FLAMENCO

FLAMENCO: A privacy-preserving Federated Learning Application for diagnosis of communication disorders in child development

About Project

The FLAMENCO project applied Federated Learning to an existing software application suite, which is used to diagnose communication skills development in children, detecting potential deficiencies timely and accurately.
The software suite collects data from a child's responses to an animation game, along with heart rate readings obtained from a smartwatch. The data is then sent and stored into a cloud data hub and analyzed using AI-based classification techniques. The outcome is a risk indicator that suggests the likelihood of a child developing any learning and communication disorders. As the data collected was sensitive and personal, and the predictive model requires continuous and incremental training, more sophisticated techniques were required.
The FLAMENCO project employed several algorithms to ensure users' privacy and prediction accuracy. First, the project used Fully Homomorphic Encryption during the model aggregation step to protect user data from potential breaches. Second, client selection techniques were utilised to remove users with corrupted or missing data to improve the model's predictive accuracy. Third, state-of-the-art aggregators were introduced to manage data imbalance, ensuring that the Federated Learning model can converge effectively. The project outcome simulates a Federated Learning process using real-world data from IoT edge devices and incorporating the proposed extensions. The communication between FLAMENCO's hardware modules utilised the MQTT protocol to seamlessly integrate with TERMINET's existing hardware infrastructure.
Additionally, a web application was developed to enhance the overall user experience. Overall, the project's results demonstrate the potential of this decentralised approach to pave the way towards personalised AI-enabled healthcare solutions while respecting patient privacy.

Dataset

In the context of the FLAMENCO project, we have released a dataset designed for predicting potential deficiencies in children's communication skills, tailored for Federated Learning. This dataset specifically focuses on addressing two prevalent deficiencies in communication skill development in children: autism and intellectual disability. The dataset includes scores, such as Verbalization, Voicing, Syntax, etc., that are derived from the child's performance in specialised gamified exercises and have been computed with the assistance of expert clinicians. The target value per each case can be -1 (no clinician's diagnosis available for the case), 0 (no diagnosed deficiency in the case), 1 (indicates a positive diagnosis of communication deficiency by a clinician).
The machine learning task associated with this dataset involves predicting the probability of a case being diagnosed with a specific communication deficiency. We have 5 clinicians with 451 total cases. The dataset train-test split was performed based on our proposed solution architecture that handles this problem as an anomaly detection task using a semi-supervised learning approach.

