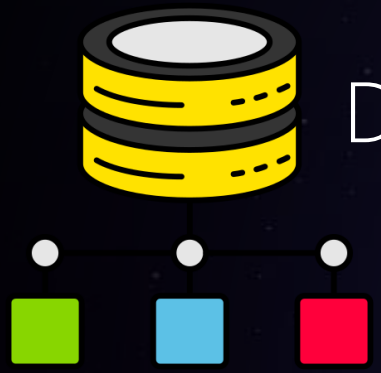
FEDERATIVE LEARNING

Sajjad Ranjbar

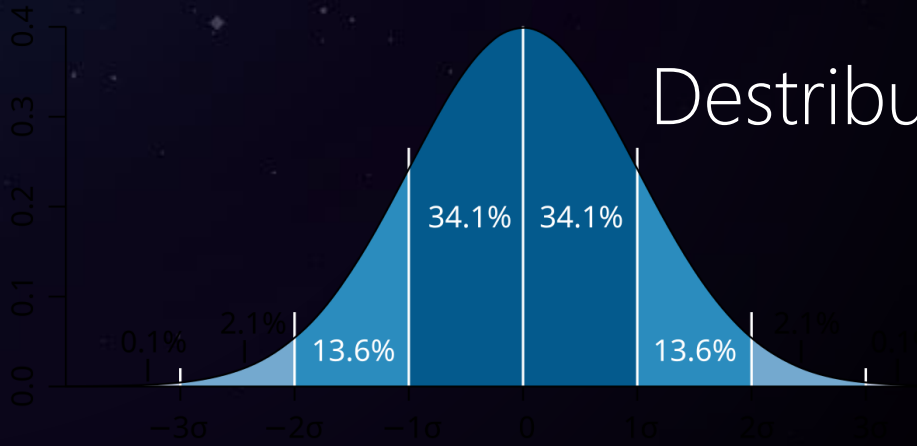
CLASSICAL MACHINE LEARNING



Dataset

model

Destribution



CLASSICAL MACHINE LEARNING

The volume of data is enormous

Dataset

Data is confidential

CLASSICAL MACHINE LEARNING

Dataset

The volume of data is enormous

A single server must compute all processes

Data is confidential

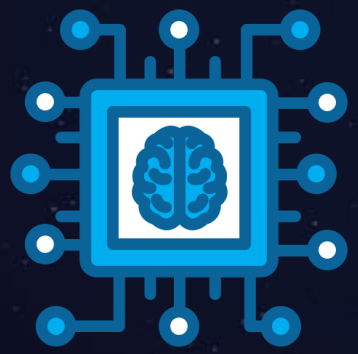
Medical information, banking transactions, etc.

FEDERATED LEARNING

Federated learning is a distributed machine learning approach where multiple devices or entities collaborate to train a shared model without exchanging raw data. This allows for data privacy and security while still enabling collaborative model improvement.

