

For the prediction of the diamond price, I have chosen the random forest model. More specifically, I have identified the price variable in the training set to be the independent variable, and chosen the rest of the variables to be the dependent variable. I used the ordinal encoder to transform three ordinal categorical variables (cut, color, clarity) into different levels. I then trained the random forest model on the training set with 1000 estimators and a random state of 42. To get the prediction results, I applied the model on the training set which contained the exact same variables but independent entries from the training dependent variables and got a list of predicted diamond prices.