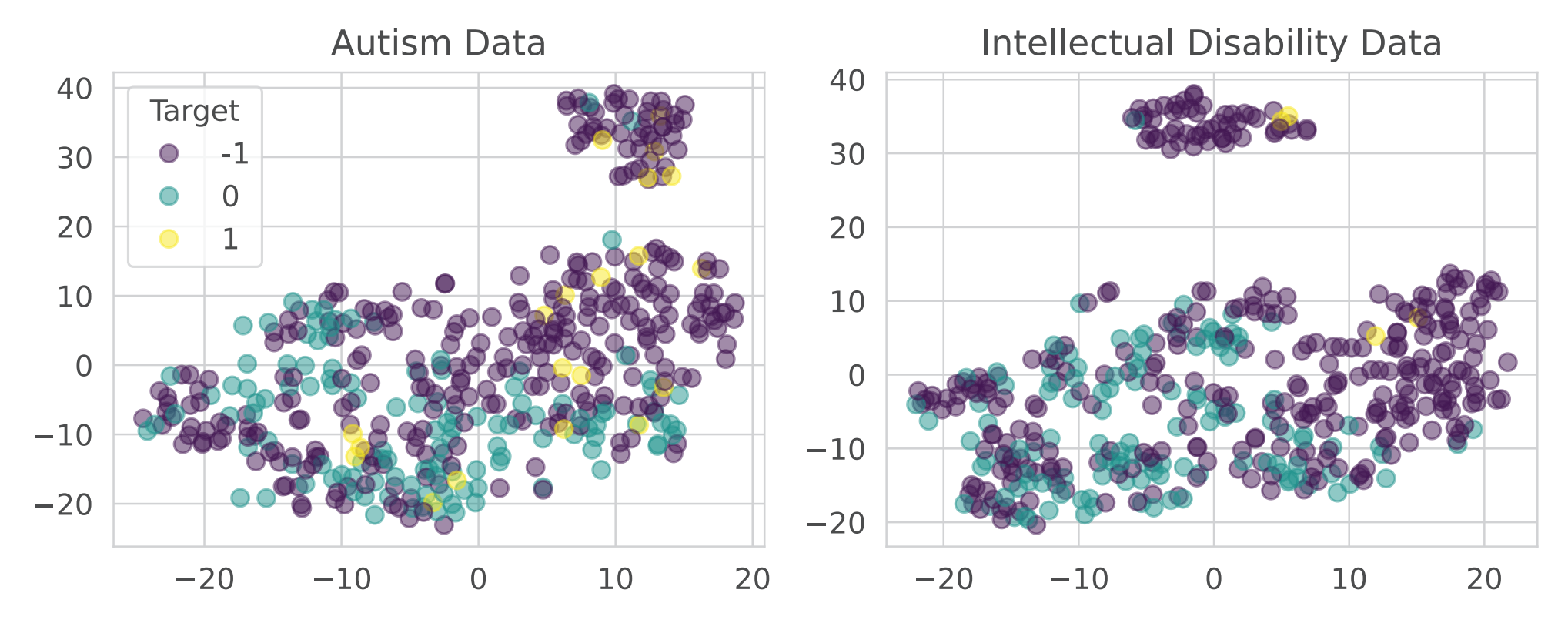
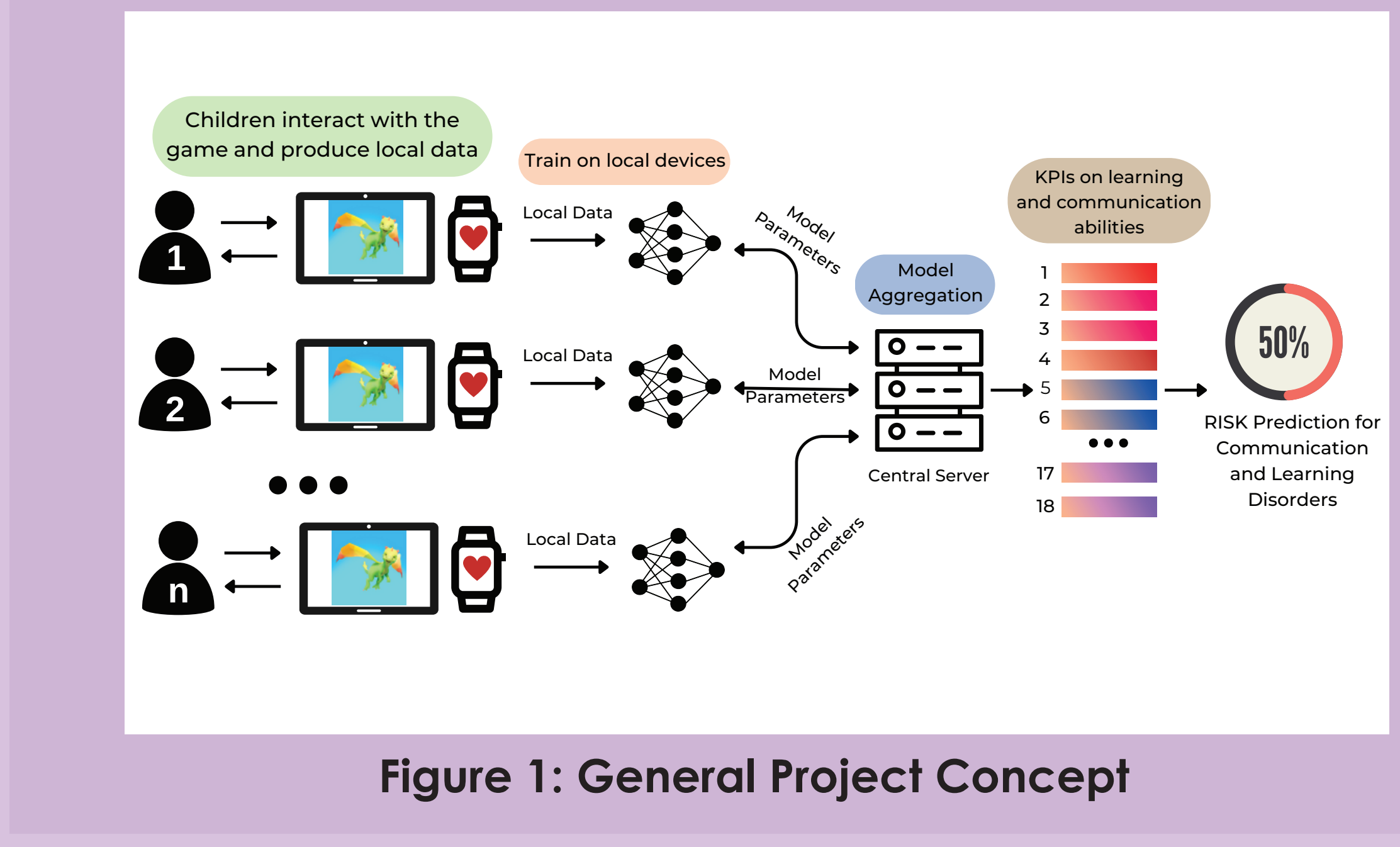


