

Machine Learning

Spring Semester 2023
Prof. Dr. Peter Zaspel

Assignment Sheet 5.

Submit on **Tuesday, March 14, 2023, 10:00.**

Exercise 1. (Mathematics recap)

In this problem, you will need to remind yourself about the definition of a positive definite matrix. Moreover you will look for local maxima and minima of a function of two variables.

- a) Compute for the following matrices, whether they are positive definite.

$$M = \begin{bmatrix} 3 & 5 \\ 5 & 7 \end{bmatrix}, \quad N = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 2 & -3 \\ -2 & -3 & 5 \end{bmatrix}$$

- b) You are given the following function:

$$f(x, y) = 2x^3 + 6xy^2 - 3y^3 - 6x$$

Use the Knowledge 4.1 statement from the lecture notes to find and classify all locations of minima and maxima of this function.

(4 Points)

Exercise 2. (Linear regression by least squares)

You are given the training data $\mathcal{T} = \{((2, 1)^\top, 3), ((2, 3)^\top, 1), ((3, 3)^\top, 6), ((4, 3)^\top, 5)\}$.

- a) Use linear regression by least squares to compute a predictor for the output. Do this by hand, i.e. don't use the computer and show all computation steps.
- b) Predict the output for $\mathbf{x} = (2.5, 5)^\top$.

(4 Points)

Exercise 3. (Bayes classifier)

Prove Theorem 3.2 from the lecture. If you consider this as too hard, consult the literature from this field to solve the problem. In case you use external literature, make sure that you give a citation for that literature (i.e. say where you took it from) and rephrase the proof appropriately.

(4 Points)

Programming Exercise 1. Consider the Examples 4.2 and 4.3 from the lecture, for which you also have access to the source code. Complete the following tasks:

- a) (Re-)implement Example 4.3. This time, however, you need to implement the linear regression yourself, without a machine learning library. (If you implement in Python, just start from the available Jupyter notebook.) Verify the correctness of your implementation by cross-checking it with Example 4.3.
- b) Apply your implementation to the Energy efficiency Data Set from the UCI Machine Learning Repository. Build the predictor for the required heating load on the full data set and predict the load on the first three samples of the data set.
- c) Implement Example 4.5 starting from Example 4.2 from the lecture notes.

Reference solutions will only be provided in Python+Matplotlib. The submission format for Python is a Jupyter notebook. The submission format for C/C++ is standard source files. Choose an appropriate format for the Gnuplot-related submission.

(4 Points)