Ideation Phase

Brainstorm & Idea Prioritization

Date	19 September 2022
Team ID	PNT2022TMID54402
Project Name	Efficient water quality analysis and prediction using machine learning
Maximum Marks	4 Marks

1. IoT-based real time-online water quality monitoring system

The tested algorithms would predict the water quality immediately based on the real-time data fed from the IoT system. The proposed IoT system would employ the parameter sensors of pH, turbidity, temperature and TDS for parameter readings and communicate those readings using an Arduino microcontroller and ZigBee transceiver. It would identify poor quality water before it is released for consumption and alert concerned authorities. It will hopefully result in curtailment of people consuming poor quality water and consequently de-escalate harrowing diseases like typhoid and diarrhea.

2. Water quality index modelling using random forest

The water quality index is one of the prominent general indicators to assess and classify surface water quality. Random forest is a combination of predictive trees, which works with a collection of unrelated trees and averages them, in which each tree depends on the values of a random vector in the sample independently and with the same distribution of all trees in the forest. Using a random selection of features to split each node produces error rates that compare favourably to the AdaBoost algorithm, but are more robust with respect to noise. Internal estimates monitor error, strength, and correlation. Further-more, these are used to show the response to the increase in the number of characteristics used in the division. The ideas are also applicable to regression. Reducing the number of possible predictors when constructing a random forest gives a greater opportunity to variables that do not have such a strong predictive power, and, therefore, there will be a greater variability between the trees generated. This procedure is particularly useful when you have a high number of correlated predictors.

3. Supervised machine learning algorithms to estimate the water quality index (WQI) and also water quality class (WQC)

Water quality index (WQI) is the singular measure that indicates the quality of water and it is calculated using various parameters that are truly reflective of the water's quality. To conventionally calculate the WQI, nine water quality parameters are used, but if we do not have all of them, we could still estimate the water quality index with at least six defined parameters. water quality class (WQC), which is a distinctive class defined on the basis of the WQI. Once WQI has been estimated, water quality class (WQC) of each sample will be defined using the WQI in classification algorithms. We can use both regression and classification algorithms. We can use the regression algorithms to estimate the WQI and the classification algorithms to classify samples into the previously defined WQC.