**Water Quality Monitoring in Kutch Region - Omdena Project**

## 1. Introduction

Water is a fundamental resource for all forms of life, and monitoring its quality is essential to ensure the health and safety of both humans and the environment. The "Water Quality Monitoring" project, conducted under the Omdena platform, focuses on the Kutch region in India. It aims to assess and predict the water quality of water bodies using satellite imagery, machine learning techniques, and data visualization dashboards. This report summarizes the complete workflow, contributions, methodologies, and results from all the major tasks undertaken during the project lifecycle.

## 2. Objectives

* Identify and list essential water quality parameters and their standards.
* Review literature to understand previous methodologies and best practices.
* Collect satellite imagery and relevant data from sources like Google Earth Engine.
* Analyze satellite imagery and derive water indices using remote sensing techniques.
* Identify important water quality parameters and associated spectral band formulas.
* Apply machine learning models to predict water quality indices based on collected data.
* Create an interactive dashboard for the public to visualize and interpret the results.

**Project Overview**

This project aims to monitor the **water quality** of water bodies in the **Kutch region** using **remote sensing data from Google Earth Engine (GEE)** and predictive **machine learning models**. The final objective is to create a **dashboard** that allows users to select a region of interest and visualize predicted water quality parameters across different time frames.

**Project Dashboard**

A publicly available dashboard allows users to select:

* **Type:** Area of interest (e.g., Water Body)
* **Coordinates:** Longitude (72.6026), Latitude (23.0063)
* **Time Range:** With constraints (min. 15 days, max. 3 months)

**Project Structure**

**├── LICENSE**

**├── README.md**

**├── original**

**├── reports**

**│ └── README.md**

**├── src**

**│ ├── data**

**│ ├── docs**

**│ ├── references**

**│ ├── tasks**

**│ ├── visualizations**

**│ └── results**

* **original**: Historical reference code from the original Omdena challenge.
* **reports**: Final reports and summaries.
* **src/data**: Datasets collected during the project.
* **src/docs**: Presentations, documents, and diagrams.
* **src/references**: Manuals, research papers, etc.
* **src/tasks**: Contains all individual task folders and their contributions.
* **src/visualizations**: Data visualizations and dashboards.
* **src/results**: Final machine learning results and model outputs.

**Contributions by Task**

**Task 1: Water Quality Standards**

* Listed standard parameters related to water quality.
* Categories include:
  + Inorganic parameters (e.g., Nitrate, Nitrite)
  + Organic parameters (e.g., BOD, COD)
  + Radiological and Microbiological indicators
  + Acceptability parameters (taste, odour, appearance)

**Task 2: Literature Review**

* Reviewed relevant academic and scientific literature.
* Identified different satellite sources like **Sentinel-2**, **MODIS**, **Landsat-8**.
* Compiled information on spectral band formulas for water quality indices.
* Explored previous machine learning applications to water quality estimation.

**Task 3: Satellite Data Collection**

* Utilized Google Earth Engine (GEE) for satellite data retrieval.
* Selected relevant datasets based on:
  + Spatial resolution
  + Availability in the region
  + Spectral bands supporting water quality analysis

**Task 4: Satellite Image Analysis**

* Preprocessing of satellite images.
* Raster to vector conversion for segmentation of water bodies.
* Applied standard formulas for water quality-related color indices:
  + NDWI (Normalized Difference Water Index)
  + Turbidity index
  + Chlorophyll-a estimation

**Task 5: Water Quality Parameters Identification**

* Finalized a list of measurable water quality parameters derivable from remote sensing:
  + Turbidity
  + Chlorophyll-a
  + NDVI/NDWI
* Documented the formulae for each index and how to compute them from satellite bands.

**Task 6: Machine Learning Modeling**

* Built predictive models for water quality estimation.
* Tested and evaluated multiple models:
  + Linear Regression
  + Random Forest
  + Gradient Boosting
  + XGBoost
* Selected the best-performing model based on validation accuracy and generalization

 Evaluation metrics used:

* MAE (Mean Absolute Error)
* RMSE (Root Mean Square Error)
* R² Score (Coefficient of Determination)

.

**Task 7: Dashboard Development**

* Created a web-based dashboard using **Streamlit** and **Folium**.
* Dashboard capabilities:
  + Allows users to input area and dates
  + Fetches satellite data from GEE
  + Visualizes predicted water quality indicators

**Key Technologies Used**

* **Google Earth Engine (GEE)** - for satellite data access and image processing
* **Python (Streamlit, Pandas, Scikit-learn, Folium)** - for modeling and dashboard
* **GeoPandas, Shapely, Geopy** - for geographical data handling
* **Seaborn, Matplotlib** - for plotting and visualization
* **Joblib** - for model persistence

## Outcomes & Impact

* Created a scalable pipeline to extract and process satellite imagery.
* Successfully predicted water quality parameters with high accuracy.
* Deployed an easy-to-use public dashboard.
* Provided a low-cost, satellite-based water monitoring solution.
* Set groundwork for further expansion to other Indian districts or countries.

**Future Work**

* Integrate real-time data from IoT sensors and WSNs.
* Add temporal forecasting for water quality prediction.
* Extend project to other regions and expand dashboard capabilities.

## References

* WHO Guidelines for Drinking Water Quality
* Central Pollution Control Board (CPCB) Reports
* Research Papers on Remote Sensing and Water Quality Monitoring
* Google Earth Engine and Sentinel/Landsat Documentation

## 10. Conclusion

This project showcases a data-driven approach to monitor and predict water quality using satellite imagery and machine learning. By leveraging the power of open data and collaborative intelligence, this initiative provides a cost-effective and scalable method for environmental monitoring in water-scarce regions like Kutch. The dashboard acts as a bridge between complex data analysis and actionable public information

**Contributors**

This project was built as part of the Omdena AI challenge and involved collaboration between data scientists, engineers, environmentalists, and designers.