Assignment 5: Random Forest Regression

Steps:

(i) Load the required Libraries

(ii) Import the dataset

(iii) Separate dependent and independent variables

(iv) Split the dataset into testing and traiing dataset

(v) Sort the dataset values according to X for visualization purpose

(vi) Manual Method:

(a) Define number of trees and maximum depth

(b) for each iteration (total trees)

(i) Construct tree using Decision Tree Regressor with atleast n splits (predefined) and max-depth

(ii) Use Bootstrap Aggregating ie select a random sample with replacement of training set

(iii) Train the tree over this random sample

(iv) Append this tree to trees[].

(c) Predict the value of y for test dataset as

(i) Predict trees for X_test from trees[].

(ii) Assign the value of most common label from trees to y.

(vii) Scikit-learn Method:

(a) Create Random Forest Regression() Object for the model, Train then predict.

(ix) Calculate the Mean Absolute Error and Root Mean Square Error for both methods and plot their regression lines separately.

(x) Compare the results for both methods.