FEDERATED LEARNING

Server

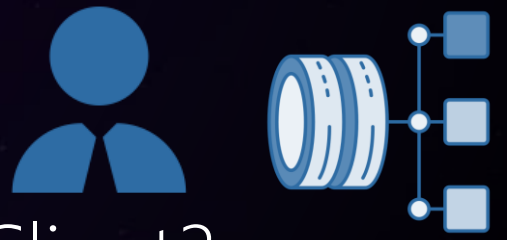


Global Model

Local data

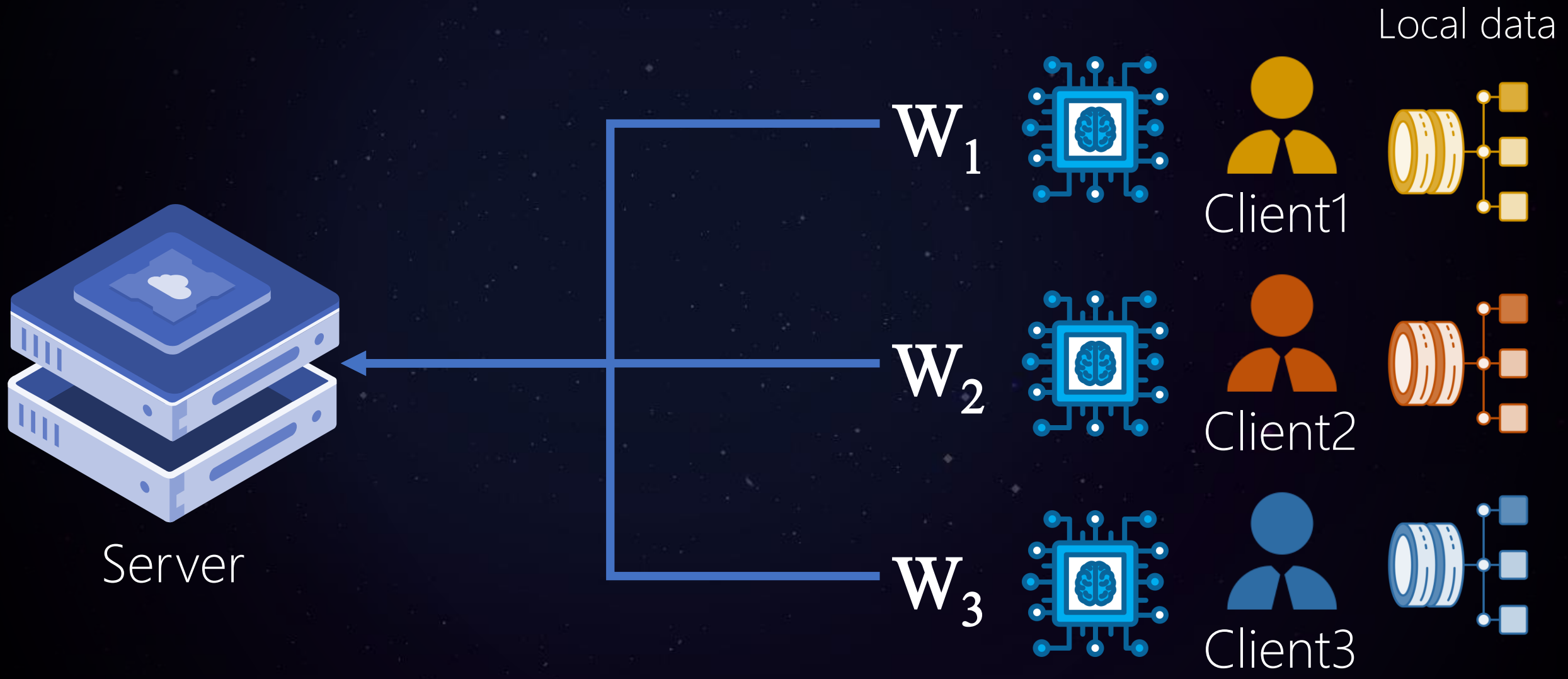
Client1

Client2



Client3

FEDERATED LEARNING



FEDERATED LEARNING

Server

$$W = F(W_1, W_2, W_3)$$

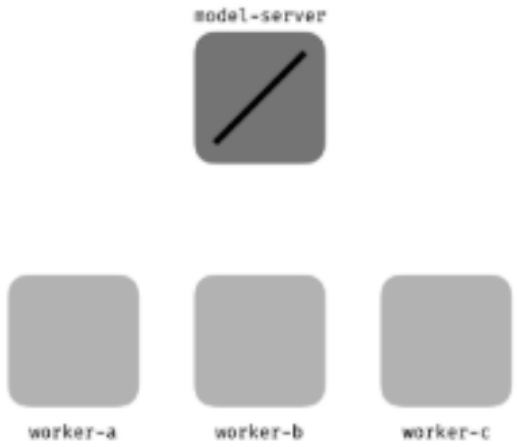
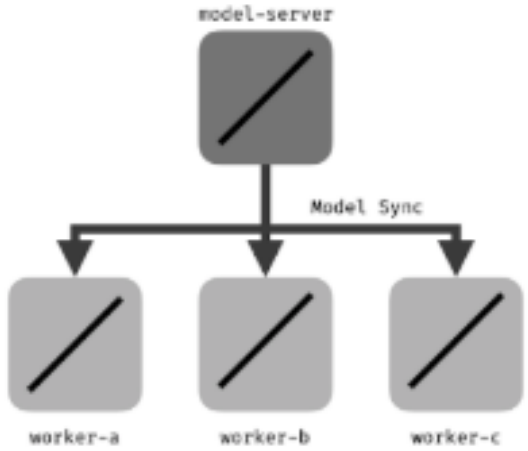
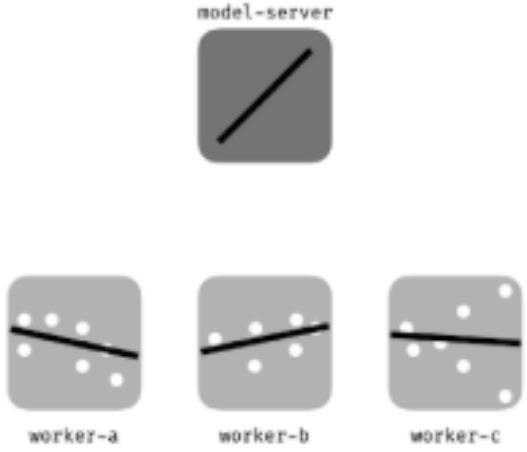
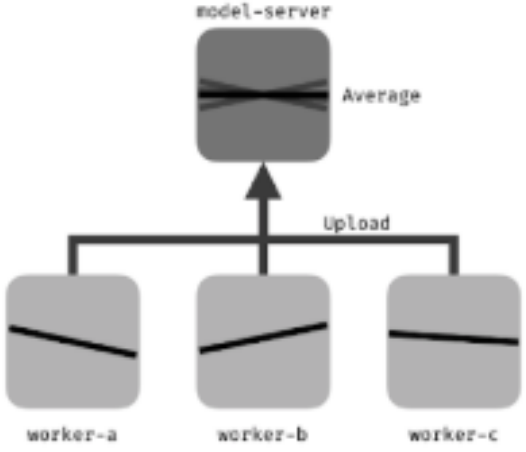
Local data

Client1

Client2

Client3

FEDERATED LEARNING

Step 1	Step 2	Step 3	Step 4
			
Central server chooses a statistical model to be trained	Central server transmits the initial model to several nodes	Nodes train the model locally with their own data	Central server pools model results and generate one global mode without accessing any data

FEDERATED LEARNING

Even in FL, model weights may provide the original data.

Solution: Encrypt weights before sending to the server.

Challenge: The server must be able to perform aggregation operations on encrypted weights.

