Federated Learning In Wireless Security

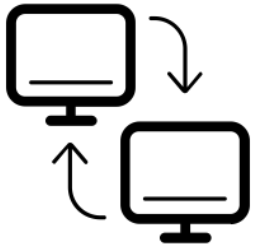
Supervisor: Prof. Harshan Jagadeesh

By:
Ekansh Singh(2020EE10490)
Tanish Singh Tak(2020EE10560)

Outline

- What is Federated Learning?
- Data Diversity
- Our Work
- Inference from the work
- Next Course of Action

Problem Statement

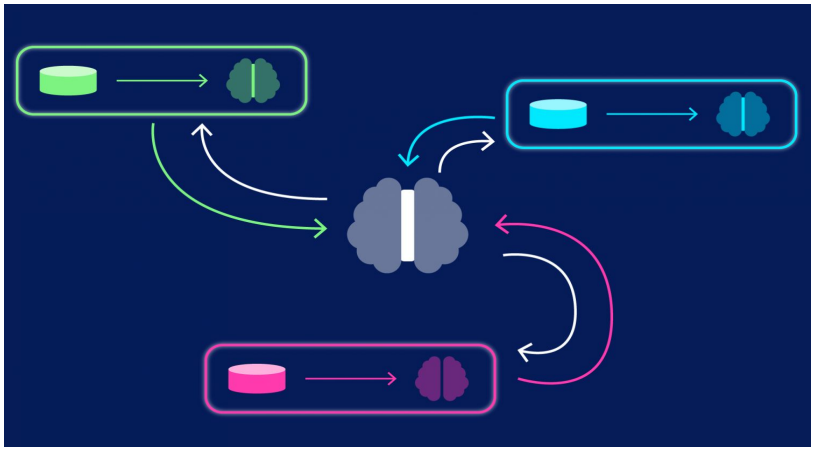


- “Our data is as important as our machine learning model”
- We need more diverse data points so as yield better results
- But the parties which possess these would not agree to provide the same without any incentive

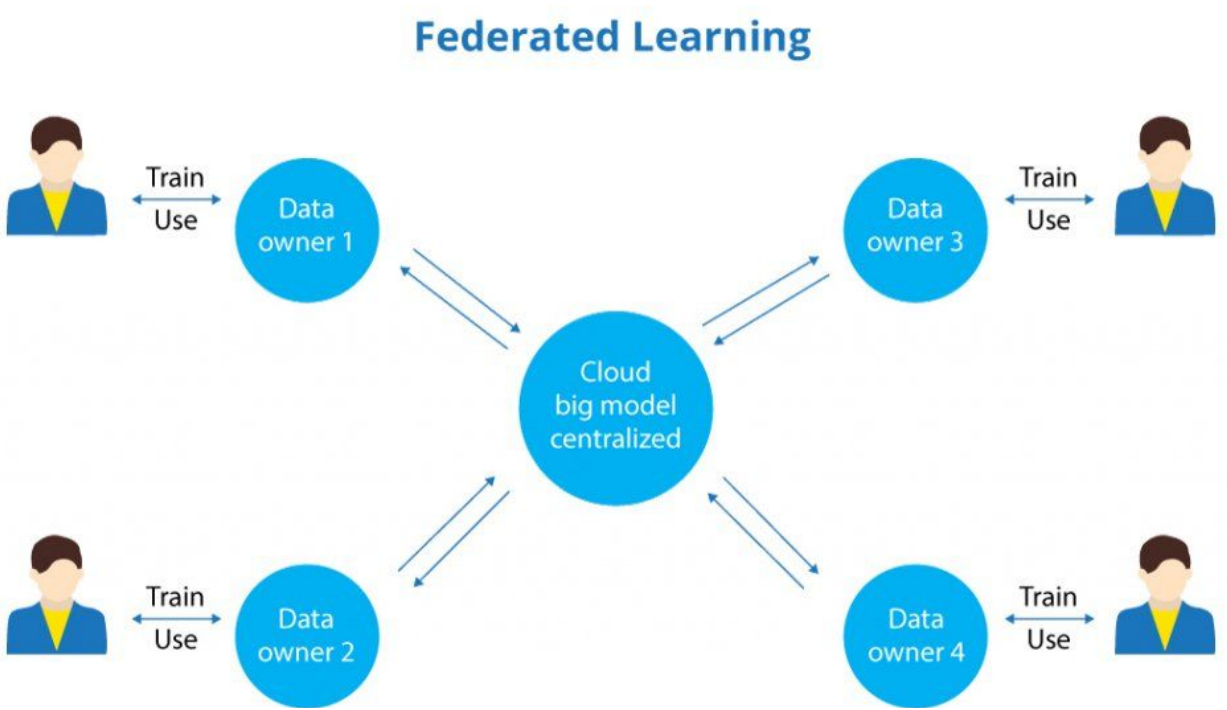
Hence we are interested to develop some sort of an algorithm or metric which could judge the encrypted version of the datapoint and provide a fair idea on how “diverse” is the new dataset

What is Federated Learning?

- It is a decentralized approach to train machine learning models which doesn't require any exchange of data from client to global servers instead we use the raw data on edge devices to train locally prioritizing privacy



Why Federated Learning?



Why do we care for Data Diversity

- **Enhanced Decision-Making:** Diverse datasets inform better decisions across business, governance, and domains, improving outcomes.
- **Bias Mitigation:** Lack of data diversity can perpetuate bias and discrimination. Diverse data helps identify and rectify biases, promoting fairness in AI and decision systems.
- **Robust Models**: Machine learning benefits from diverse datasets, leading to more adaptable and robust models.

Our Work

- Till now, we created a neural network using the MNIST dataset to observe the effects of data diversity and how sharing the weights and biases from one model to another can improve accuracy
- We created two datasets from the MNIST dataset, in which one of the data comprises the images for the numbers in the range zero to four and the second dataset includes the images in the range five to nine
- Then, we trained two separate models on these datasets and tested them on two datasets divided in the same manner as the train data and observed the accuracies

Inference from the work

- We also observed the accuracies for both test cases using an updated model created after merging the weights of both models

Model Type	Test dataset Type	Accuracy (in %)
Model_04	Dataset_04	98.988
Model_59	Dataset_59	98.478
Model_04	Dataset_04	0.0
Model_59	Dataset_59	0.0
Model_merged	Dataset_04	51.683
Model_merged	Dataset_59	62.292

Future works

- For the next part of the project, we would try to come up with a method to quantify the diversity between the two datasets.
- Then we will try to come up with a method through which one can share their data to another person in an encrypted way such that the receiver can calculate this diversity factor without getting the actual data.
- Following this, our final goal will be to extend this method of secure sharing of data to sharing of model's weights and biases to prevent the sharing of large size datasets.

Thank you