PROJECT DESIGN PHASE-II

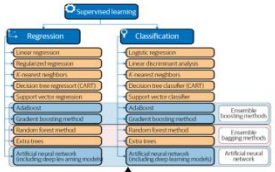
TECHNOLOGY STACK (ARCHITECTURE & STACK)

Date	15 October 2022
Team ID	PNT2022TMID39608
Project Name	Efficient Water Quality Analysis and Prediction using Machine Learning.
Maximum Marks	4 Marks

TECHNICAL ARCHITECTURE:

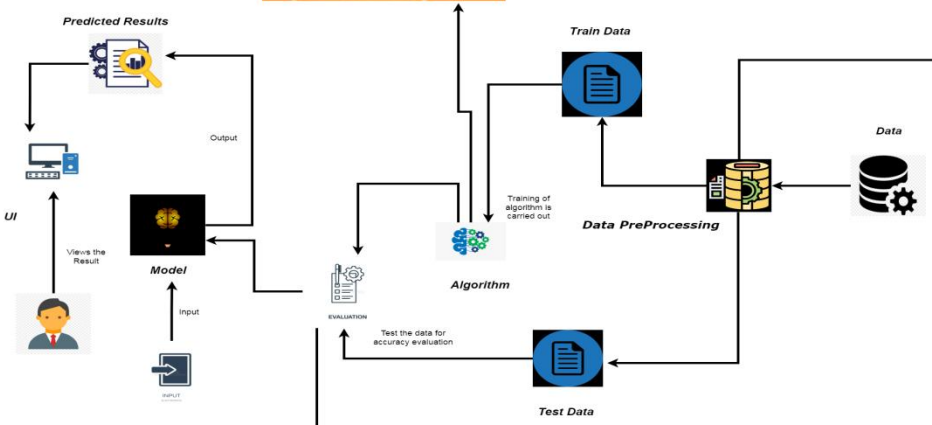
TECHNICAL ARCHITECTURE

METHODOLOGY APPLIED

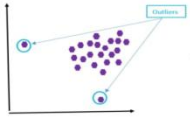


MODELS USED FOR PREDICTION

CORRELATION ANALYSIS



DATA EXPLORATION



OUTLIER DETECTION

DATA NORMALISATION

STEPS IN DATA PREPROCESSING

REGRESSION EVALUATION METRICS (FOR WQI PREDICTION)

CLASSIFICATION EVALUATION METRICS (FOR WQC PREDICTION)

EVALUATION TECHNIQUES APPLIED

Table-1: COMPONENTS & TECHNOLOGIES:

S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Variety of frameworks, libraries and supports are required to develop the project.	Java / Python
3.	Application Logic-2	Helps in predicting the Water Quality Index (WQI) using various Regression and Water Quality Classification using various Classification algorithms based on various parameters involved. It also helps in predicting the potability of water samples and also recommends various purification methods based on the impurities present in the water sample.	IBM Watson STT service, Machine Learning Algorithms.
4.	Application Logic-3	Provides fast, accurate and consistent results of water quality analysis and interprets the results in a easy understandable manner.	IBM Watson Assistant
5.	Database	It can be numerical, categorical or time series data.	MySQL, NoSQL, etc.
6.	Cloud Database	Enables the user to host the database on his/her own hardware without buying additional hardware.	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage should be highly flexible, scalable, effective and a reliable one.	IBM Block/Object Storage or Other Storage Service or Local Filesystem
8.	External API-1	Used to access the information in the cloud.	IBM Weather API, etc.
9.	External API-2	Used to access the information for data driven decision making.	Aadhar API, etc.

10.	Machine Learning Model	Purpose of Machine Learning Model	Regression and Classification Model, etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Install the windows version and execute the installer.	Local, Cloud Foundry, Kubernetes, etc.

Table-2: APPLICATION CHARACTERISTICS:

S. No	Characteristics	Description	Technology
1.	Open-Source Frameworks	The Frameworks used in the project are:	Anaconda Navigator, Tensor Flow, Keras, Flask.
2.	Security Implementations	The security / access controls are implemented using firewalls, etc.	SHA-256, Encryptions, IAM Controls, OWASP etc.
3.	Scalable Architecture	The scalability of architecture (3 – tier, Micro-services)	Data, models operate at different sizes, speed, consistency and complexity.
4.	Availability	The availability of application (e.g. use of load balancers, distributed servers etc.)	It can be availed by all kinds of customers who wish to test the quality of water they consume.
5.	Performance	Design aspects for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Gives correct and effective prediction, easy accessibility to the results using Machine Learning.