1 Patten Recognition Project

1.1 General

important dates:

- Date for hand in of the project: 20.01.23 12:00
- Date for the defense of the project: 23.01-27.01.23

hand in:

- you should hand in all necessary files on moodle
- your hand in should include:
    1. all needed python files together with a proper documentation
    2. Written part (approx. 5-10 pages)

1.2 First Part

On moodle you will find a data set called simply dataset. It contains 512 10-dimensional data points. With that data set you should solve two problems:

1. Determine a reasonable number of clusters inside the data. Decide for your own which measure you will choose to evaluate this reasonable number. Justify your choices in the written part.
2. Fit a GMM model with the data and choose the number of components regarding your outcome of the first part.

1.3 Second Part

On moodle you will find a second data set called classification data set. Further you find 3 papers regarding different LVQ modells. Individual hints are given in the seminar.
The assignment of the groups to the individual algorithms will be done in the seminar. You should implement the algorithm assigned to you and train it on the given training data. Try to find a optimal number of prototypes using cross validation and and describe your procedure in the written part. Furthermore, there is another set of data which is unlabeled. You have to classify it with your trained result and should hand in you're predicted labels together with the pythons files.