Prof. Dr. Jan Rolfes, Ehsan Waiezi, Lars Weidner
Winter term 23/24

Mathematics of Learning – Worksheet 7 - Repetition

- The exercise sheets will be uploaded every Monday. Solution sketches will be uploaded one week later.
- You can hand in your own solutions via StudOn and we correct them - this is not mandatory. Please hand in in small groups of 2-3 students.
-
- For questions, please use the forum on StudOn since other students may have similar questions. If you have a more personal question about the exercises please send an email to ehsan.waiezi@fau.de or lars.weidner@fau.de respectively.

This sheet is easier than the others, it serves as a little motivation to scan over the previous lectures and remember what the key points in these have been. Nevertheless, to prepare for the exam, it is also wise to solve previous sheets if not yet done so. This is ***not*** *a mock exam.*

Exercise 1 [K-means].
Consider the data set $X := \{x_1, ..., x_6\} := \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}, \begin{pmatrix} -2.5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \end{pmatrix} \right\}$, clustered in Clusters $C_1 := \{1, 2, 3\}$ and $C_2 := \{4, 5, 6\}$, with cluster means $m_1 = \begin{pmatrix} 0 \\ 0.5 \end{pmatrix}$ and $m_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Calculate three iterations of the k-means algorithm, starting with the given data.

Exercise 2 [Linear PCA].
Given a set of data points $X := \{x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}\} := \left\{ \begin{pmatrix} 7 \\ -3.5 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 4 \\ -3 \\ 2.5 \end{pmatrix}, \begin{pmatrix} 9 \\ -5 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ -2 \\ 1 \end{pmatrix} \right\}$.

Calculate the first two (linear) principal components of the data set.

Exercise 3 [Kernel methods].
Let data points $X := \{x^{(1)}, x^{(2)}, x^{(3)}\} := \left\{ \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 2 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^4$ be given.
a) Specify the polynomial kernel depending on parameters $a \in \mathbb{R}$ and $d \in \mathbb{N}$ and the gaussian kernel depending on a parameter $\sigma > 0$.
b) Calculate the kernel matrix of X for a polynomial kernel with $a = 1$, $d = 2$ and a gaussian kernel with $\sigma = 0.5$.
c) Apply the kernel PCA to the data points and calculate the 2-dimensional kernel principal components of the data set X for a polynomial kernel with $a = 1$ and $d = 2$.