

Statistisches Data Mining (StDM)

Woche 10

Aufgabe 1 Lab

Read and do the excersises of section 9.6.1 and 9.6.2 in ISLR.

Aufgabe 2 SVM for orange juice

This excercise is taken from ISLR:

- a) Use the OJ data and create a training set containing a random sample of 800 observations, and a test set containing the remaining observations.
- b) Fit a support vector classifier to the training data using `cost=0.01`, with Purchase as the response and the other variables as predictors. Use the `summary()` function to produce summary statistics, and describe the results obtained.
- c) What are the training and test error rates?
- d) Use the `tune()` function to select an optimal cost. Consider val- ues in the range 0.01 to 10.
- e) Compute the training and test error rates using this new value for `cost`.