

Home assignment 1

February 26, 2015

Deadline 9.03.2015 23:59 PDF file of the report should be sent to sven.nommm@gmail.com
Scripts, codes, functions etc. should be presented during the practice on
12.03.2015.- practice will take place in ICT-403

Please present your **own** original work. Strictly no plagiarism!!!
Results should be supported by necessary computations, scripts, functions
and codes.

Preferred implementation language is MATLAB. You may use MATLAB
functions from the Statistics ToolBox or program your own functions. You
are expected to be able to explain your implementation.

Report (preferably in L^AT_EX) should be structured as follows:

1. Student name and code;
2. Problem statement;
3. Short description of the used method(s);
4. Description of the implementation;
5. Description of the experiments and results;
6. Remarques.

1 Decision trees

Develop and draw decision tree on the basis of the dtreedata.mat

2 k-Nearest Neighbours

Implement k-Nearest Neighbours for the data given by the

```
kNN_data in HomeAssignment1_studentX.mat,
```

here X is the number assigned by OIS.

2.1 Distance & Metric

Verify if Euclidean distance / Manhattan (cityblock) distance/ Chebyshev distance / Mahalanobis distance / Cosine distance /Standardized Euclidian distance/ Correlation distance / Spearman distance / Hamming distance /Levenshtein /Jaccard distance is a metric. Each student is assigned two distances to verify. The cases are selected randomly while assignment is given. If you missed the class please contact the lecturer by e-mail.

3 K-means

Determine "optimal" number of clusters and perform clustering of the data given by

```
Kmeans_data in HomeAssignment1_studentX.mat
```