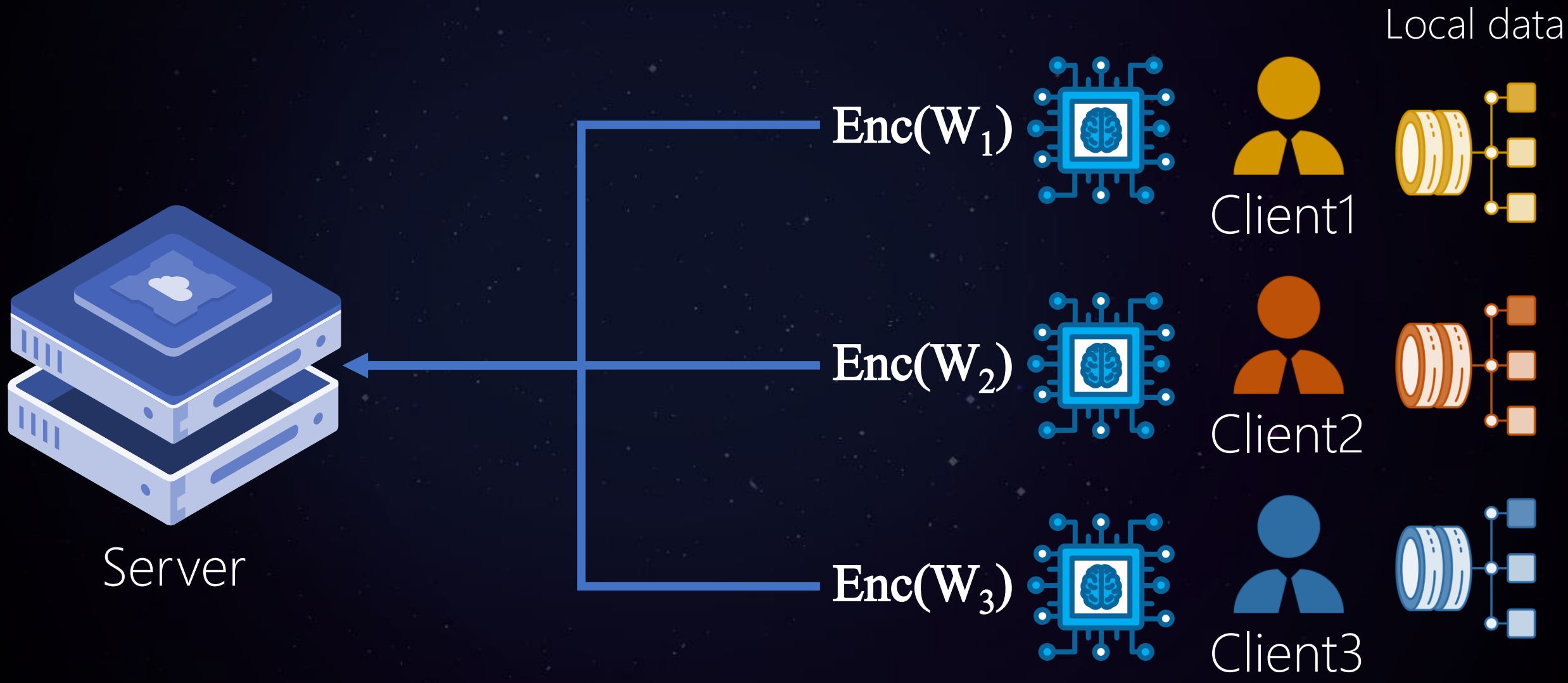
Homomorphic Encryption

HOMOMORPHIC ENCRYPTION

A type of encryption that allows calculations to be performed on encrypted data without the need to decrypt it.

$$\mathrm{Enc}(x \curlywedge y) = \mathrm{Enc}(x) \curlywedge \mathrm{Enc}(y)$$

FEDERATED LEARNING



FEDERATED LEARNING

Server

Local data

Client1

Client2

Client3

$$W = F(Enc(W_1), Enc(W_2), Enc(W_3))$$

FEDERATED LEARNING

Real example

id	text	label
1	…	Suicide
2	…	Non-suicide

Link

EXAMPLE

Many Classical models have been tested on this dataset. You can see a good example in this notebook. The following methods have been tested on this netbook:

- Naive Bayes (Voting Classifier)
- Random Forest
- Decision Tree
- Gradient Boosting
- XG Boost

FEDERATED LEARNING BASED ON NAIVE BAYES

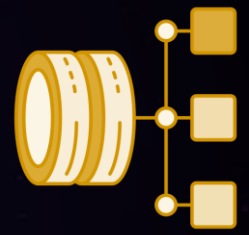
Server

Naive Bayes

W_1

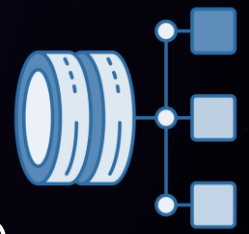
Client1

Local data



W_{20}

Client20



$$W = \frac{1}{20}\sum_{i=1}^{20} w_i$$

RESULTS

	Classical model	Federated model(20 clients)
Train Accuracy	88.02%	-
Test Accuracy	88.06%	88.53
Precision	0.86	0.89
Recall	0.91	0.88
F1-score	0.88	0.89

Thank you so much