Heinrich Heine Universität Düsseldorf

HEIDERLAB

Exercise 3 on Machine Learning WS 2023/24
Prof. Dr. Dominik Heider
M. Sc. Arsam (Mohammad) Tajabadi
Submission: Until 15.11.2023, 23:59 on Ilias.

Task 1. Principal Component Analysis (PCA) (5 Points)

Implement a principal component analysis using only the Python package **Numpy** and (optionally) **Pandas**. Specify scores (position), loadings (rotation), and explained variance. Plot the results appropriately (you may use the **Matplotlib** package). Use the dataset *dat.csv*.

Task 2. Multi-dimensional scaling (3 points)

Implement a PCoA (Classical multidimensional scaling, principal coordinates analysis) in two dimensions. Use the packages **Scikit-Learn**, **Numpy**, **Matplotlib** and the dataset *eurodist.csv*. Plot the results.
Bonus: Is the implementation of MDS in Scikit-Learn a classic MDS (PCoA)? (With justification)

Task 3. Comprehension questions (2 points)

a) Are there conditions under which the results of a PCA and a PCoA are equivalent? If so, what are they? (1 P.)

b) What options exist to determine a meaningful number of dimensions for a PCoA? (1 P.)