RAJALAKSHMI
ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

Department of Computer Science and Engineering

WATER QUALITY PREDICTION

Mrs M. Divya M.E

220701186 | Niyathi S

Problem Statement and Motivation

Problem Statement

1. **Inconsistent and delayed water quality assessments** hinder timely identification of contaminated water sources, posing risks to public health and the environment.

2. **Manual water testing methods are often time-consuming and resource-intensive**, making it challenging to monitor water quality at scale and in real time.

Motivation

1. **Ensuring access to clean and safe water** is essential for human health, economic development, and sustainable ecosystems.

2. **Leveraging machine learning for predictive analysis** enables faster, more accurate, and cost-effective water quality monitoring, aiding proactive decision-making.

Existing System

- **Enhanced Accuracy Through Machine Learning Models**
  Unlike traditional systems that rely on fixed thresholds or limited rule-based checks, this project uses advanced machine learning algorithms trained on historical data to predict water quality more accurately, even under varying environmental conditions.

- **Utilization of Multiple Input Parameters**
  By considering a wide range of water quality indicators (such as pH, turbidity, dissolved oxygen, conductivity, and more), the model achieves a more comprehensive understanding of water health, leading to **more stable and reliable predictions**.

Objectives

- To design and implement a machine learning model capable of predicting water quality based on physicochemical parameters such as pH, turbidity, conductivity, and dissolved oxygen.
- To build a user-friendly system that allows real-time or near-real-time analysis of water samples without requiring extensive laboratory resources.
- To support public health and environmental safety by enabling early detection of water contamination through data-driven insights.

Abstract

- Access to clean and safe water is essential for human health and environmental sustainability. Traditional water testing methods are often time-consuming, costly, and require laboratory infrastructure. This project aims to develop a machine learning-based system to predict water quality using easily measurable parameters such as pH, turbidity, conductivity, and dissolved oxygen. By training models on historical data, the system can classify water as safe or unsafe with high accuracy. The goal is to provide a fast, affordable, and scalable solution for early detection of water contamination, supporting proactive water management in both urban and rural areas.

