Federated Learning with SGD Variants: A Comparative Study

Course: [Your Course Name]
Team Members: [Name 1, Name 2]
Date: [Submission Date]
GitHub: [Repository Link]

Abstract

This project implements and evaluates five federated learning algorithms derived from different SGD optimizer variants. We compare these algorithms against the FedAvg baseline on the FEMNIST dataset under various non-IID data distributions. Our results demonstrate that [brief key finding]. The implementation provides a modular framework for federated learning experimentation.

Keywords: Federated Learning, SGD Variants, Non-IID Data, FedAvg, Adaptive Optimization

---

1. Introduction

1.1 Motivation

Federated learning enables training machine learning models on decentralized data while preserving privacy. However, the standard FedAvg algorithm faces challenges with:

- Heterogeneous (non-IID) data distributions across clients
- Slow convergence in high-variance settings
- Sensitivity to hyperparameters

Different SGD optimizer variants (Adam, AdaGrad, RMSprop, etc.) have been successful in centralized settings, but their adaptation to federated learning remains an active research area.

1.2 Research Questions

1. How do different SGD variants perform when adapted to federated learning?
2. Is it better to apply adaptive optimization on clients or on the server?
3. How robust are these algorithms to non-IID data distributions?
4. What are the trade-offs between convergence speed and final accuracy?

1.3 Contributions

- Implementation of 5 federated learning algorithms with different SGD variants

- Comprehensive experimental evaluation on FEMNIST dataset
- Analysis of performance under varying non-IID conditions (Dirichlet α)
- Open-source framework for