Ryan Gosiaco
912819444
5/24/2018

STA 141C HW 2 Report

1) The run time for the original code when ran on the server was 167.881621 seconds with an accuracy of 0.794000. After implementing multiprocessing on the outer loop, the run time was reduced to 43.909356 seconds with the same accuracy of 0.794000. I chose the outer loop because it made more sense to split up the data into equal parts and run the nearest neighbors algorithm on each respective part and then finally averaging the final values. I also found that the reduction in time seemed to be linearly related to the number of cores I decided to use.
2) Because I did not fully complete HW1, I did not have a finalized/correct implementation of logistic regression. With the code I had i got a weight of -1.56096031e-05 with a time of 152.401907 seconds without multiprocessing. After trying to implement multiprocessing on the inner loop where the gradient is calculated, I ended up with a really long time. I think the way I implemented the multiprocessing is wrong because the code took much longer to run that I expected. I might have changed the algorithm accidentally when messing with the code.