Lab 7

Exercise 1 - Linear Discriminant Analysis (LDA)

a) Load the data `Auto.rda` and install the package `ISLR`.

b) Create a fuel consumption rating variable named `Economy` that will be treated as categorical based on the following info. For `mpg<17` mark as `Heavy`. For `mpg>=17 & mpg<22.75` mark as `OK`. For `mpg>=22.75 & mpg<29` mark as `Eco`. For `mpg>=29` mark as `Excellent`.

c) Perform LDA using the `lda` function and using all the available data. Interpret the output based on the theory discussed in class. Use 'library(MASS)'.

d) Perform cross-validation using the `CV=TRUE` option. Construct the confusion matrix as well as the proportion of correctly classified counts. Option `CV=TRUE` is used for "leave one out" cross-validation; for each sampling unit, it gives its class assignment without the current observation. This is a method of estimating the testing classifications rate instead of the training rate.

e) Specify our own prior distribution; `c(0.25,0.25,0,25,0.25)` lists prior probabilities in the same order the classes are listed. Construct the confusion matrix as well as the proportion of correctly classified counts. What do you observe on the results?

f) For this part use the priors `c(0.4,0.3,0.2,0.1)`. What do you observe on the results?