Machine Learning

Homework No. 05

Go to https://rpubs.com/Hgoswami/368878 and by following the instructions there download german_credit data with the corresponding column names **(score = 10).**

Assume that the output is german_credit$response with 2 classes “1” and “2”. Assume that the positive class is response = "2".

Divide dataset into training (80%) and test sets (20%) **(score = 10).**

1. Apply LDA on the training set. Draw the ROC curve and calculate the AUC **(score = 10).**

2. Apply QDA on the training set. Draw the ROC curve and calculate the AUC **(score = 10).**

3. Apply Naïve Bayes on the training set. Draw the ROC curve and calculate the AUC **(score = 10).**

4. Apply Logistic Regression on the training set. Draw the ROC curve and calculate the AUC **(score = 10).**

5. Apply k-NN on the training set and by 10-fold cross validation find the optimal value of the parameter k. For the optimal model draw the ROC curve and calculate the AUC **(score = 15).**

6. Compare AUC measures of different models. Find the best model **(score = 10).**

7. For the best model calculate the test Accuracy, Balanced Accuracy, Sensitivity and Precision of the positive class **(score = 15).**