Federated Learning where machine learning and data privacy coexists
AML Project

Giuseppe Salvi — Basiten Bartholom — Alfredo Baldó Chamorro

February 2022

here we should describe the algorithms we used why lenet why average

Federated learning applied in the real world, would take into account multiple devices with different characteristics (CPU capacity and speed...). The model and algorithm selection was chosen counting on the variety of devices if the model would be deployed.

With that in mind, for the neural network we chose a short one: LeNet5 (put reference), with the following structure:

Regarding the aggregation of the loss function on the server side, we chose the Federated Average Algorithm (FedAvg) (PUT REFERENCES). The weights of the FedAvg where proportional to the number of images each client saw (PUT FORMULA?).

Moreover, in order to simulate real world data, we implemented into the code a variability regarding the distribution of the data among clients. The variable added some randomness to the quantity of data that each client would be seeing.