

Name: Suraj Rimal

Answers:

- a. R file attached
- b. Answer: We can explore the relationship between mpg01 and other predictors in the following ways by looking at the scatter plot (Figure 1) and boxplots (Figure 2-5):
 - i. Cylinders: Since the graph does not display any cylinder boundary values that can predict 0 or 1 for cylinder, there is no meaningful relationship. They are distributed evenly on both sides.
 - ii. Displacement: Though the graph appears to be superior to the cylinders, value 1s and 0s are totally overlapping. This cannot possibly be a reliable indicator.
 - iii. Horsepower: Since we can draw a vertical sigmoid line to separate points, this can be considered a strong indicator.
 - iv. Weight: Same as horsepower case.
 - v. Acceleration: Same case as weight.
 - vi. Year: Since the points are evenly distributed on both sides, the predictor is ineffective.
 - vii. Origin: Somehow sparse but same case as Year. Cannot be a good predictor.
- c. Answer is in the R file attached. The test and training data are separated randomly.
- d. The test error for the model obtained in Linear Discriminant Analysis is 10.20%.
- e. The test error for the model obtained in Quadratic Discriminant Analysis is 12.25%.
- f. The test error for the model obtained in Logistic Regression Analysis is 14.29%.
- g. The error values for various values of K is given below:

K=1, error = 14.29%
K = 2, error = 14.28%
K = 3, error = 13.26%
K = 5, error = 15.30%
K = 10, error = 10.20%
K = 12, error = 9.18%
K = 15, error = 10.20%
K = 20, error = 11.22%
K = 25, error = 12.24%
K = 50, error = 12.24%

From the observation, we can say that the error value is least for K = 12, which is 9.18%