Introduction to Machine Learning and Artificial Intelligence, Summer 2022 (ET1550)

Project 1

In project 1, a multivariate linear regression model will be investigated. According to the lectures, in linear regression, we would like to find the relation between several inputs and an output. This relation (model) will be used to predict the output, given any new input.

$$h_w(x) = w_0 + w_1 x_1 + \cdots + w_n x_n$$

where $h_w(x)$ is the prediction, $x_1, \ldots, x_n$ are the input features and $w_0, \ldots, w_n$ are the parameters of the model.

The Dataset

The provided code for this project is designed for the given industrial dataset of an air production unit in a utility plant from the petrochemical industry.

The Code

The project may be implemented in both Google Colab and Jupyter Notebook (please refer to the Project IDE document).

The Report

For the project report, you only need to provide answers to the questions. In addition, include your written codes (from the places that are specified in the code) for each question (if required). The project report should be a PDF file containing your answers, codes and possible outputs from the specified parts.

The deadline for submission is **22 July 2022**.