Assignment 1

Mustererkennung/Machine Learning WiSe 20/21

Aufgabe 1. Implementation of a k-NN-Classifier

Implement a k-NN-Classifier in Python (incl. Numpy, Pandas, Matplotlib) on the Jupyter Notebook Environment. Use the "ZIP-Code"-Dataset[1] with the training data as reference for neighborhood. Evaluate the model on the test data.

(a) Print out the accuracy.

(b) Using Matplotlib, plot some of the numbers that are classified incorrectly.

(c) Which k is optimal for the accuracy?

(d) What are advantages and disadvantages for the k-NN-Classifier?

Submit as a fully executed Jupyter Notebook as PDF and as .ipynb in KVV. ***Deadline:*** *16.11.20, 10.00h.*

[1] `https://web.stanford.edu/~hastie/ElemStatLearn/data.html`