**Water Quality Analysis**

**Introduction:**

This report presents an analysis of a water quality dataset and the development of machine learning models for predicting water potability. The dataset contains information on various water quality parameters, and the goal is to build models that accurately classify water samples as potable or non-potable.

**Steps Taken:**

1. **Data Exploration:**
   * The initial exploration of the dataset revealed missing values, which were addressed by dropping rows with missing values for simplicity.
2. **Univariate Analysis:**
   * Univariate analysis was performed on the 'pH' column, including the creation of a frequency distribution table and visualizations such as a bar chart, pie chart, and frequency polygon.
3. **Model Development:**
   * Three classification models were selected for comparison: Logistic Regression, Decision Tree, and Random Forest.
   * The models were trained on the dataset, and their accuracy was evaluated on a test set.
4. **Model Comparison:**
   * The accuracy of each model on the test set was compared.
   * Additional models, including XGBoost and Support Vector Machine (SVM), were suggested to potentially improve accuracy.
5. **Normalization:**
   * Feature normalization using **StandardScaler** was introduced to see if it improves model performance.
6. **Confusion Matrix and Metrics:**
   * Confusion matrices and performance metrics (accuracy, precision, recall, F1-score) were computed for each model to provide a more detailed evaluation.

**Normalization:**

Feature normalization was applied using **StandardScaler** before training models to ensure consistent scales.

**Confusion Matrix and Metrics:**

Confusion matrices and performance metrics were calculated for each model to evaluate their overall performance.

**Conclusion:**

This analysis provides insights into water quality prediction using machine learning models. By continually iterating and refining models, we aim to achieve higher accuracy and better understand the factors influencing water potability.