FEDERATED LEARNING

Aggregation strategies for Federated Learning

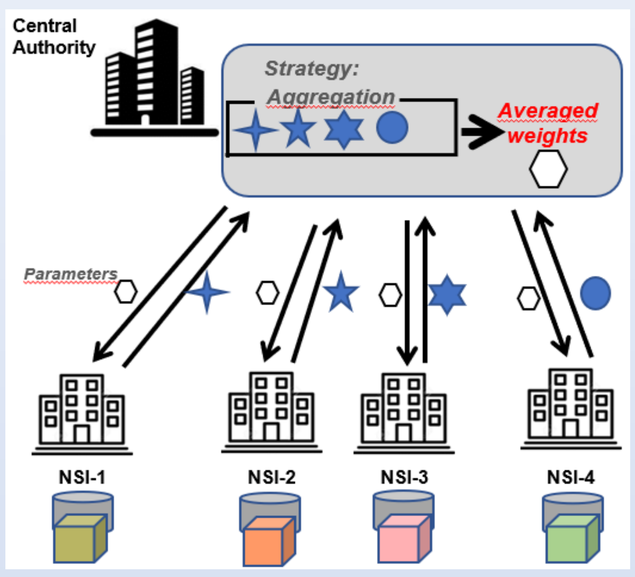
Conferenza Nazionale di **Statistica**

MASSIMO DE CUBELLIS - ISTAT decubell@istat.it | **MAURO BRUNO -** ISTAT mbruno@istat.it | **FRANCESCO PUGLIESE** – ISTAT frpugli@istat.it | **ERIKA CERASTI -** ISTAT cerasti@istat.it | **JULIAN TEMPLETON –** Statistics Canada julian.templeton@statcan.gc.ca | **BENJAMIN SANTOS –** Statistics Canada benjamin.santos@statcan.gc.ca **RAFIK CHEMLI –** Statistics Canada rafik.chemli@statcan.gc.ca | **MATJAZ JUG -** CBS m.jug@cbs.nl

Objectives

This poster presents a collaborative research conducted with Statistics Canada and Statistics Netherlands on Federated Learning (FL). It evaluates various FL aggregation strategies to determine their effectiveness in scenarios involving heterogeneous datasets and Homomorphic Encryption (HE) approaches. The research assesses different FL aggregation strategies and explores new approaches to protect the privacy of local models during aggregation. The findings of this research will provide National Statistical Institutes (NSI) with a more comprehensive understanding of the adaptability and potential of FL in various contexts, particularly when combined with other Privacy Enhancing Technologies (PET).

FL overview



FL is a Machine Learning (ML) methodology which enables ML models to be trained across distributed devices while keeping the training data stored locally on each device.

Using FL, an organization can train a ML model held by some central authority without requiring the training data to be shared with the authority. This permits analytics to be derived from distributed private data sources.

During the training process, the global model held by the central authority is sent to all clients (individuals or organizations) participating in training that model for the training round.

The locally updated models are then sent to and aggregated by the central authority who stores the updated global model for future use or further training.

