Michael Miller – msmille3

1/30/2016

CS 498DF, HW 1

Four scripts were created to showcase different classifiers:

a) A simple naïve Bayes classifier

b) The same as above but values of 0 were changed to NA in columns 3, 4, 6, and 8

c) A naïve Bayes built with packages caret and klaR

d) An SVM created with klaR and SVMLight

Scripts were created almost entirely from example code provided on the course website. Some corrections and modifications were made, particularly to allow for multiple trials and plotting. All classifiers were run on 10 different partitions of the same data. Their averages are displayed in Table 1. The individual scores of trials 1-10 are displayed in Figures 1 – 4 on the next page.

| Classifier Type	Accuracy
a) 0s ignored	74.4%
b) 0s replaced with NA	72.2%
c) klaR and caret	76.0%
d) klaR and SVMlight	75.6%

Table 1. Accuracy results of the 4 created classifiers

As seen in Table 1, the classifier with the best accuracy was the one created with klaR and caret. This increased accuracy is most likely due to these libraries choosing the best probability fit instead of being constrained to a normal distribution.
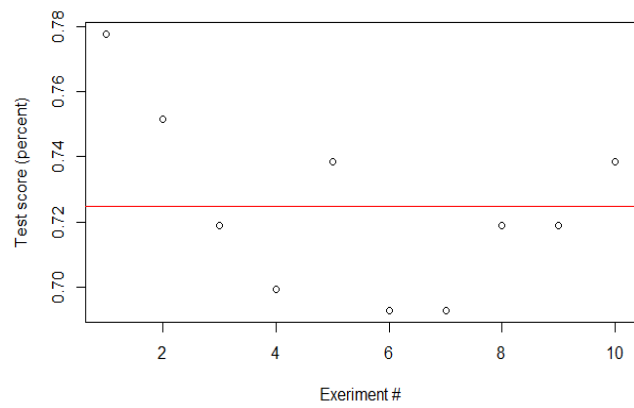
Results for 10 trials of NBC with 0s ignored
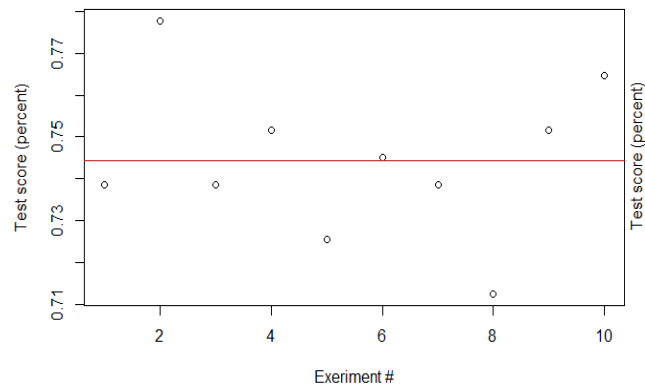
Results for 10 trials of NBC built with klaR and caret libraries
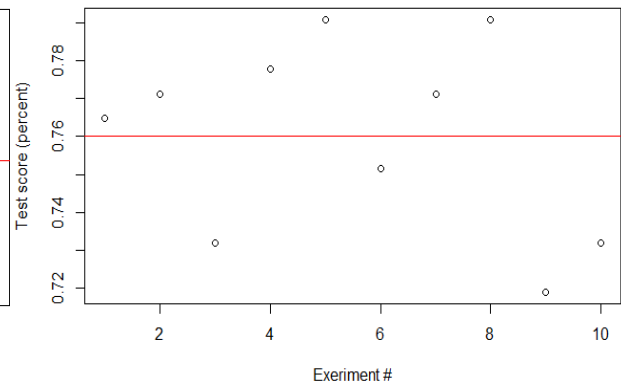
Results for 10 trials of NBC with 0s set to NA

Results for 10 trials of NBC built with SVMlight