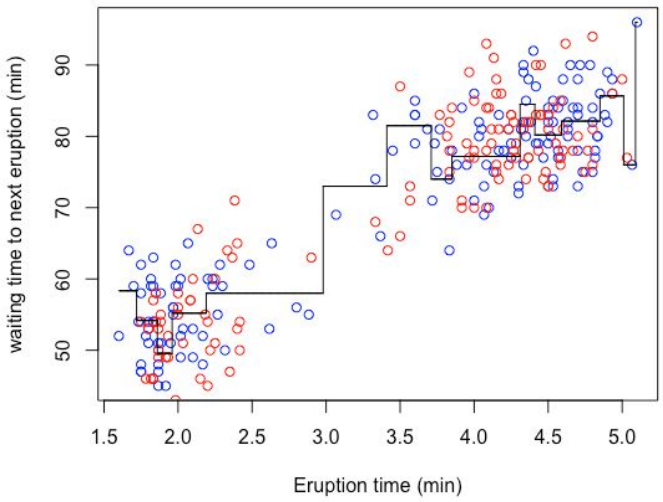
Kağan Cenan-28462

HW05 - REPORT

Q1-Q2: In the 1st and 2nd question I read the csv and and split data into two dataset which are train and test data. The data was a univariate regression dataset. It was about duration of the eruption and waiting time between eruption for a geyser.

Q3: In this question I wrote a function which is a decision tree regression alghoritm. To inhibit overfitting I added preprunning parameter as 'n'. It was for deepness of tree selection. I covered split scores for selection best split.

Q4: I selected P as 25 to see result. And I put the result on plot as expected:



Q5: I calculated RMSE which is covered in the question. My result was:

"RMSE is 6.45408341335209 when P is 25"

Q6: I got the results of RMSE like question 5. This time was for an array which are for P equal 5, 10, 15, ..., 50. I put it on the graph.