Federated Learning

& securing communication with Solidity

A diabetes dataset case

Created and Presented by : ***Abdelatif Mekri***

Table of contents

01

Introduction

Federated Learning

a decentralized machine learning approach that enables multiple entities to collaboratively train a model without sharing raw data. Unlike traditional centralized learning methods, where data is aggregated in a single location,

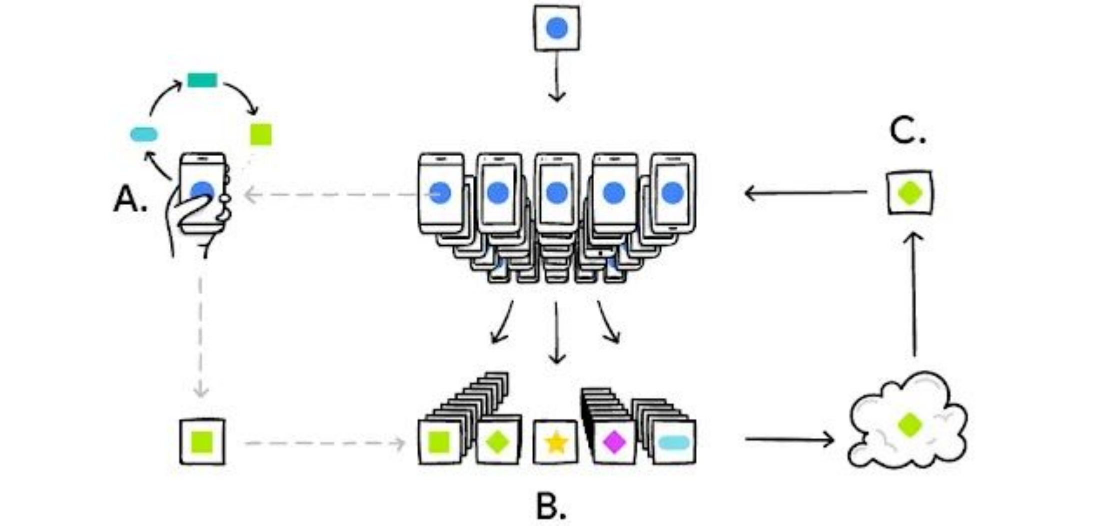
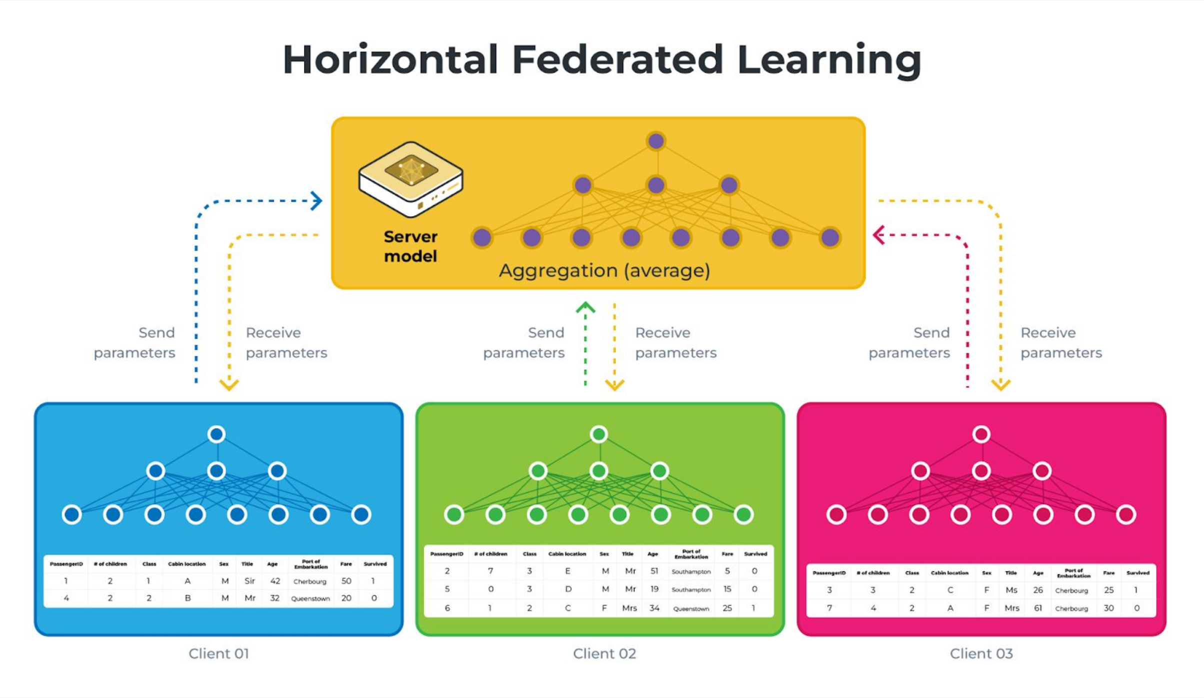