Proposed System

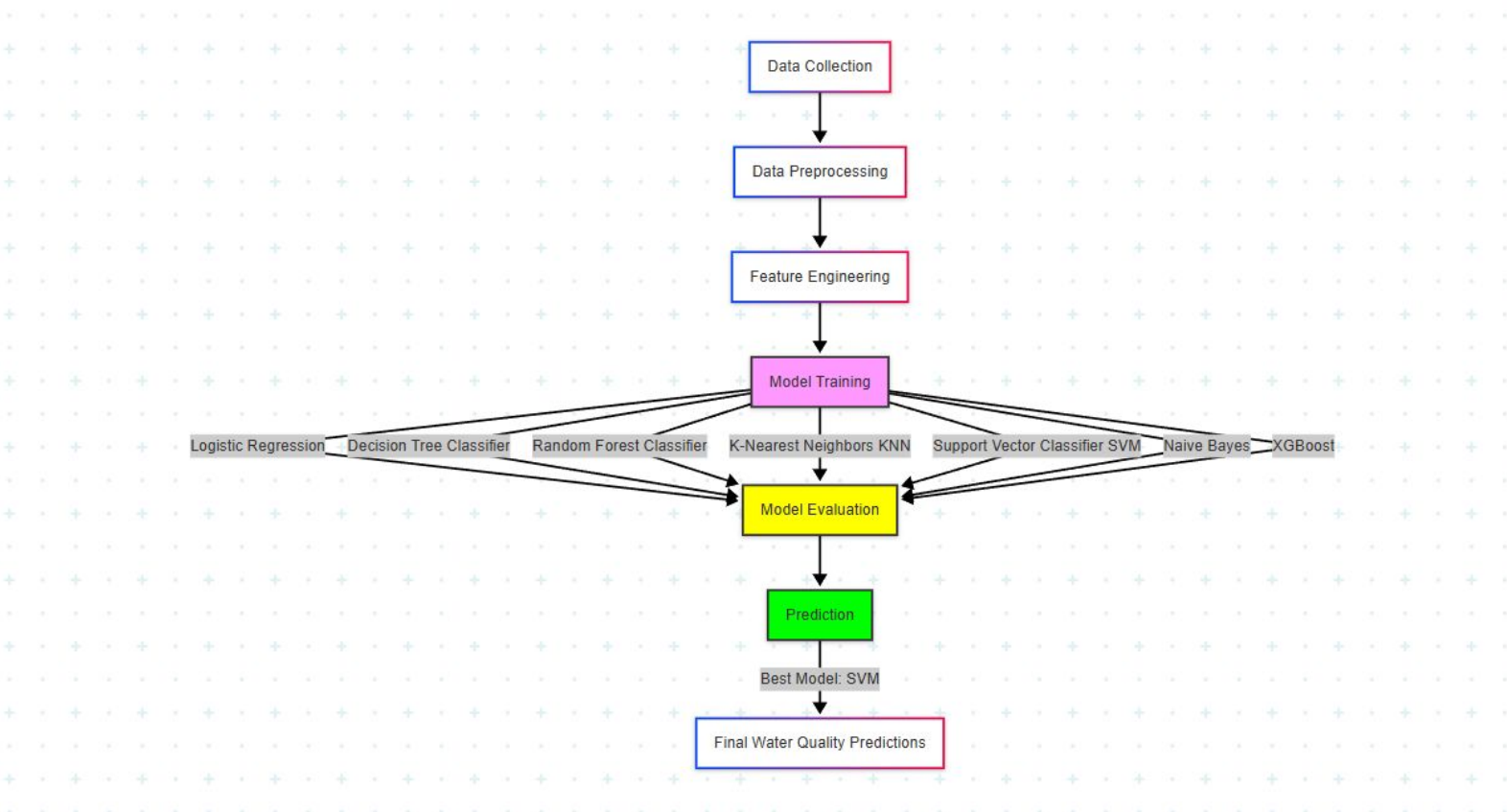
Data-Driven Prediction Model

Utilizes machine learning algorithms (e.g., Random Forest, Decision Tree) trained on historical water quality datasets to accurately classify water as "Potable" or "Not Potable" based on various physicochemical parameters.

Feature Engineering and Preprocessing

The solution performs data cleaning, handles missing values, and applies standardization to ensure high model performance and robustness across varying water samples.

