Name: Victor Olawale-Apanpa

DATE: 19 NOV 2025

- Github Username: vapanpa
- How to Run: Notebook should be in the same directory as 'HW3-data' folder.

Run as is to produce output folders 'output/part1' and 'output/part2' with all required plots.

Federated Learning and Differential Privacy — Report

Project: FedAvg + Differential Privacy Analysis

---

# 1. Introduction

Federated Learning (FL) allows multiple clients to collaboratively train a global model without sharing raw data. Instead, each client trains locally and sends model updates to a central server. This preserves privacy and reduces communication of sensitive information.

In this project, we implement: 1. FedAvg (Non-DP) — baseline federated averaging 2. Differentially Private FedAvg — adding Laplace noise to client model updates

We evaluate: • Model convergence behavior • Privacy–utility tradeoffs • Impact of noise levels on training stability and performance

---

# 2. Dataset and Non-IID Client Distribution

The training dataset is partitioned across 100 clients in a non-IID (label-skewed) manner. Below are sample visualizations of the label distributions for clients 0–4.

Client Label Distributions

Alt text Alt text Alt text Alt text Alt text

Global Label Distribution

Alt text

Observation: • Each client has a unique, uneven label distribution. • Some labels are heavily overrepresented for certain clients. • This confirms a highly non-IID scenario, which makes FedAvg training more challenging.

# 3. Part 1 — FedAvg (No Differential Privacy)

## 3.1 Model Convergence

Below are the accuracy and loss curves for standard FedAvg.

FedAvg Accuracy vs Rounds Alt text

FedAvg Loss vs Rounds Alt text

Analysis • Training accuracy and test accuracy increase steadily across rounds. • Final test accuracy reaches ~0.63. • Loss decreases smoothly, indicating stable convergence. • This demonstrates the baseline performance without privacy noise.

FedAvg successfully learns despite non-IID distributions.

# 4. Part 2 — Differential Privacy (DP-FedAvg)

In this section, we add Laplace noise with scale parameter b to each client's model update before aggregation.

## 4.1 DP Training (Example: b = 0.1)

DP Loss Curve (b = 0.1) Alt text

DP Accuracy Curve (b = 0.1) Alt text

Analysis • Loss decreases early on but eventually plateaus or increases slightly. • Accuracy is lower than FedAvg and improves more slowly. • Noise introduces instability, especially in early rounds.

Still, the model learns meaningful structure even under DP.

# 5. Privacy–Utility Tradeoff Analysis

We evaluate model utility (final test accuracy) at four different Laplace noise levels: • b = 0.0 (no privacy) • b = 0.01 • b = 0.05 • b = 0.1

Final Accuracy vs Noise Scale b Alt text

Interpretation • b = 0.0 (no noise) achieves the highest accuracy (~0.63). • Accuracy drops sharply at b = 0.01, indicating strong privacy → strong utility loss. • Increasing noise from 0.01 → 0.1 slightly improves accuracy, showing the model can tolerate moderate noise once it stabilizes. • However, all DP levels underperform the baseline.

Conclusion

Higher noise improves privacy but reduces accuracy. There is a clear privacy–utility tradeoff.

# 6. Final Summary

FedAvg (No DP) • Achieves stable convergence and highest accuracy. • Handles non-IID data reasonably well.

Differentially Private FedAvg • Adding Laplace noise reduces accuracy. • Lower noise (b=0.01) hurts performance the most. • Moderate noise (b=0.05–0.1) allows limited recovery. • Overall: privacy comes at a meaningful cost to model utility.

Key Takeaways • FL works well in label-skewed environments but benefits from privacy-aware strategies. • Differential privacy is effective but requires careful tuning. • The tradeoff between privacy and accuracy must be balanced depending on the application.