Figure 4: TSNE Representation of the dataset

Problem Formulation

Our methodology involves training ML models on "normal" cases labeled as 0 (no diagnosed deficiency) and some selected unlabeled cases (-1) that are presumed to be normal. During training, the model is evaluated on all possible labels, i.e., negative, unknown and positive cases. To this end, we utilized AutoEncoder-based architectures, a type of neural networks that try to reconstruct the original input. The goal throughout the project was to train the model on "normal" data and learn a hidden representation of the normal distribution. When the network encounters data that significantly differ from the training samples (in the project's context, cases with disability), we expect that it will result in higher reconstruction error since the associated patterns differ from what the model learned as normal. During inference, when the model encounters samples that are hard to reconstruct accurately, it will result in high reconstruction error, indicating a potential disability.
In our respective scernario we have 5 clinicians with 451 total cases. Each one owns a local dataset without ever being transmitted. A central server is responsible for communication between clients and orchestrating the FL process in general. It starts every Federated round by transmitting the learnable weights of a common global model across the clients. Then, each client performs local training on its private data using the classical gradient descent optimization. After that, every client retransmits its local model weights back to the central server, which aggregates them to produce the new global weight. During the project we implemented the following key modules for federated learning procedure:
-**Machine Learning Models**: AutoEncoder (AE), Variational AE (VAE)
-**Metrics**: Precision@k, AUC-ROC, AP, MSE, SIREOS
-**Aggregators**: FedAvg, FedNova, FedAvgM, FedAdagrad, FedYogi, FedAdam, SimpleAvg, MedianAvg
-**Fully Homorphic Encryption**
The basic idea of the project is depicted in Fig. 1



Children interact with the game and produce local data
Train on local devices
1
2
n
Local Data
Local Data
Local Data
Model Parameters
Model Parameters
Model Parameters
Model Aggregation
Central Server
KPIs on learning and communication abilities
1
2
3
4
5
6
17
18
50%
RISK Prediction for Communication and Learning Disorders

Figure 1: General Project Concept

Web application

To demonstrate a basic operational instance of utilizing TERMINET's infrastructure within the FLAMENCO project, we created a proof of concept. This proof of concept employs an MQTT broker to coordinate communication between Federated Learning clients and the application server. TERMINETs MQTT broker server, offers a straightforward replacement for the one utilized in our initial proof of concept. Our final implementation scheme is shown in Fig.2. A UI mockup of our application is depicted in Fig. 3.

