**WATER QUALITY ANALYSIS – PHASE 2**

**ALGORITHM:**

**LINEAR REGRESION:**

* Linear regression is a statistical method that can be used to model the relationship between two variables. It can be used to predict the water quality index based on parameters like temperature, dissolved oxygen, pH value, Turbidity, Nitrates and fecal coliform. It relates one or more explanatory variables to a response variable using linear coefficent.
* Urbanization and industrialization have contributed to a deterioration of water quality at an unprecedented pace, contributing to harrowing diseases. The dataset that has been used for WQI assessment consists of a total of 1992 samples with six columns corresponding to six parameters mentioned earlier. Based on the simulation results it has been observed that the proposed regression model is able to predict the WQI effectively.
* It’s important to note that linear regression is just one of many statistical methods that can be used for water quality analysis. Other methods such as logistic regression, decision trees, and neural networks may also be appropriate depending on the specific problem at hand

**STEPS:**

**1.Data Collection:**

* Collect data on the water quality parameters of interest. This data should include both normal and unusual values for the parameters.
* It includes pH value, Hardness, Solids, Chloramines, Sulfate, Conductivity, Organic Carbon, Trihalomethanes.

**2.Data Preparation:**

* Prepare the data for analysis by cleaning and organizing it. This may involve removing outliers, filling in missing values, and transforming the data if necessary.

**3.Model Selection:**

* Choose an appropriate linear regression model to fit the data. This may involve selecting a simple linear regression model or a more complex model that includes multiple variables.

**4.Model fitting:**

* Fit the selected model to the data using a statistical software package such as R or Python.

**5.Model Evaluation:**

* Evaluate the performance of the model by examining its goodness of fit and statistical significance.

**6.Prediction:**

* Once the model has been fitted and evaluated, it can be used to predict unusual values in water quality parameters.

**7.Deployment:**

* If the model is satisfied and meets your requirements you can deploy it in a production environment in real time.