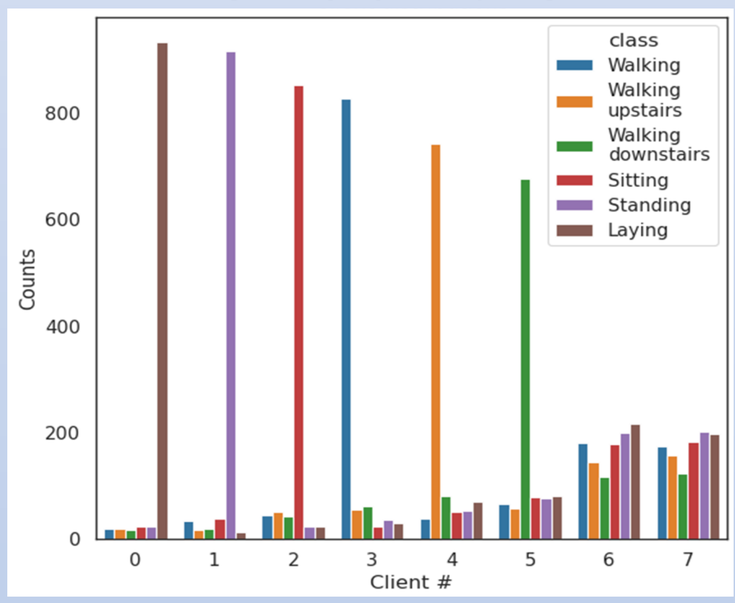
FL aggregation strategies: the use case

This research utilizes a simplistic Human Activity Recognition (HAR) public dataset comprising of accelerometer and gyroscope smartphone data (3-axial linear acceleration and 3-axial angular velocity, rate of 50Hz) from 30 volunteers (19-48 years).

The ML model aims to ascertain human activity types from the following six categories: walking, walking upstairs, walking downstairs, sitting, standing, and laying.

Within these tests, the HAR dataset has been partitioned among eight clients, aiming to establish a set of clients with a mix of heterogeneous and homogeneous local datasets. We have prepared the following four partitioning methodologies to distribute the data:
Random: samples are randomly distributed among clients
Majority even: each client has one majority class, same number of records
Majority: each client has one majority class, different number of records
Pick two: each client has two majority classes, same number of records
Note that each class can only be assigned as a majority class once, where remaining clients without a majority class are given a distribution of all classes.

Human activity class distribution among clients using the 'Majority Even' splitting method

