Technische Hochschule Ingolstadt, M-APE, Prof. A. Schiendorfer Lukas Lodes M.Sc.

# Intro to AI and Neural Networks (Summer 2023)

## Assignment 06

**Exercise 1** ***(Linear Regression – A real dataset)***

In this exercise, we want to apply linear regression to a real dataset. In our case, it's a dataset containing records about the weather in Szeged (Hungary) from 2006 to 2016[1]. We have prepared a Jupyter notebook (`weather_analysis.ipynb`) that guides you through the implementation.

a) Take a look at the CSV file. How is the dataset represented? Which features may be interesting for the application of linear regression?

b) As a first task, let's stick to one feature and try to predict the temperature based on the humidity. Take a look at the visualizations and try to make assumptions about the regression coefficients.

c) Train the model and check how well the actual coefficients match the data.

d) Let's evaluate our model with a seperate test dataset that's not used during training. Use the `train_test_split` function provided by scikit-learn to split the dataset and the `mean_squared_error` function to evaluate the model.

e) Add more features: Take another look at the dataset and select additional features that might be helpful in predicting the the temperature. How well does the model perform?

[1] Source: `https://www.kaggle.com/datasets/budincsevity/szeged-weather`