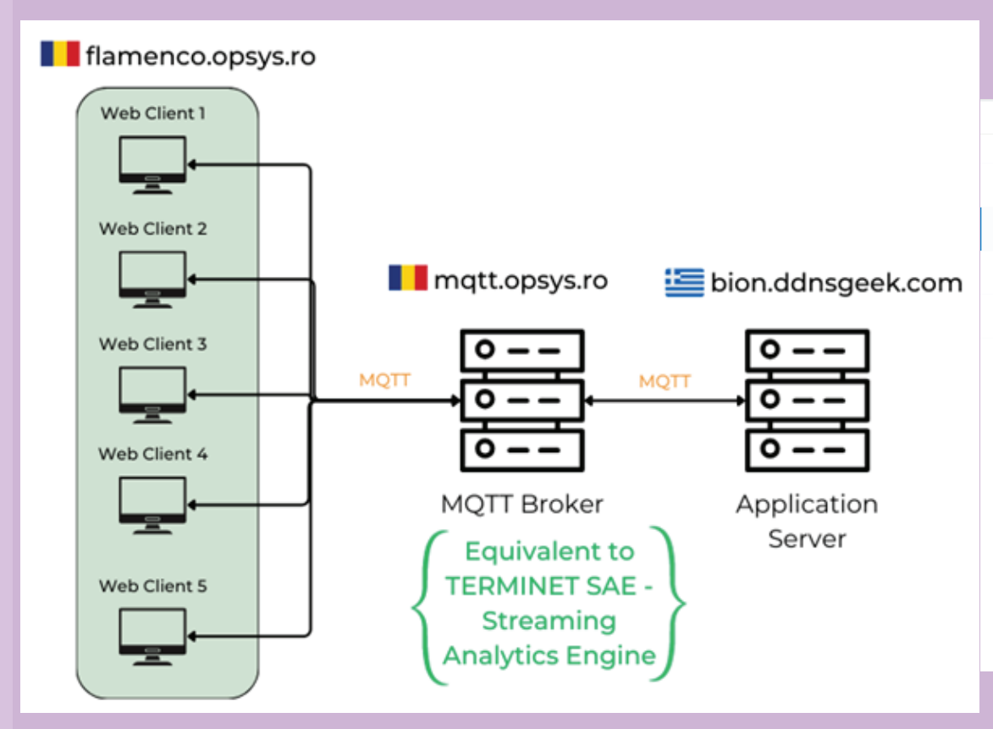
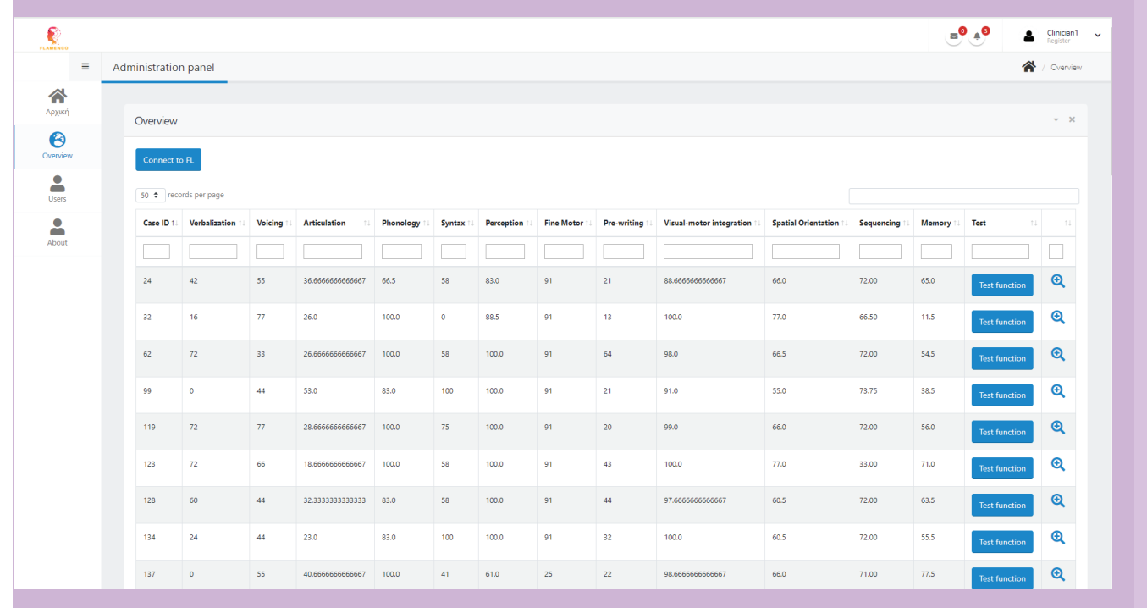


Figure 2: Application Architecture



Administration panel

Overview

Case ID	Verbalization	Voicing	Articulation	Phonology	Syntax	Perception	Fine Motor	Pre-writing	Visual-motor integration	Spatial Orientation	Sequencing	Memory	Test	
24	42	55	36.6666666666667	66.5	58	83.0	91	21	88.6666666666667	66.0	72.00	65.0	Test function	
32	16	77	26.0	100.0	0	88.5	91	13	100.0	77.0	66.50	11.5	Test function	
62	72	33	26.6666666666667	100.0	58	100.0	91	64	98.0	66.5	72.00	54.5	Test function	
99	0	44	53.0	83.0	100	100.0	91	21	91.0	55.0	73.75	38.5	Test function	
119	72	77	28.6666666666667	100.0	75	100.0	91	20	99.0	66.0	72.00	56.0	Test function	
123	72	66	18.6666666666667	100.0	58	100.0	91	43	100.0	77.0	33.00	71.0	Test function	
128	60	44	32.3333333333333	83.0	58	100.0	91	44	97.6666666666667	60.5	72.00	63.5	Test function	
134	24	44	23.0	83.0	100	100.0	91	32	100.0	60.5	72.00	55.5	Test function	
137	0	55	40.6666666666667	100.0	41	61.0	25	22	98.6666666666667	66.0	71.00	77.5	Test function	