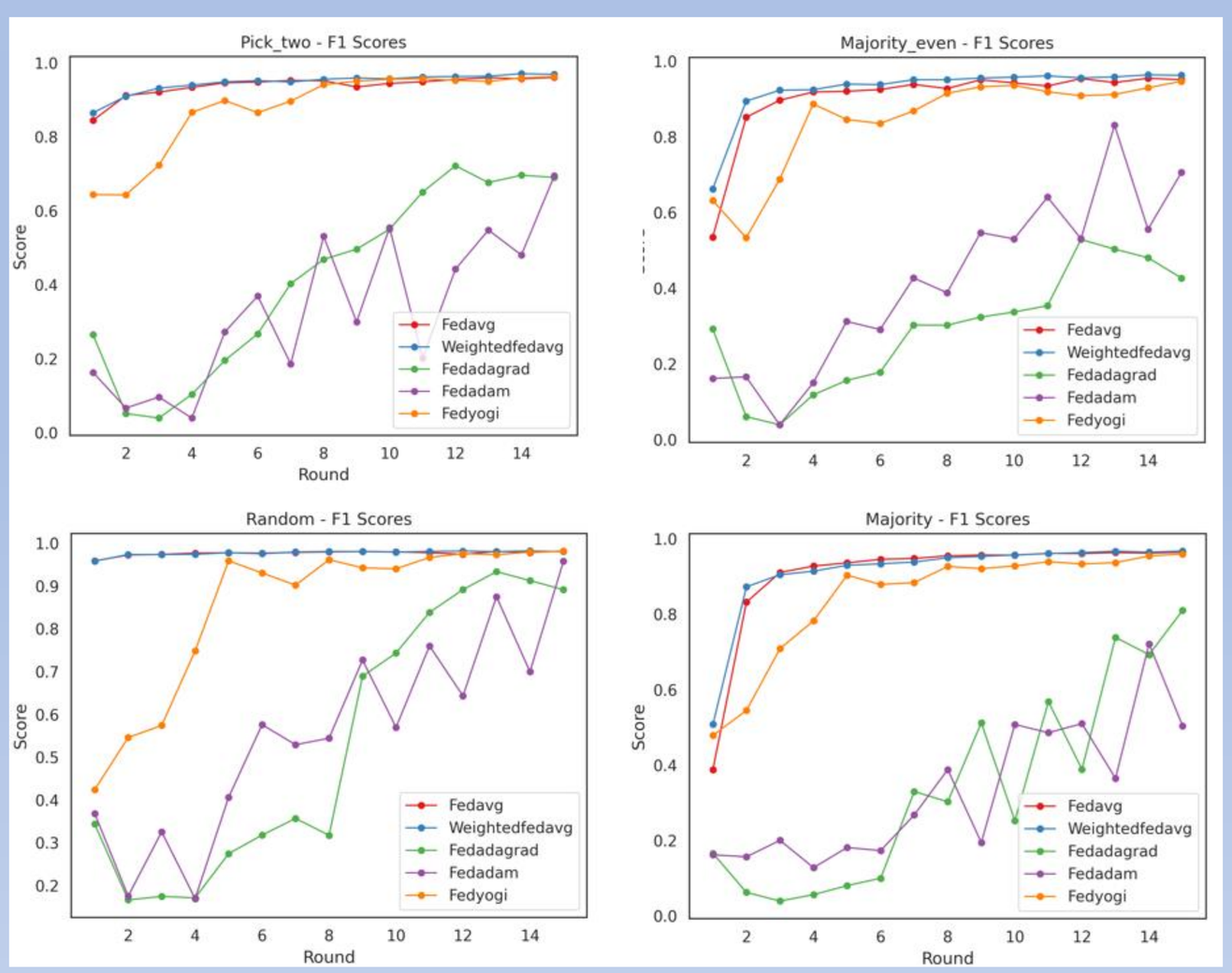


Aggregation strategies:

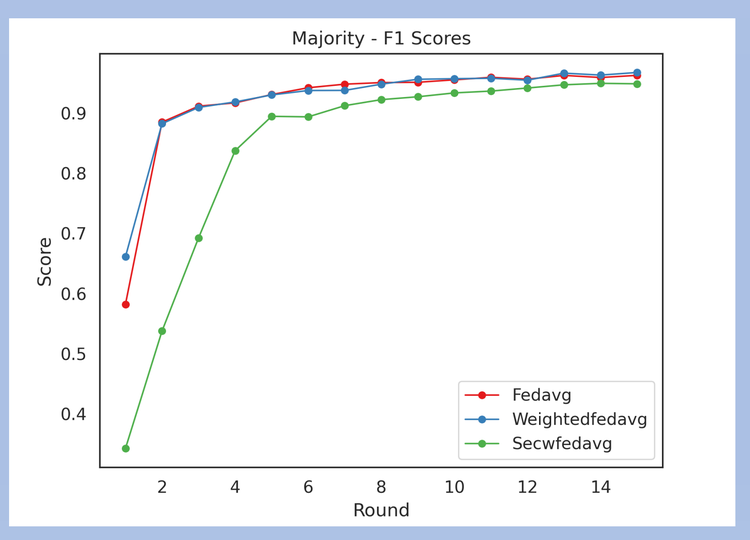
- ✓ Federated Averaging (FedAvg): is a FL algorithm that aims to train a global model by aggregating the local model updates from multiple clients by calculating the average of the model parameters.
- ✓ Federated Adaptive Gradient (FedAdagrad): is a variant algorithm that exploits the adaptive gradient descent method called Adagrad. It adapts the learning rate for each model parameter based on its historical gradients, allowing the model to converge faster and achieve better performance.
- ✓ Federated Adam (FedAdam): is another FL algorithm that combines the advantages of the Adam optimizer with the FL setting. It employs adaptive learning rates and momentum to efficiently update the global model using the local updates from clients. The gradients computed locally by the devices are aggregated in the central server.
- ✓ Federated Yogi (FedYogi): is a FL algorithm inspired by the Yogi optimizer. It incorporates elements of both adaptive learning rates and momentum to handle non-convex optimization problems in FL scenarios.
- ✓ Weighted FedAvg (WFedAvg): is a customized weighted FL aggregation strategies, named WFedAvg with the aim of better understanding the complexity of homogeneous and heterogeneous data scenarios in FL approach.

Results

Comparative analysis between different federated aggregation strategies



Performance of FedAvg and WFedAvg compared to FedAvg with HE



Conclusions

With reference to the comparative analysis (on the left), the best models are FedAvg and WFedAvg, which are very similar. Adaptive methods (FedYogi) do not show a better performance with heterogenous data. It seems that the simpler aggregation technique (FedAvg) work better for this dataset. With reference to HE(on the right), it makes the convergence slower. HE can add significant time and communication costs, scaling with the amount of encrypted weights/gradients. Heterogeneity of the training dataset seems to not affect performance in this example.