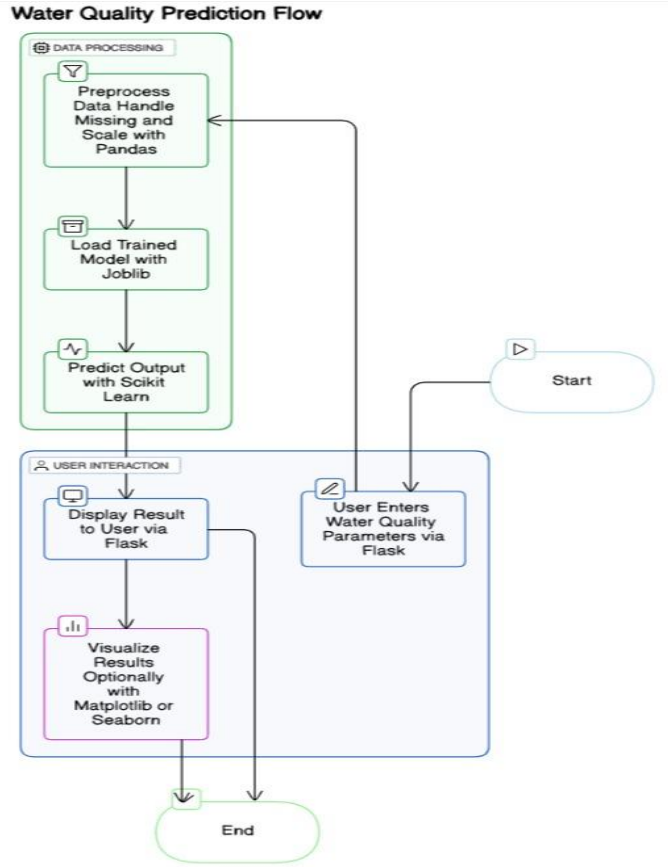
System Architecture

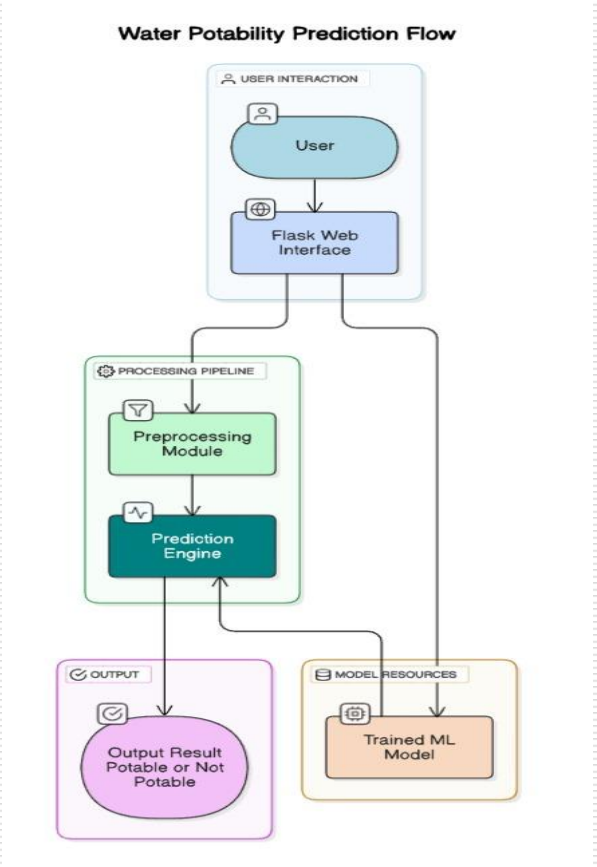


List of Modules

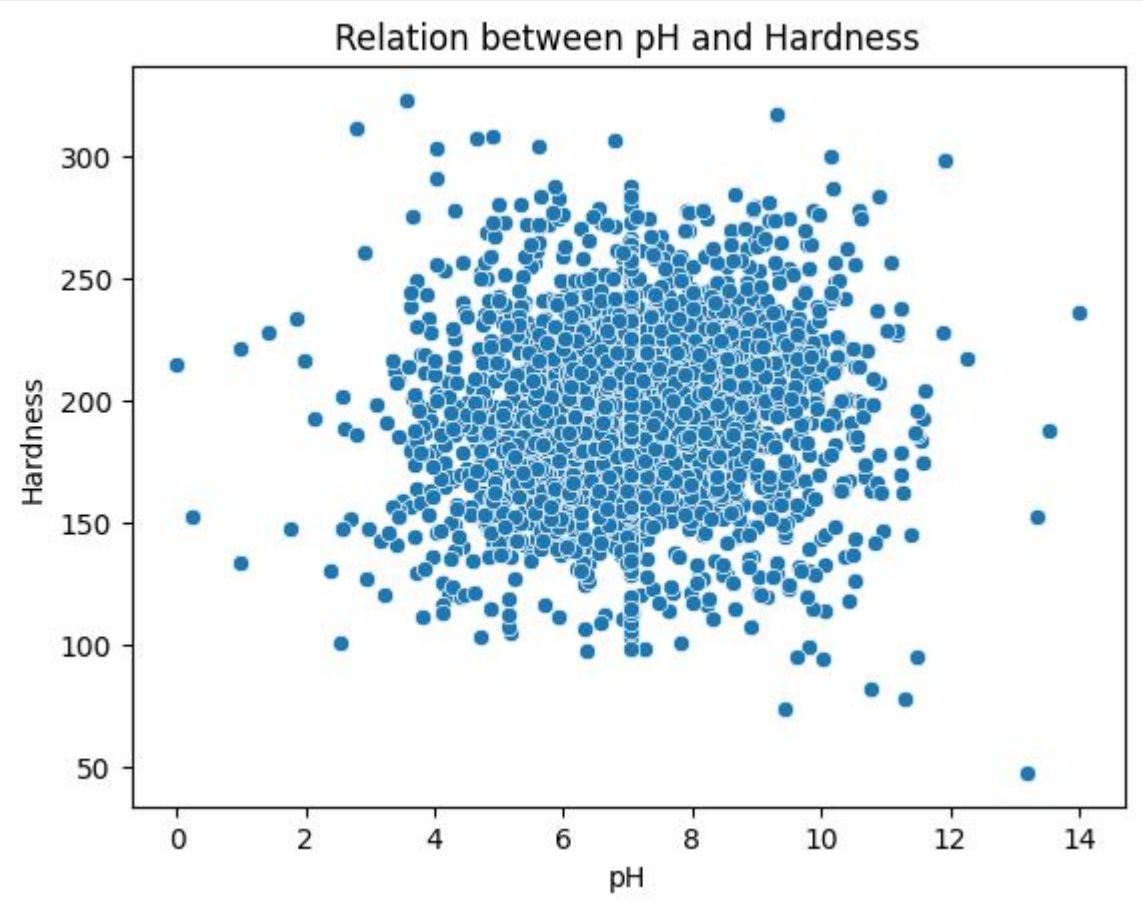
- flask – web deployment framework for applications
- pandas – data handling and analysis tool
- scikit-learn – machine learning and model building
- matplotlib – data visualization through static plots
- seaborn – advanced statistical data visualizations
- joblib – model saving and loading utility
- gunicorn – production server for Flask apps
- notebook – run and edit Jupyter notebooks
- nbconvert – convert notebooks to other formats

