**1. Approach Used:**

I have used Xgboost Classifier for the problem. It is an ensemble learning technique, which creates multiple tress from random subset of data. It is also known as Boosting. First a model is created by choosing the random subset of data. Errors for this model are calculated by comparing the actual values and the predicted values. The next model is then created by taking most of the misclassified data, hence increasing the probability of correctly classifying them. Every model contributes to the input to the next model.

I have implemented this algorithm in Python language. Python has very rich libraries for machine learning algorithm and its very simple to implement these algorithm.

**2. Model performance:**

I have used a library in python called “sklearn.preprocessing” for evaluating the performance of the model. For this problem measuring the “accuracy\_score” would be ideal. It simply compared the the correctly classified data to the whole data set.

By doing so, I got 98.8 accuracy for the model.