Figure 1. Simple showcase of federated workflow

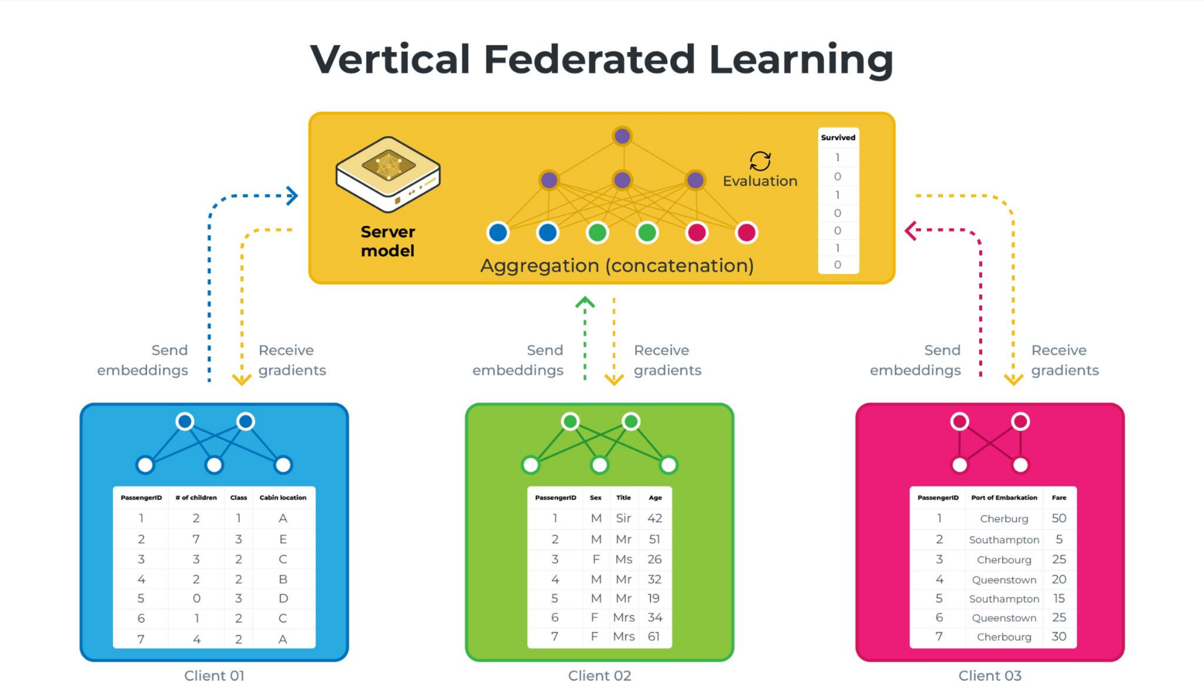
Horizontal Federated Learning (HFL)

Clients have datasets with the same feature space but different user samples. This setup is common when organizations operate in the same industry and collect similar types of data.



Vertical Federated Learning (VFL)

Clients share common user samples but have different feature spaces. This setup is useful when organizations hold complementary data about the same users.



Advantages of Federated learning (FL)

Privacy Preservation

Data remains on local devices, avoiding direct sharing of raw sensitive information

Regulatory Compliance

Facilitates adherence to strict data protection laws (e.g., HIPAA) by minimizing data transfer and central storage

Scalability

Enables training across thousands of devices without centralized infrastructure bottlenecks

Enhanced Security

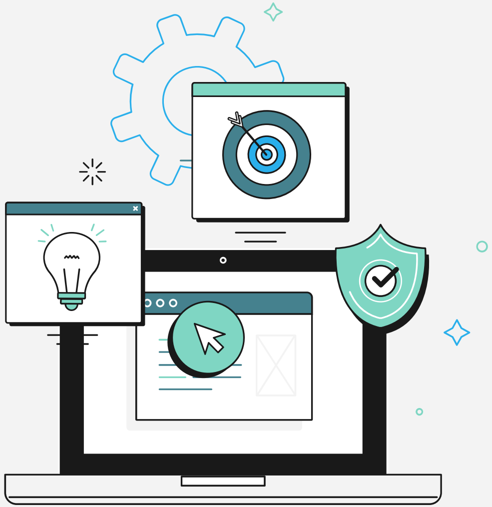
Decentralized data storage limits attack surfaces

Fault Tolerance

Resilient to device dropouts or network issues, as training continues with available participants

User Trust and Participation

Increased user willingness to contribute to model training when privacy is assured, enhancing dataset diversity and model quality



