Project Design Phase-II

Technology Architecture

Date	05 November 2022
Team ID	PNT2022TMID14106
Project Name	Project - Visualizing and Predicting Heart Diseases with an Interactive Dash Board

Technology Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

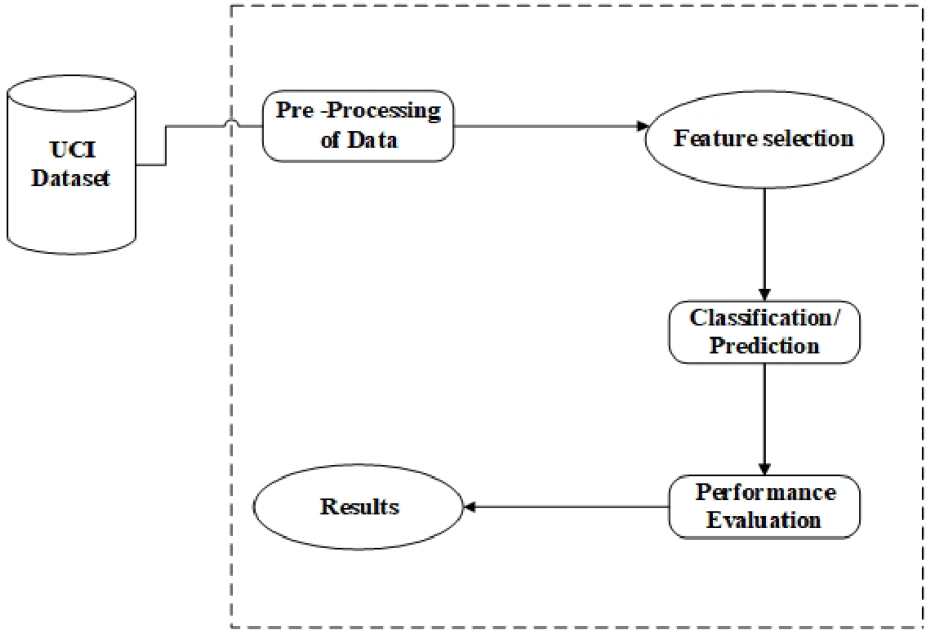

FIGURE 1. Experiment workflow with UCI dataset.

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	Fetching Data	Import Data infers the data type of each column based on the values it contains, and loads the data into your designer pipeline. The output of Import Data is a dataset that can be used with any designer pipeline.	Python , Numpy , Pandas
2.	Data Cleaning	Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.	Python
3.	Data Pre-processing	Data preprocessing is the concept of changing the raw data into a clean data set. The dataset is preprocessed in order to check missing values, noisy data, and other inconsistencies before executing it to the algorithm.	Python
4.	Training data	Training data is an extremely large dataset that is used to teach a machine learning model. Training data is used to teach prediction models that use machine learning algorithms how to extract features that are relevant to specific business goals.	Python
5.	Testing data	Data created or selected to satisfy the execution preconditions and input content required to execute one or more test cases.	Python
6.	Machine Learning model	A machine learning model is a program that has been trained to recognize certain types of patterns. You train a model over a set of data, providing it an algorithm that it can use to reason over and learn from those data.	Python
7.	Checking Accuracy	Model accuracy is defined as the number of classifications a model correctly predicts divided by the total number of predictions made. It's a way of assessing the performance of a model, but certainly not the only way.	Python

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Open source framework is a software for which the original source code is made freely available and may be redistributed and modified according to the requirement of the user.	Technology of Open source framework
2.	Security Implementations	Security analytics combines data from the various sources and looks for correlations and anomalies within the data. A security analytics tool may use different methods for analyzing the data. These include traditional rules-based methods, as well as statistical analysis and machine learning.	SHA-256, Encryptions, IAM Controls, OWASP etc.
3.	Scalable Architecture	Scalability is the property of a system to handle a growing amount of work by adding resources to the system. In other words, a scalable architecture supports higher workloads without any fundamental changes to it	Microservices 'Smart Edpoints and Dump pipes' , AWS Lambda, API Gateway
4.	Availability	Availability is achieved through load balancing multiple content hosts and failover support. Load Balancer acts as the enterprise's single point-of-presence on the Internet.	Server load Balancers
5.	Performance	Caching allows you to efficiently reuse previously retrieved or computed data. The data in a cache is generally stored in fast access hardware such as RAM and a cache is a high-speed data storage layer which stores a subset of data.	Cache