Project Title: Federated Learning for Secure Healthcare Data Analysis in Distributed Systems

Team Members:

Dipesh Tripathi
Ujjwal Bhatta
Sumaly Bajracharya

Abstract:

Federated Learning (FL) offers a robust framework for secure healthcare data analysis within distributed systems, addressing the critical need for privacy-preserving methodologies in collaborative model training. In the context of healthcare, one significant challenge is the limited availability and diversity of datasets, which is further compounded by stringent privacy regulations that inhibit data sharing among institutions. FL effectively mitigates these challenges by enabling models to learn from diverse domains and sources without the necessity of transferring or centralizing sensitive data. This approach not only enhances data richness but also ensures compliance with regulatory standards. Employing advanced privacy-preserving techniques, such as differential privacy and secure aggregation, this framework safeguards data confidentiality while harnessing a wide array of healthcare data types.

This project aims to develop a secure FL framework tailored for scalable and accurate healthcare data analysis within distributed systems. The objective is to identify potential optimization areas for model performance and resource allocation. To achieve this, experimental evaluations will simulate a multi-institutional environment, rigorously assessing model accuracy, communication efficiency, privacy preservation, and scalability in real-world healthcare settings. By exploring strategies to manage data heterogeneity, enhance model accuracy, and reinforce security measures, this project aspires to bridge existing gaps in FL research, promoting effective collaboration among healthcare institutions while upholding the highest standards of data security and regulatory compliance.

Key Words: Federated Learning, distributed systems, healthcare data, data heterogeneity