Kenneth Woodard
Lab4
CS 425
October 28, 2022

Question 1

Displacement, horsepower, and weight are the best features for predicting miles per gallon out of the five features. You can tell easily by just looking at the initial scatterplots for each feature. Cylinders and acceleration are not as consistent at predicting MPG. Cylinders would perform much better if 3 cylinder engines didn't exist. Acceleration doesn't perform terribly, but acceleration positively correlates with miles per gallon only because vehicles with higher acceleration generally weigh less. And if a vehicle weighs less, it generally has a higher mpg. Acceleration is more helpful as a part of the five feature model because weight is factored in with it.

Using all of the features improves the performance over all of the features individually except for weight. The R2 score for testing data of weight alone was better than the R2 score for testing data on all of the features. It makes sense too because weight should directly impact miles per gallon the most.

Question 2

Out of all of the number of iterations, 1 million iterations was the best. Obviously, more iterations should lead to better performance so no surprise there. As for why C = 100 performed the best, I can't be sure when given the fact that C = 0.01 performed the second best. I suppose that specific amount of forgiveness (for lack of a better word) on classification happened to work better.