Cryptographically secure federated learning for healthcare consortia

Federated learning (FL) is an approach to build global models without the need to share data to increase data privacy. Models are trained on edge devices and only updates are shared with a central server for aggregation. However, many researchers have proved that Federated is subjected to many inference attacks and poisoning attacks that can compromise data privacy or model performance.

This research and development project focused on identifying the causes and solutions of privacy vulnerabilities in Federated Learning and implementing a Federated Learning model with an optimized privacy-preserving technique in a healthcare setting.

In recent years, some popular encryption techniques were applied with FL which included **Homomorphic encryption, Secure MultiParty Computation** and **Differential Privacy.** However, many research papers revealed that these techniques are either computationally expensive or negatively impact model accuracy.

We proposed combining **cryptograms** with Federated Learning to address these issues. This technique is based on the principles of **elliptic curve cryptography (ECC)** and aims to encrypt client model updates, achieving competitive accuracy with lower computational expense.

To test our proposed approach we developed a Federated Learning model using **Convolutional Neural Network (CNN)** on the **ChestMNIST** dataset. Our FL setting consisted of three clients whose extracted weights of CNN model were encrypted by implementing cryptograms. The encrypted weights were then aggregated by a server.

In comparison with other encryption techniques when applied to the same FL model our approach showed no drop in accuracy and the iterations were significantly faster.

The final product of this project is a **web-based application** which allows registered clients to upload their local model's weights. The back-end system performs encryption using cryptograms and aggregates the updates on a centralized server. Once the FL completes aggregations, the combined model gets available for the clients to download and use.