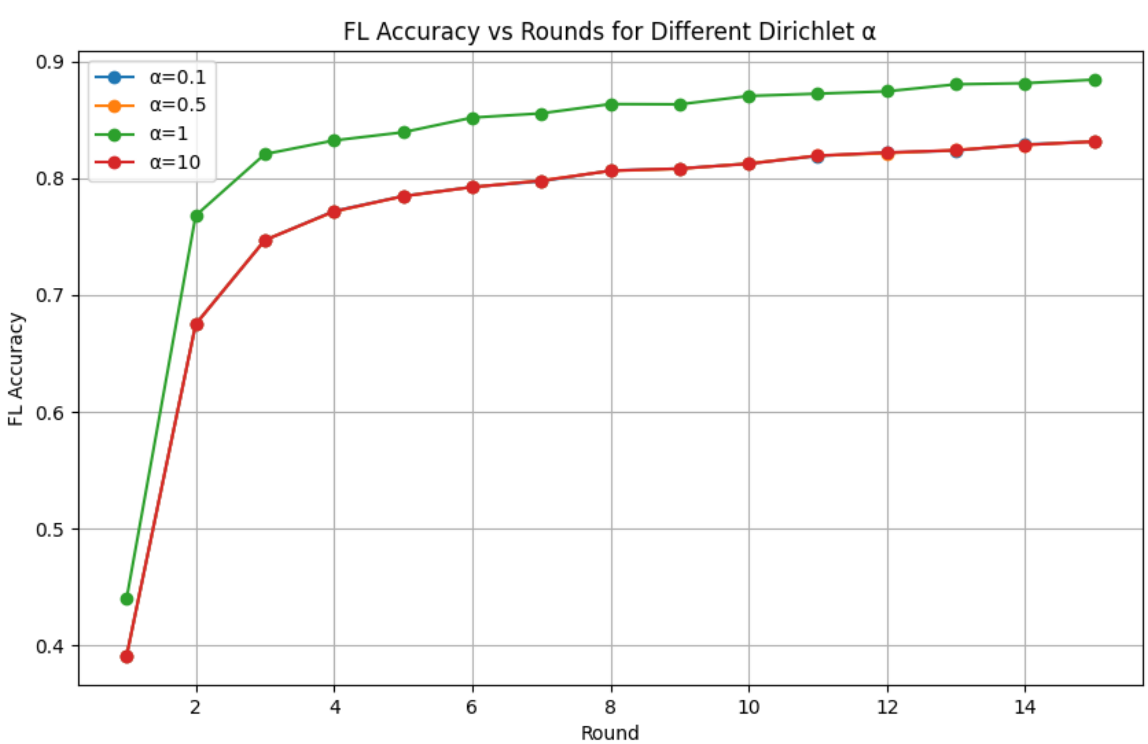

Figure 11: FL Accuracy vs Rounds for Different Dirichlet α

	Round	Acc (α=0.1)	Acc (α=0.5)	Acc (α=1)	Acc (α=10)
0	1	0.391	0.391	0.440	0.391
1	2	0.675	0.675	0.768	0.675
2	3	0.747	0.747	0.821	0.747
3	4	0.772	0.771	0.832	0.772
4	5	0.785	0.785	0.839	0.784
5	6	0.792	0.792	0.852	0.792
6	7	0.797	0.798	0.855	0.798
7	8	0.806	0.806	0.863	0.806
8	9	0.808	0.808	0.863	0.808
9	10	0.813	0.813	0.870	0.812
10	11	0.819	0.819	0.872	0.819
11	12	0.822	0.821	0.874	0.822
12	13	0.824	0.824	0.880	0.824
13	14	0.829	0.829	0.881	0.828
14	15	0.831	0.831	0.884	0.831

Figure 12: Accuracy Table for Different Dirichlet α

4 Conclusions

- **Rounds:** More rounds can speed up convergence and increases accuracy.
- **Epochs:** More epochs per round can speed up convergence but may risk overfitting.
- **Clients:** More clients improve generalization but can slow convergence if data is highly non-iid.
- **Learning Rate:** Needs to be tuned for stability and speed; 0.01 worked best in our case.
- **Batch Size:** Moderate batch sizes (32 or 64) balanced stability and speed.
- **Dirichlet α:** Lower α increases heterogeneity, making FL harder but more realistic.