University of Stuttgart | ANALYTIC COMPUTING

Machine Learning (SS 23)

Assignment 03: Classification

Mojtaba Nayyeri
Mojtaba.Nayyeri@ipvs.uni-stuttgart.de

Akram Sadat Hosseini
Akram.Hosseini@ipvs.uni-stuttgart.de

Nadeen Fathallah
Nadeen.Fathallah@ipvs.uni-stuttgart.de

Rodrigo Lopez Portillo Alcocer
rodrigo.lopez-portillo-alcocer@ipvs.uni-stuttgart.de

Tim Schneider
timphillip.schneider@ipvs.uni-stuttgart.de

Osama Mohammed
osama.mohammed@ipvs.uni-stuttgart.de

Daniel Frank
daniel.frank@ipvs.uni-stuttgart.de

This assignment sheet consists of 5 pages with 4 Questions and 1 Task:

Submit your solution in ILIAS as a single PDF file.[1] Make sure to list full names of all participants, matriculation number, study program and B.Sc. or M.Sc on the first page. Optionally, you can *additionally* upload source files (e.g. PPTX files). If you have any questions, feel free to ask them in the exercise forum in ILIAS.

Submission is open until Monday, 8th May 2023, 12:00 noon.

[1] Your drawing software probably allows to export as PDF. An alternative option is to use a PDF printer. If you create multiple PDF files, use a merging tool (like pdfarranger) to combine the PDFs into a single file.

Question 1: kNN for Data Classification

Consider a classification problem with one input variable (attribute) x, and the following training data points:

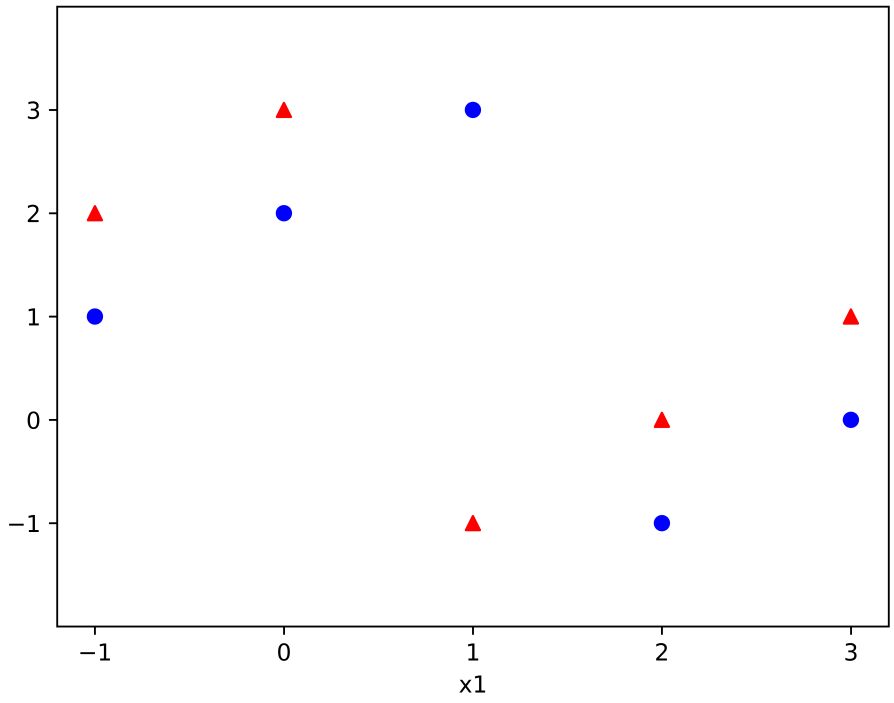
attribute	x	-3	-1	1	7.2
target	y	0	1	1	0

- Find the number of miss-classified points in the training data for each of the following classifiers (use the Euclidean distance function): (1) K-nearest Neighbor with $K = 1$ and (2) K-nearest Neighbor with $K = 3$. State the reasons in each case.

Question 2: kNN and cross-validation

For the data given below with 10 points and 2 classes, answer the following questions:

1. What is the leave-one-out cross-validation error when using K-nearest neighbor classifier with $K = 1$?
2. Which of the following values of K leads to the minimum number of leave-one-out validation errors: 3, 5 or 9?



Question 3: kNN for Image Classification

Research and discuss how you could use a k-nearest neighbor classifier for image classification. You should at least answer these questions:

- How do you represent the images?
- What distance function do you use?
- What decision rule do you use?

Provide an example for your representation of the images and how your classification decision is made based on the distance function and decision rule. Explain the advantages and disadvantages of your approach.

Task 1: kNN and Classification

Please download the Jupyter notebook *assignment2.ipynb*. Follow the instructions in the Jupyter notebook.