OKAGBUE ONYEKA FRANCIS
CMPT 317
ASSIGNMENT 9
30TH MARCH 2023
11279373

QUESTION 3

Results for data set 1
Learned weights were: [1139.2099999999996, 1899656.9000000001]
Error on TRAINING set: 5394737415232127657
Error on TEST set: 5169175143432937971

Results for data set 2
Learned weights were: [1042.589999999999, 1733241.7799999993]
Error on TRAINING set: 4163751331693655566
Error on TEST set: 4227436391489350294

Results for data set 3
Learned weights were: [1126.5199999999995, 1856591.7099999997]
Error on TRAINING set: 4940981122063962172
Error on TEST set: 4503832301111161933

Results for data set 4
Learned weights were: [947.3199999999999, 1525473.7]
Error on TRAINING set: 2942806730317815867
Error on TEST set: 3058764686557531350

Results for data set 5
Learned weights were: [1002.1800000000001, 1608427.8399999999]
Error on TRAINING set: 3411204065050371309
Error on TEST set: 3136123643197063058

****
Total average test error: 1.0047666082894023e+19

Process finished with exit code 0

The reported test errors for the 5 different data sets are as follows:

Error on TEST set: 5,169,175,143,432,937,971
Error on TEST set: 4,227,436,391,489,350,294
Error on TEST set: 4,503,832,301,111,161,933
Error on TEST set: 3,058,764,686,557,531,350
Error on TEST set: 3,136,123,643,197,063,058

How confident are you in the reported test error now, as compared to your results from the previous question? **For the randomized data partitions, the results between the datasets are much closer to each other than the previous question and I'm more confident in the test error for these results. Both the training and test errors are much closer to each other now than the previous question.**

Do you see any evidence of over-fitting in your results here? And for these, I see there is less over-fitting in the randomized data partitions. The training and test errors are much closer to each other for each of the data set. I would suggest there is not evidence of over-fitting in the result. When there is less evidence of overfitting that would show that the test and training datasets are small and are at a minimal point.