C | A | U

Kiel University
Christian-Albrechts-Universität zu Kiel

Exercise Sheet 4
Intelligent Systems
November 18, 2019

Representation

Exercise 1 - Representation

A. Explain the idea of the *Shape Definition Language* and its application?

B. Approximate the time series in Figure 1 with the following approximations:
   - *Piecewise Aggregate Approximation* (*PAA*) with 4 segemtns.
   - *Clipping* to binary values ($\rightarrow$ search the procedure on the internet).
   - *Picewise Linear Approximation* with 4 segments.
   - *Run-Length Encoding* (*RLE*).

C. Aggregate the timeseries to the following statistical measures:
   - *Mean*
   - *Standard deviation*
   - *Mode*

D. What are the advantages and disatvantages of the *clipping* procedure?

E. What is the main difference between the *Adaptive Picewise Aggregate Approximation* (*APAA*) and the *PAA*?

Exercise 2 - Data Adaptive Representations

A. What is the goal of the *Principal Component Analysis* (*PCA*) and what is its basic assumption.

B. What is the benefit of the *PCA*?

C. Describe the following items:
   - Zero-mean feature
   - Variance
   - Standard deviation
   - Covariance matrix
   - Arithmetic mean
   - Eigenvector
   - Eigenvalue
   - Projection onto new feature space

D. How can we get a dimensionality reduction with the means of Eigenvalues?
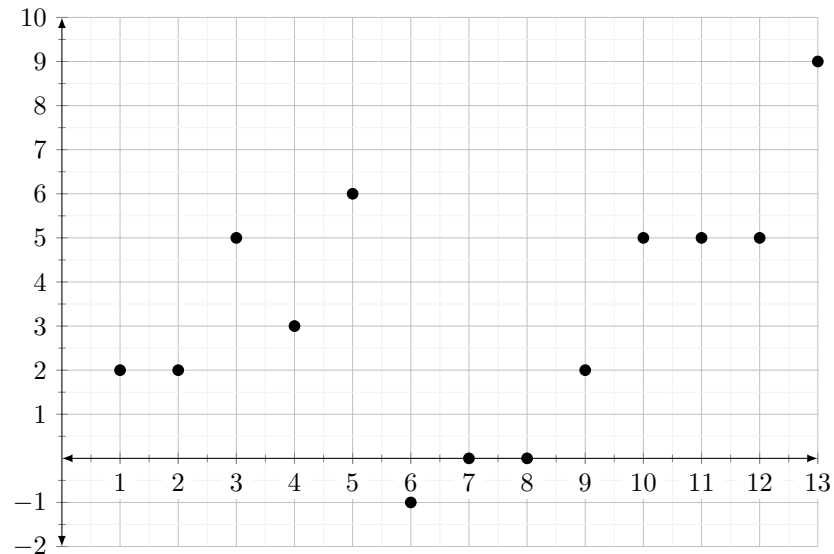
Figure 1: Point sequence.

Exercise 3 - Data Adaptive Representations

A. Download the file *04_Representation.ipyn* from *OpenOlat*.

B. In order to solve the tasks, you can use the library *numpy*.

C. Compare your results afterwards with the help of *sklearn*.

Signature Task

Officially the *Signature Task* has been started! You can already find data set called *INT_Signature.zip* within the project folder *00_SignatureTask* on OpenOlat. Since you already have learned about *Preprocessing* and *Representation*, you are now able to proceed these tasks for the *Signature Task*. In detail, the following should be done first:

Preprocessing

A. Download the data *INT_Signature.zip* and get familiar with the file structure.

B. Preprocess the raw data and create numerical dataframes for every user by calling a function *get_signature_df(user_id)*.

Representation

A. Create a function that shows an image of the signature of a specific user by *calling show_signature(user_id)*.

B. For every user upload a example signature to *OpenOlat*.

C. Think about sensful features

D. Project the samples in your feature space onto a two-dimensional space by means of PCA.