Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	27 October 2023
Team ID	Team-592697
Project Name	Diabetes Prediction using Machine Learning
Maximum Marks	4 Marks

Technical Architecture:

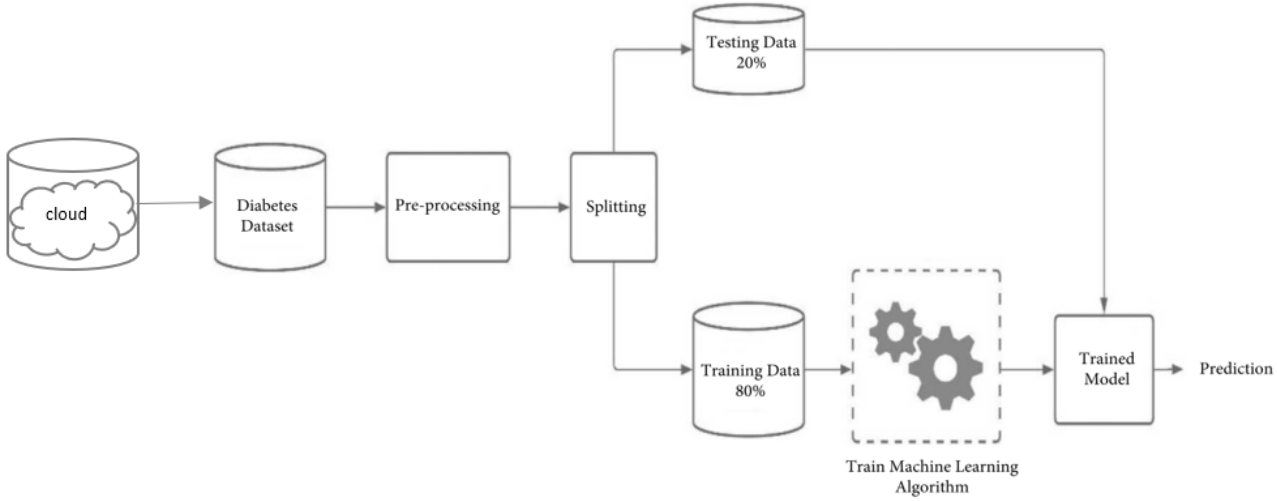

Table-1: Components & Technologies:

S.No	Component	Description	Technology
1.	Data Collection	Fetch and store data from various sources	Cloud – based storage(eg: AWS S3), data collection tools
2.	Data Storage	The data is often stored in a suitable format, such as a database or a distributed data storage system. This allows for easy retrieval and access by the ML model during training and prediction phases.	Cloud – based data bases(eg: AWS RDS), local data bases.
3.	Data Preprocessing	Raw data is often messy and unstructured. Data preprocessing involves cleaning, transforming, and organizing the data to make it suitable for ML analysis. This step includes handling missing values, removing outliers, normalizing or scaling features, and encoding categorical variables.	Python (Pandas, NumPy), Data Cleaning Tools
4.	Feature Engineering	Feature engineering is a crucial aspect of data ingestion for ML. It involves selecting or creating relevant features from the data that will be used to train the model. In the case of diabetic prediction, features might include patient demographics, medical history, lifestyle factors, and biomarker measurements.	Python (Pandas), Feature Engineering Tools .
5.	Machine Learning Models	Build and train ML models for Diabetic Prediction.	Python (Scikit-Learn, ARIMA), ML Frameworks
6.	Web Application (Django)	Provide a user interface for accessing data.	Python (Django), Web Development Tools
7.	Security and Compliance	Ensure the security and compliance of sensitive data, including customer information.	encryption, authentication, and access control measures.
8.	IBM Deployment (Cloud):	Cloud-based infrastructure, such as IBM Cloud, for hosting the Django application andmachine learning models.	IBM Cloud services for hosting and scaling the application.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Utilize open-source frameworks for development, machine learning, and data analysis.	Python, Scikit-Learn, Django)
2.	Security Implementations	Implement security measures to protect data and user interactions within the application.	SSL/TLS, Encryption, Authentication
3.	Scalable Architecture	Design the architecture to be scalable, allowing the application to handle growing data and user loads.	Cloud Services (e.g., AWS Auto Scaling), Load Balancing
4.	Availability	Ensure high availability of the application, minimizing downtime and disruptions	Redundancy, Failover, Monitoring andAlerting
5.	Performance	Optimize application performance for responsiveness and efficient use of resources	Caching, Database Indexing, EfficientAlgorithms.