**Project Design Phase-I**

**Solution Architecture**

|  |  |
| --- | --- |
| Date | 15 November 2023 |
| Team ID | Team-591881 |
| Project Name | River Water Quality Forecasting |
| Maximum Marks | 4 Marks |

**Solution Architecture:**

Solution Architecture: River Water Quality Forecasting

1. Problem Statement:

The business problem is to forecast river water quality accurately to support environmental management, public safety, and sustainable water usage.

2. Goals:

Find the best tech solution for water quality forecasting.Describe the structure, characteristics, and behavior of the software. Define features, development phases, and solution requirements. Provide specifications for solution definition, management, and delivery.

3. Components and Processes:

3.1 Data Collection:

Sources: Monitoring stations, weather stations, satellite imagery.

Parameters: Chemical concentrations, physical characteristics, meteorological data.

3.2 Data Preprocessing:

Clean and preprocess data, handle missing values, and engineer features.

Transform data for model compatibility.

3.3 Training Data Split:

Split the dataset into training and validation sets.

3.4 Machine Learning Model (Random Forest):

Train a Random Forest Classifier using historical water quality data.

Evaluate the model on the validation set (accuracy, precision, recall, F1-score).

3.5 Real-Time Data Integration:

Integrate real-time data from monitoring stations for live predictions.

3.6 Prediction:

Utilize the trained Random Forest model for predicting current water quality.

Implement algorithms for efficient and real-time predictions.

3.7 Alerts and Notifications:

Establish an alert system to notify stakeholders in case of predicted water quality issues.

3.8 Visualization (Dashboard):

Develop a user-friendly dashboard for different user groups.

Include interactive maps, charts, and historical trends.

3.9 Continuous Improvement:

Implement a feedback loop for continuous model improvement.

Update the model based on new data and user feedback.

4. Technology Stack:

Programming Language: Python

Data Processing: Pandas, NumPy

Machine Learning: scikit-learn

Web-Hosting: Flask

5. Specifications:

Define clear specifications for data sources, model parameters, real-time data integration, and visualization requirements.

6. Documentation:

Provide detailed documentation for all components, processes, and technologies used in the solution. Include user manuals and guides for stakeholders.

7. Security and Compliance:

Implement security measures for data privacy and integrity. Ensure compliance with relevant environmental and data protection regulations.

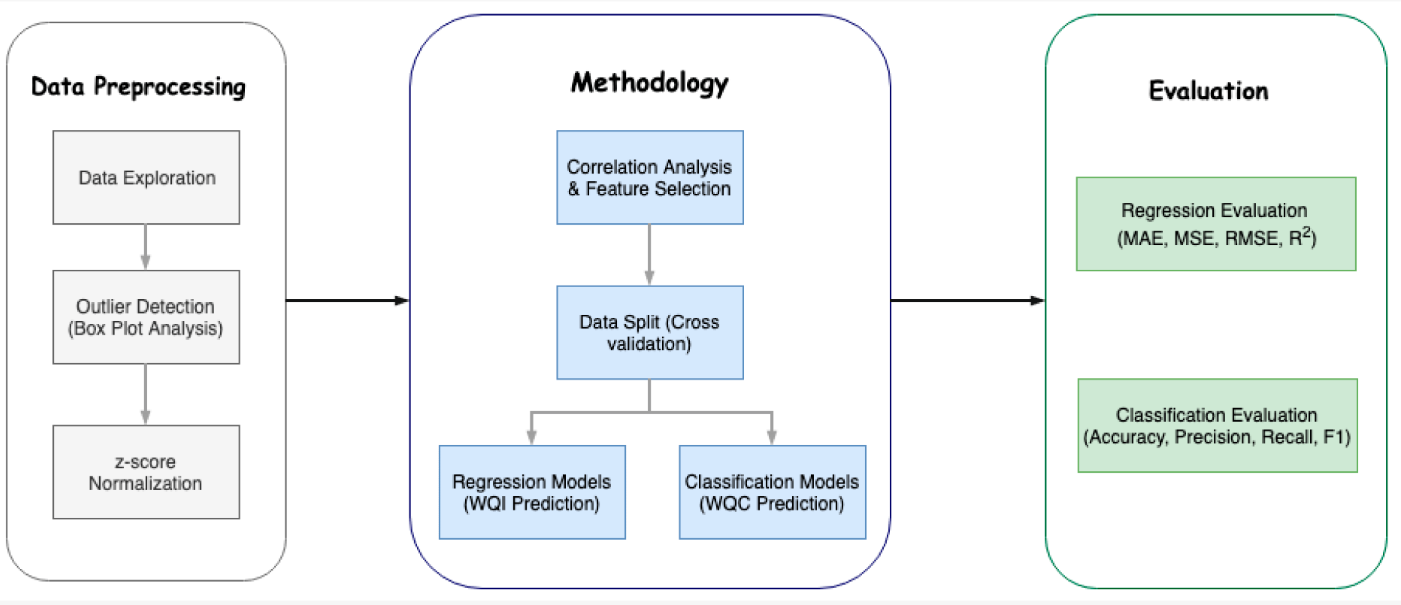
8. Scalability and Performance:

Design the solution to scale horizontally and vertically to handle increasing data and user demands. Optimize algorithms and infrastructure for performance.

9. Monitoring and Maintenance:

Implement monitoring tools for system performance. Establish regular maintenance procedures for data updates, model retraining, and system improvements.

**Solution Architecture Diagram:**

****