Team -3 – Binary Mind

Members = Rahul V Rao, Bidura Sarkar, Pankaj Devjani



Introduction:

Our project aims at detecting the purity of water obtained from a water resource by receiving parameters from an HTML form regarding the amount of arsenic, barium, cadmium, viruses, bacteria, and so on present in the water sample. The model uses the Random Forest Classifier module to predict if the water is safe for drinking. Flask module has been used to receive data from the front-end form.

The oneAPI platform has been used to develop the system. The sklearnex module has been used to import the RandomForestClassifier module.

## Random Forest is one of the most popular and commonly used algorithms by Data Scientists. Random forest is a *Supervised Machine Learning Algorithm* that is *used widely in Classification and Regression problems*. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

## One of the most important features of the Random Forest Algorithm is that it can handle the data set containing *continuous variables,* as in the case of regression, and *categorical variables,* as in the case of classification. It performs better for classification and regression tasks. In this tutorial, we will understand the working of random forest and implement random forest on a classification task.

## 