02

Problematic & Dataset

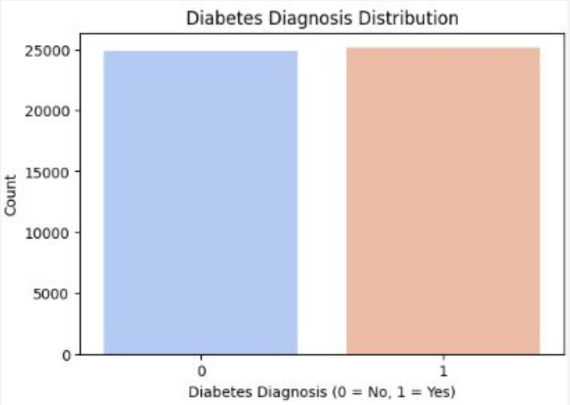
Why choosing Diabetes Dataset ?

Diabetes prediction relies on sensitive medical data, making **centralized machine learning** impractical due to **privacy risks and data-sharing restrictions**. **Federated learning** enables decentralized model training without exposing patient data, but it introduces **security concerns** in communication and participant authentication. To address this, **blockchain** ensures a **secure, transparent, and tamper-proof** consent mechanism, enhancing trust and data integrity in the training process.

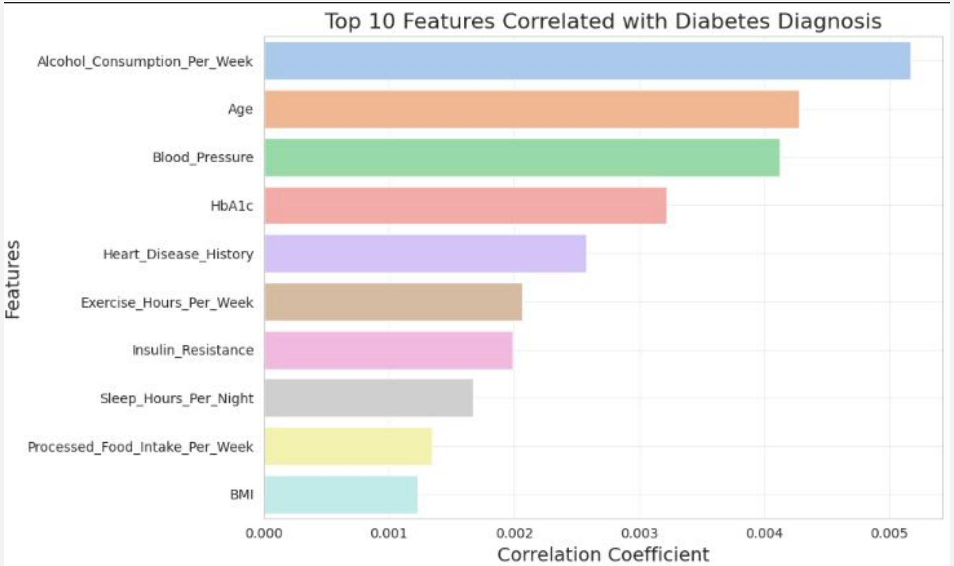
Dataset

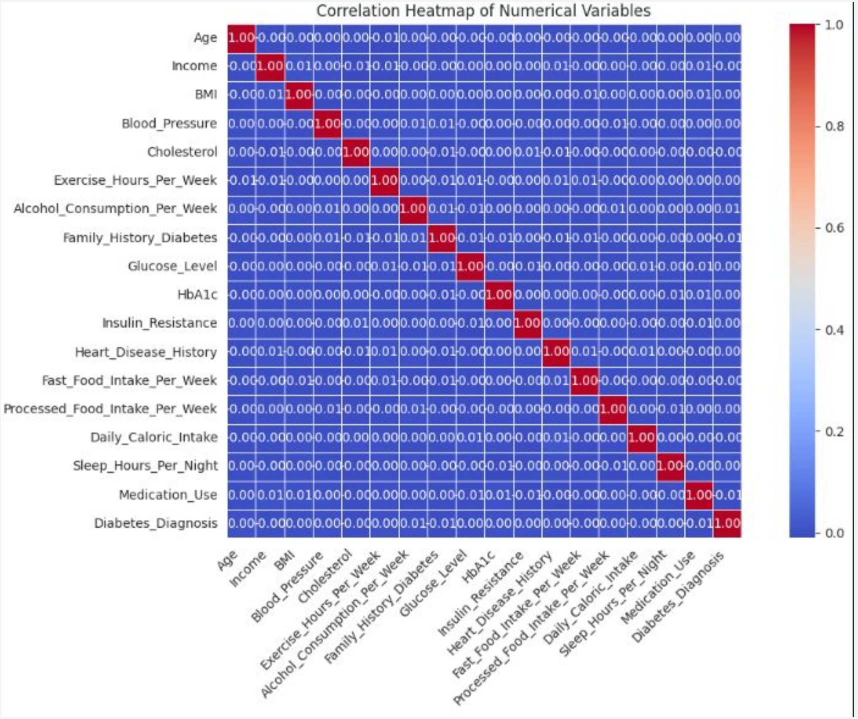
	Age	Gender	Ethnicity	Income	BMI	Blood_Pressure	Cholesterol	Exercise_Hours_Per_Week	Alcohol_Consumption_Per_Week	Smoking_Status	...	Insulin_Resistance	Heart_Disease_History	Physical_Activity_Level
0	69	Female	Other	39557	38.2	94.6	252.9	3.3	4	Never	...	5.1	0	Low
1	32	Male	Black	90663	33.6	167.0	282.6	4.6	7	Never	...	1.7	1	Moderate
2	89	Male	White	116180	39.4	100.6	106.8	6.1	5	Former	...	4.9	1	Low
3	78	Male	Other	73059	40.6	111.1	169.7	7.4	9	Never	...	9.8	0	High
4	38	Female	White	35389	29.7	143.3	296.5	2.6	6	Never	...	1.7	1	Moderate

