

Machine Learning

Exercise Sheet 7

Prof. Dr. Dr. Lars Schmidt-Thieme, Osman Akcatepe
Information Systems and Machine Learning Lab (ISMLL)
University of Hildesheim

10. January 2012
Deadline 16. January, 14.15

Nearest-Neighbor-Method (5 Points)

Given 12 cities with the following coordinates:

city i	x_i	y_i
1	11	5
2	6	4
3	4	10
4	4	2
5	2	4
6	7	7
7	8	8
8	9	2
9	5	7
10	7	1
11	1	6
12	11	11

The distance between the city a with the coordinate (a_1, a_2) and the city b with the coordinate (b_1, b_2) is defined through the following formula:

$$d(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$$

- Construct a coordinate system, in which you plot the 12 cities.
- Determine the distance matrix $\mathbf{D} = (d(a, b))$ for the given 12 cities. Send the solution of this part of this exercise an e-mail.
- Solve the TSP (*travelling salesman problem*, finding of the optimal route) with the 1-nearest neighbor heuristic, so that you start with the city 1. Plot the stretch of the ways in the coordinate system. How far is the covered distance? Is the found way optimal, that is, does it have the smallest possible distance? Choose one arbitrary city as start point and determine a solution with the 1-nearest neighbor heuristic from there. Plot the stretch of the ways in the coordinate system..

Programming in R (3 Points)

Read the capitals 9 and 10 from „An Introduction to R“.

- a) Which loop constructs are there in R? Why is `for()` used in R less than in Java or C?
- b) How are user defined binary operators identified in R? When is the naming of function arguments not necessary in function call?
- c) Write two R functions, that calculates the factorial of a natural number. One of the functions should act iteratively and the other one should be defined recursively. What happens, if the name of a function is input in the R console.

kNN in R (2 Points)

Start R and load the library `class`. Apply kNN with different parameters k on the Iris data set and compare the result with the LDA-/QDA results from the fifth exercise sheet. Use the function `knn.cv()` for evaluation and read the function documentation as you convert the data set adequately.