

Master BeNeFri in Computer Science

Course: Statistical Learning Methods with R
Spring 2022

Exercise #8. Classification: LDA and Logistic Regression

Download from the ILIAS website the `Vertebral` dataset (filename: `Vertebral.txt`). This dataset is composed by various biomedical variables that can be used to predict the orthopedic class of the patient (variable `Status`) which is a binary variable with values `Normal` or `Abnormal`. You have a total of 310 observations (patients) with six predictors and no missing values. Further information is provided in the file `Vertebral.pdf`.

1. Use logistic regression to predict variable `Status`.
2. Use LDA to predict variable `Status`.
3. Compare the predictions you obtained with the logistic regression and LDA. Use a fair methodology to compare the classifiers (and explain your choice). Can you estimate the error rate for those strategies? Which classifier is the best? Why?