

MSAN 621 - Homework 1

Andre Guimaraes Duarte

October 31, 2016

1 The Boston Data Set

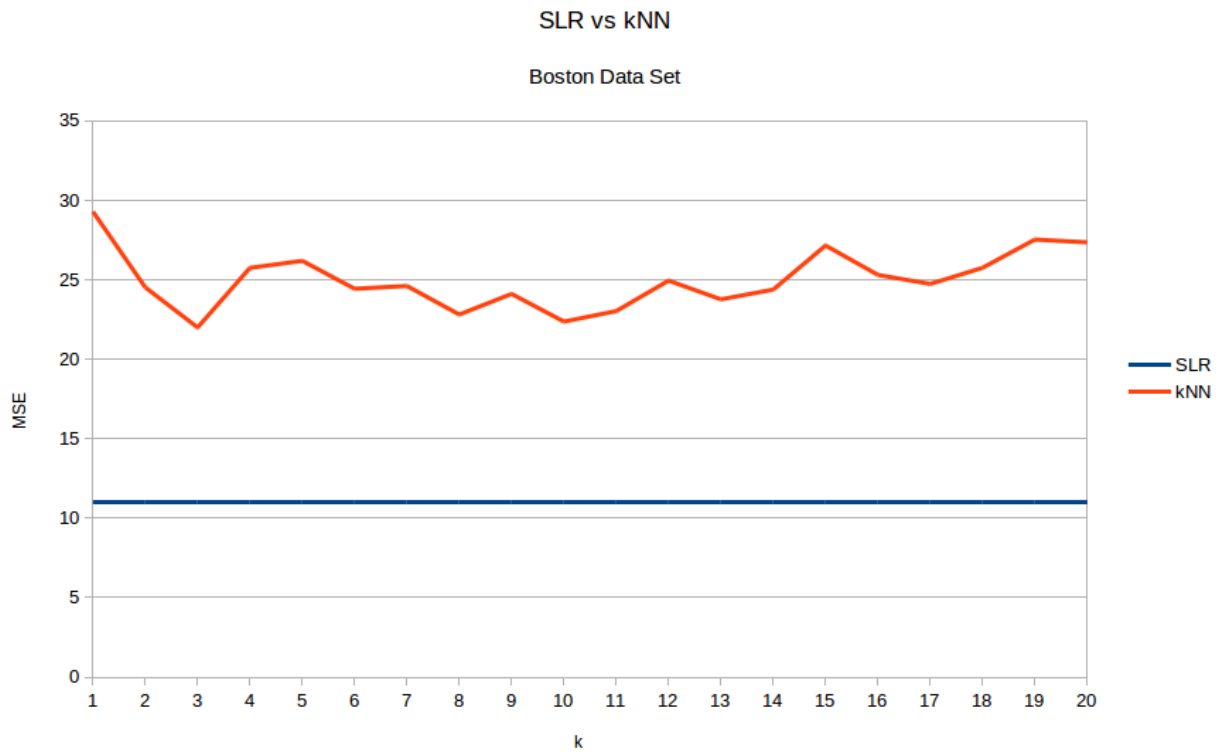
The models are trained with all data except the last 50 rows, which are reserved to test the models' performance.

For SLR, we get an MSE of 10.967.

For kNN, we get an MSE that varies according to k, as shown in the table below:

k	MSE
1	29.299
2	24.532
3	22.003
4	25.754
5	26.200
6	24.441
7	24.615
8	22.820
9	24.117
10	22.371
11	23.036
12	24.953
13	23.774
14	24.388
15	27.162
16	25.311
17	24.743
18	25.761
19	27.539
20	27.362

Plotting MSE against k for both models, we get the image below:



We can see that we get a lower test MSE with SLR than kNN, for any value of k . For this example, simple linear regression is a better algorithm for predicting new data from the Boston Data Set.

2 The U.S. Monthly Climate Normals Data Set