Approach for building the Machine Learning Model

The machine learning challenge was to predict the price range according to given mobile phone specifications. The training data had 2000 data points and 21 columns. The test data had 1000 data points across which we had to predict the price range.

A preliminary data analysis concluded that the training data had :
1. 500 data points for Price Range 3
2. 500 data points for Price Range 2
3. 500 data points for Price Range 1
4. 500 data points for Price Range 0

I removed irrelevant columns like Id which had no role in calculating the output class.
I also dealt with Null values by replacing them with zeros in the dataset in order to make the machine learning model train better.

I first implemented the Naive Bayes classifier which got an accuracy of 80%, then Random Forest Classifiers which gave accuracy of 81.5%, and also AdaBoost which had an accuracy of 64.3%.

Given that is was a classification problem I applied the boosting algorithm called Gradient Boosting Classifier which helped me in attaining an accuracy of 92.5%
I implemented the algorithm with the help of the python libraries like numpy, pandas, scikit-learn and xgboost.