

# Machine Learning

## Exercise 1: Introduction to linear regression with scikit-learn and scipy

Prof. Dr. Thomas Kopinski

July 10, 2023

### Abstract

In this week's exercise you will become familiar with using the python module scikit-learn to perform a linear regression on a given dataset. In addition you will use the more flexible `curve_fit` function from the scipy package. The data can be checked out from this courses's Git repository.

### Task 1: Getting familiar with scikit-learn

- Additional information about implementing linear regression in python with the help of scikit-learn can be found [here](#)
- The repository for this course can be found [here](#)

### Task 2: Linear regression with scikit-learn

- Import the necessary modules
- Load the file with the name "simple\_regression.csv" from the data subfolder with numpy or pandas
- Instantiate a LinearRegression model from the scikit-learn library
- Fit a linear model  $f(x) = y = mx + b$  to the given data (reshaping of the data might be necessary)
- Display the values of the coefficients and plot the data and fitted function
- Interpret the results

### Task 3: Linear regression with scipy

- Import the necessary modules
- Load the file with the name "simple\_regression.csv" from the data subfolder with numpy or pandas
- Define a model function that returns a linear mapping of the variable  $x$
- Fit that linear model to the data with `scipy.optimize.curve_fit`
- Display the values of the coefficients and covariance matrix
- What is the meaning of the covariance matrices entries?
- Plot the data and fitted function
- Interpret the results and compare it to the scikit-learn version
- Can this method be used in a more general way?