Figure 3: Web Application Mockup

Client Selection

Under federated learning setting, managing communication costs is a crucial aspect of the system's efficiency. Thus, it is essential to implement selection strategies to minimize communication overhead without compromising the predictive performance of the underlying model. Towards this goal, we have implemented four client selection mechanisms:
1) Random Sampler: This selection algorithm involves the server randomly selecting a subset of clients during federated rounds. In our case, from a total pool of five clients, the server will randomly choose up to four clients.
2) Std Sampler: This algorithm selects clients based on the variability of their data, measured by the standard deviation. Clients are chosen inversely proportional to their standard deviation, i.e., clients with higher variability are less likely to be selected.
3) Quantity Sampler: In this mechanism, clients are chosen according to the quantity of their samples relative to the number of samples of the overall dataset. This method selects clients with a larger number of samples, thereby penalizing those with fewer samples.
4) IntelliSampler: A novel introduced client selection mechanism using clients' training loss and L2 client model divergence from the global.

Evaluation Results

We performed various experiments under different settings and seed for random generators in order to validate the project's results and showcase its potential. Below (Figs. 5-8), we illustrate the most representative outcomes of our research that validate our projects functionality and verify its added value to TERMINET's ecosystem in general.

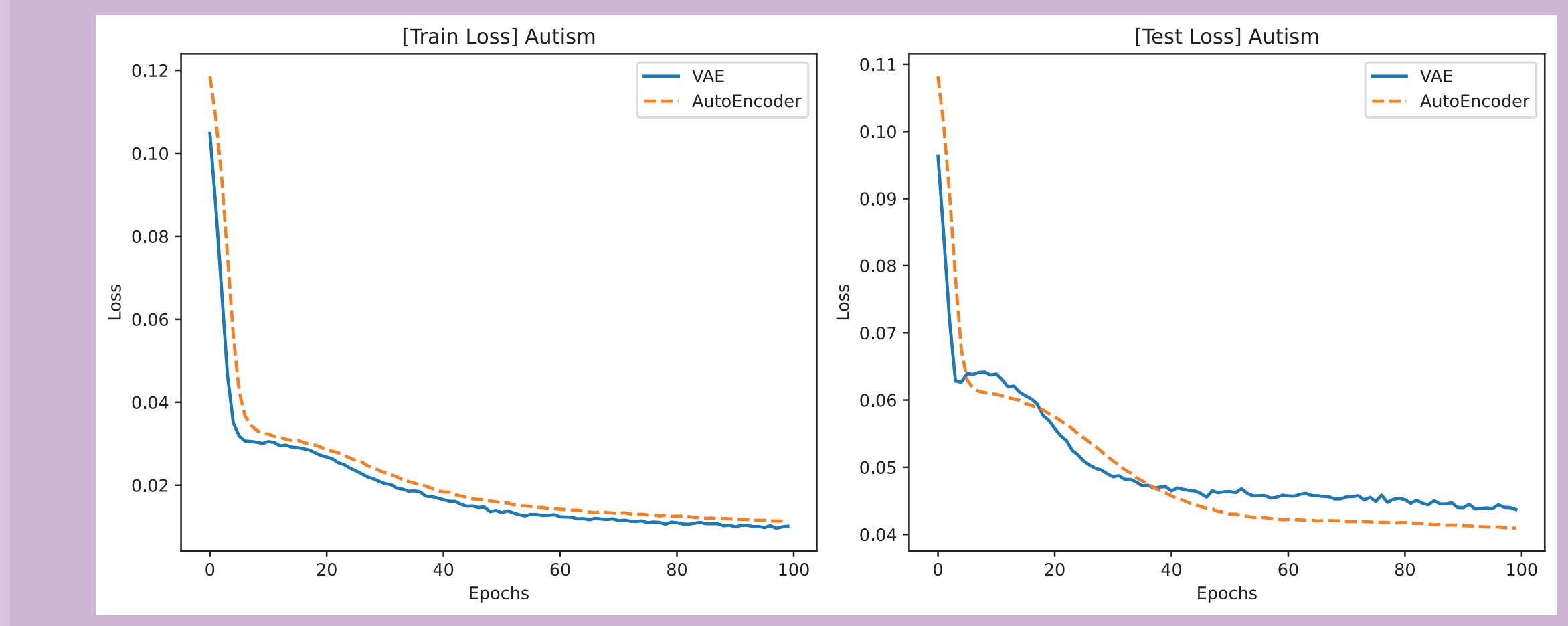


Figure 5: Comparing ML models architectures

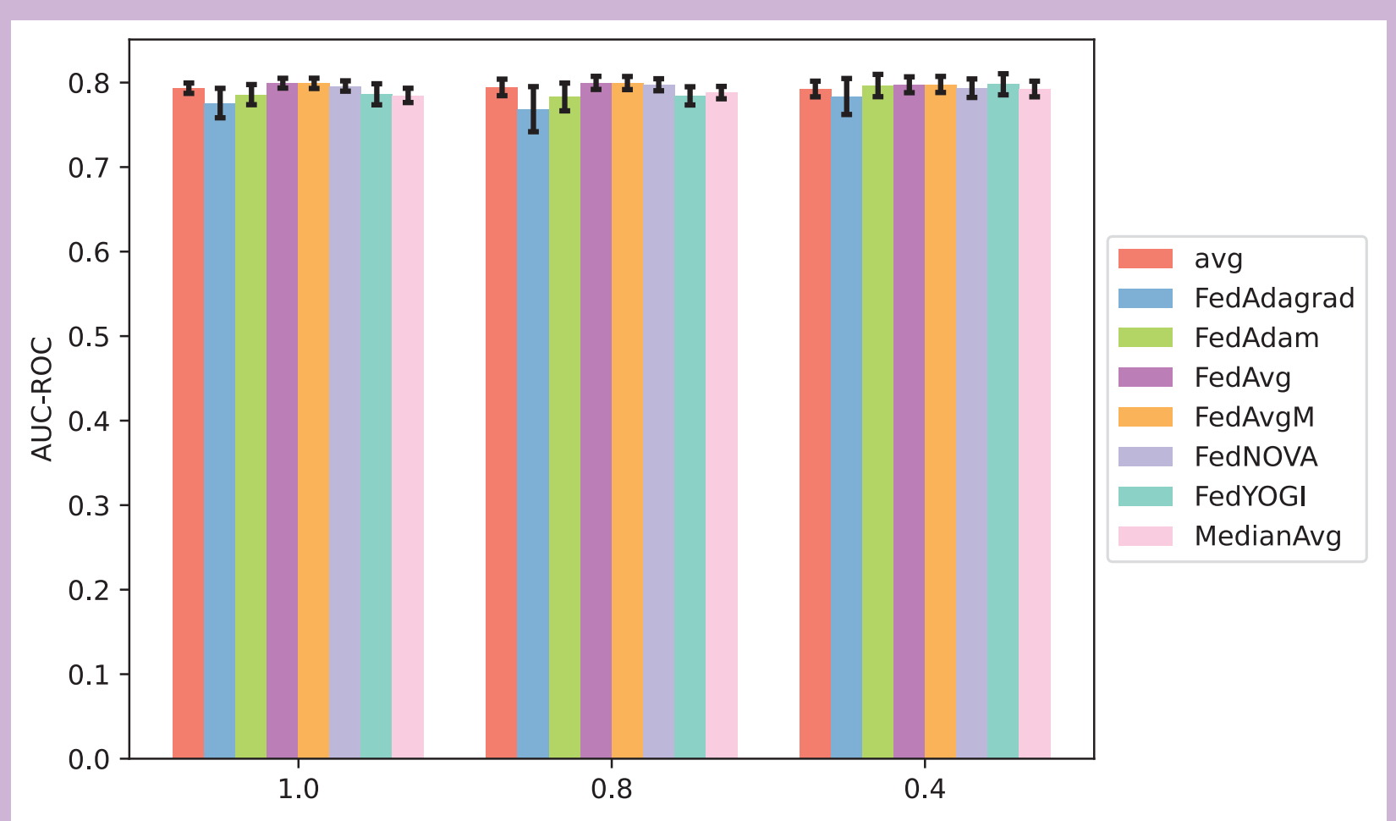


Figure 7: Overview comparison of aggregators

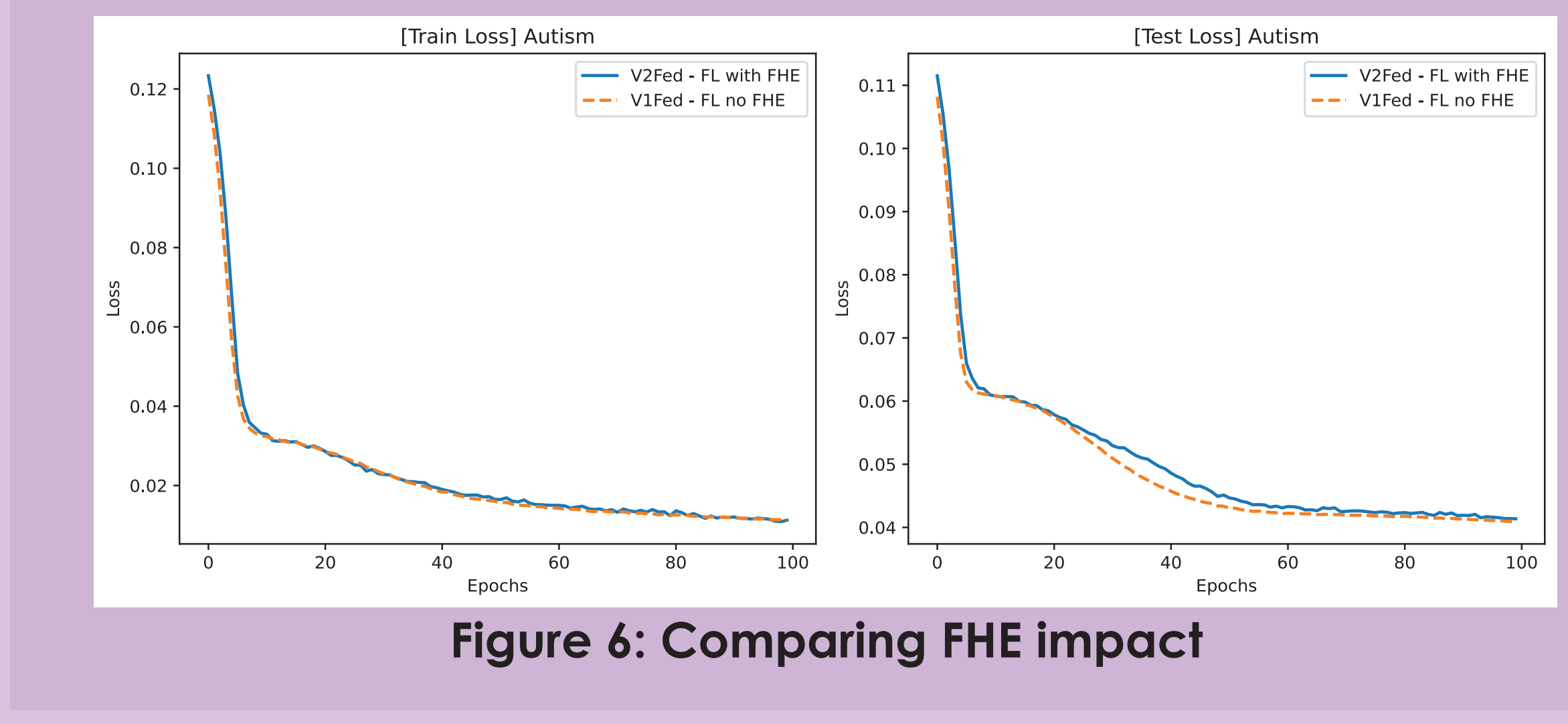


Figure 6: Comparing FHE impact

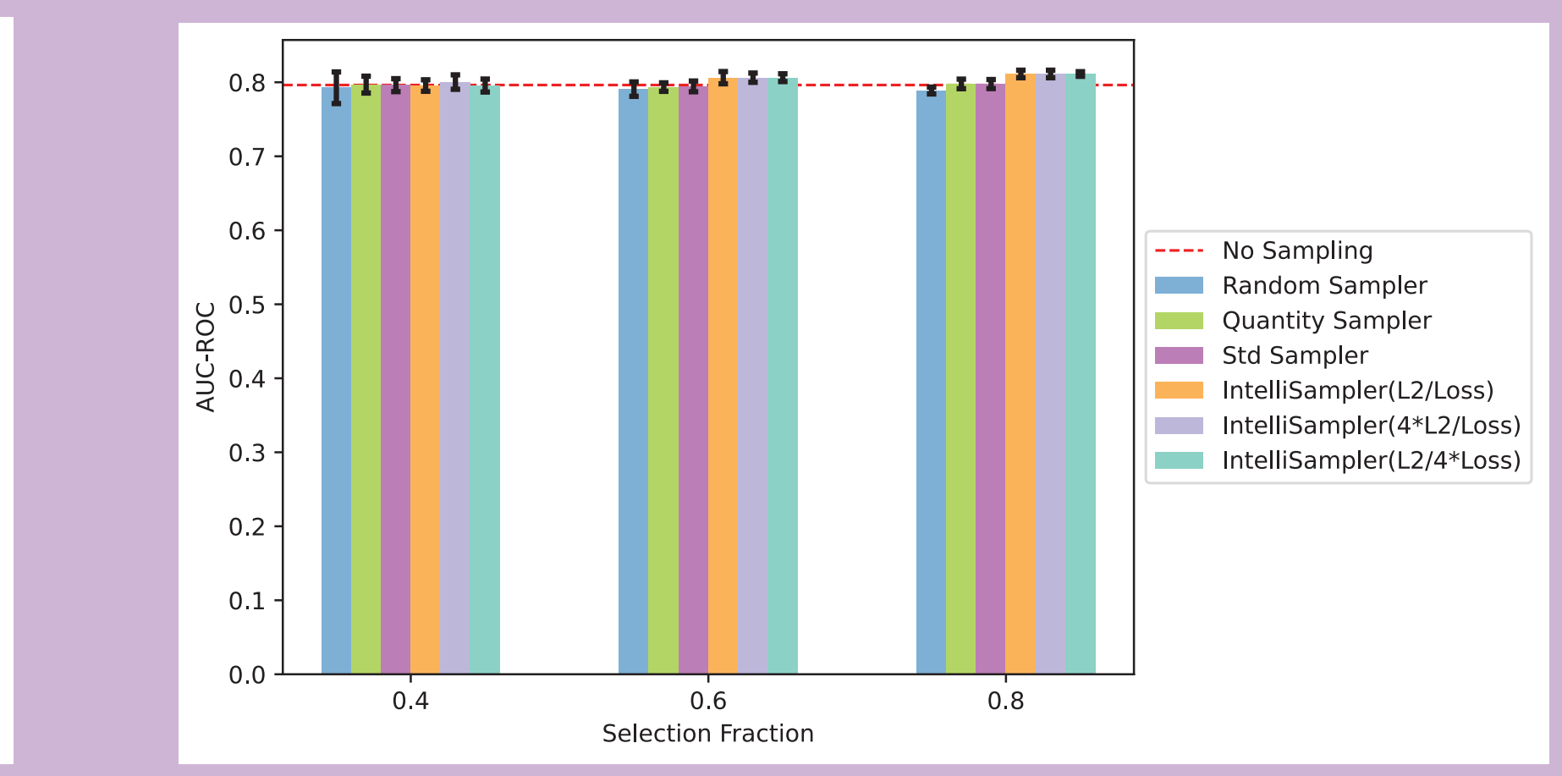


Figure 8: Overview comparison of client selection mechanisms

References

[1] FLAMENCO Dataset https://euclid.ee.duth.gr:25312/s/EyBGDfecX6g7i2m
[2] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In AI and statistics (pp. 1273-1282). PMLR.
[3] FLAMENCO, 2023, Github Repository. Available at: https://github.com/nikopavl4/FLAMENCO-Project
[4] Perifanis, V., Pavlidis, N., Yilmaz, S., Wilhelmi, F., Guerra, E., Miozzo, M., Efraimidis, P., Dini, P., & Koutsiamanis, R.A. (2023). Towards Energy-Aware Federated Traffic Prediction for Cellular Networks. In 2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC) (pp. 93-100).

TERMINET
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the TERMINET project with grant agreement No 957406

flamenco.opsys.ro

A collaboration of:

DEMOCRITUS UNIVERSITY OF THRACE | DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING