

Exercise Lesson04: CV

Machine Learning using R

Exercise 1: Cross-validation on the Taxi data using own function

Take a look at the cross-validation function shown in the lecture and try to understand all steps of the function. Use it to perform a 10-fold cross-validation on the taxi data from the last exercise (again predicting the color of the taxis). Estimate the performance of the knn classifier for different values of k (e.g. k ranging from 1 to 10).

Exercise 2: Cross-validation on the Taxi data using caret package

Perform a 10-fold cross-validation on the taxi data for different values of k using the caret package. Do the results agree with the results from our own function?