Advanced Methods in Medial Image Analysis: Exercise 2

Julia Wolleb

Deadline: 11.03.2024

Exercise 2: Linear Regression

The linear regression we have seen in the lecture can be used as a predictive model, assuming a linear relationship between input and output parameters. In this case, the model follows a linear equation $f(x) = wx + b$. The model parameters we aim to optimize ar the slope of the line w, as well as the offset b. In the given dataset, we aim to predict the score achieved by students in a test, as a function of the time they spent studying. The dataset is given in the csv file *datahours.csv*. Follow the code snippet *linearregression.py*.

- Plot the data in a scatterplot (x-axis: hours, y-axis: score)
- Split the data into training (80%) and test (20%) set.
- Compute the closed-form solution with the equations given in the lecture.
- Plot the line in the same plot, and report w and b.
- Report the Mean Squared Error Loss of the whole training and test set.
- Hand in your code, which does not throw exceptions.