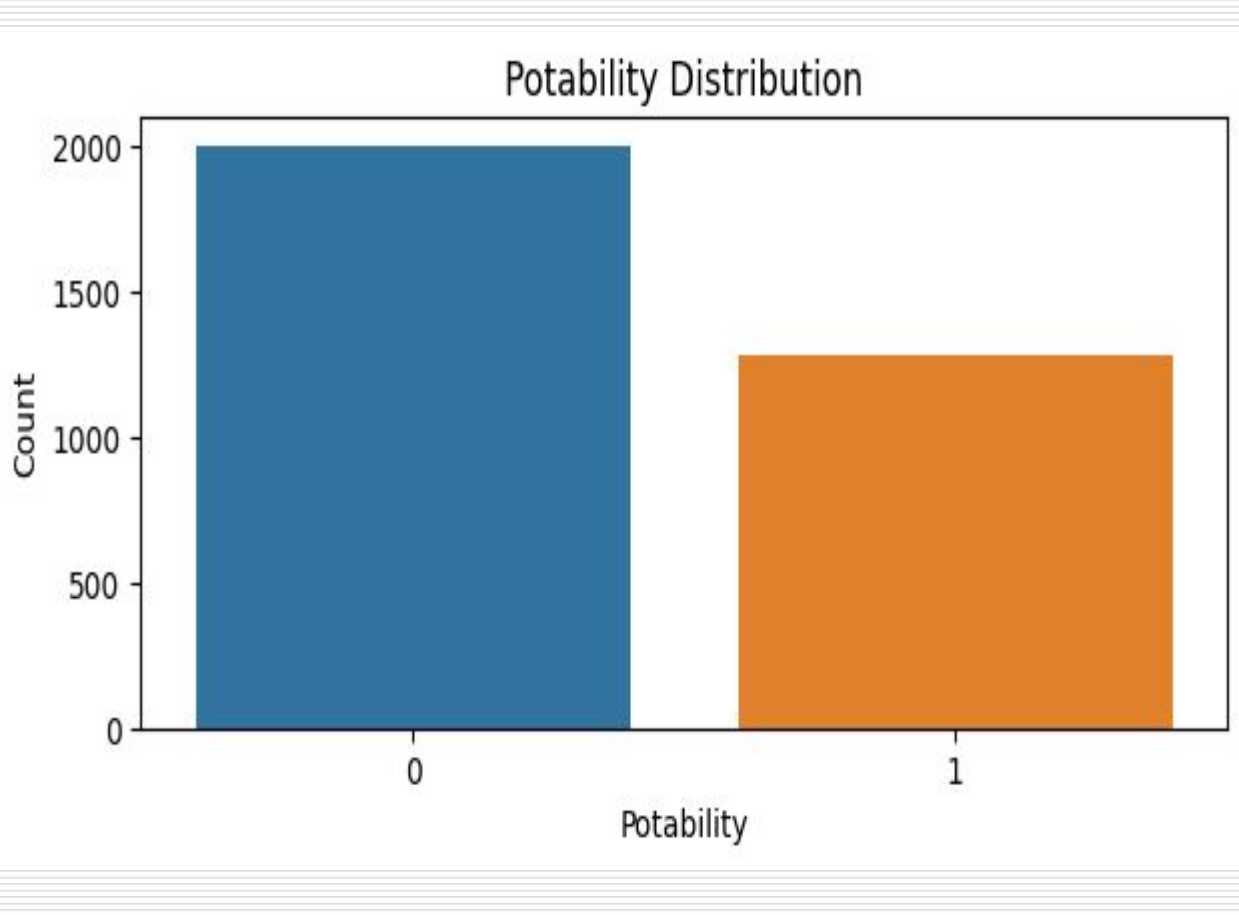
Functional Description for each modules with DFD and Activity Diagram





Implementation & Results of Module





Conclusion & Future Work

The Water Quality Prediction system effectively uses machine learning to classify water as potable or not based on chemical properties. It offers accurate results through a user-friendly Flask web interface. The current model ensures fast and reliable analysis using real-world data. In the future, image detection of pH strips can be added to enhance usability. This will enable users to input pH values via camera, making the system more accessible and automated.

References

A Machine Learning Approach for Water Quality Prediction

Authors: Narmada, M., & Venkatesh, S.

Predicting Drinking Water Quality Using Machine Learning Algorithms

Authors: Bui, D.T., et al.

Water Quality Assessment Using Machine Learning: A Review

Authors: Zainab Khan, Aftab Ahmad

Paper Publication Status

- The research paper has been fully written and is currently under mentor review for final approval.

Thank You