5 rows × 23 columns



EDA on the Dataset





03

Federated Learning

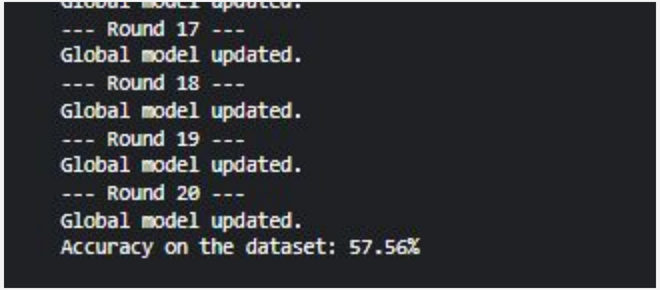
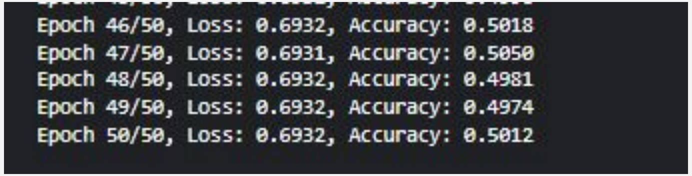
Federated Learning

Federated Learning allows multiple clients (e.g., hospitals, organizations) to collaboratively train a model without sharing their raw data. I explored two types of federated learning

Horizontal	vertical

`num_clients = 3  # Change as needed`
`client_partitions = partition_dataset(dataset, num_clients)`
	`num_features = len(feature_cols)`
	`features_A = feature_cols[:num_features // 2]`
	`features_B = feature_cols[num_features // 2:]`

Federated Learning results

Horizontal	vertical
 --- Round 17 --- Global model updated. --- Round 18 --- Global model updated. --- Round 19 --- Global model updated. --- Round 20 --- Global model updated. Accuracy on the dataset: 57.56%	 Epoch 46/50, Loss: 0.6932, Accuracy: 0.5018 Epoch 47/50, Loss: 0.6931, Accuracy: 0.5050 Epoch 48/50, Loss: 0.6932, Accuracy: 0.4981 Epoch 49/50, Loss: 0.6932, Accuracy: 0.4974 Epoch 50/50, Loss: 0.6932, Accuracy: 0.5012

04

Utilising Blockchain

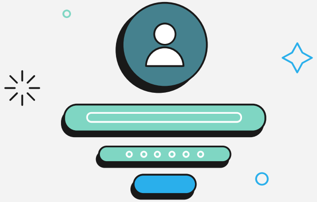
Solidity on Remix IDE

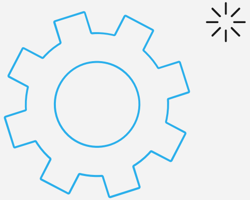
Writing contract

This Solidity smart contract, **FLConsent**, is designed to manage participant consent in a federated learning scenario

Compiling contract

-make sure to use the right compiler settings (**compiler** and **EVM** version)





05

Conclusion

Conclusion

Horizontal federated learning

Enables collaboration between organizations with the **same features** but **different users** (e.g., hospitals sharing patient records).
Learning How to train models across multiple parties without sharing raw data, preserving privacy while improving model accuracy.

Vertical federated learning

Allows different entities with **different features** but **the same users** (e.g., banks and e-commerce platforms) to collaborate.
Learning Secure feature-sharing techniques like homomorphic encryption and privacy-preserving methods to train models on complementary data.

Blockchain

nsures **trust, security, and decentralization** in federated learning by maintaining an immutable and transparent ledger.
Learning how blockchain enhances federated learning through secure aggregation, smart contracts for automation, and decentralized consensus mechanisms.