Kiel University
Christian-Albrechts-Universität zu Kiel
Faculty of Engineering

# Assignment I
# Intelligent Systems
# Preprocessing

**This assignment will be discussed on November 25, 2020**

**Instead of the regular lecture next Monday 23, 2:15 pm we offer a practice lesson:**
**You can work on this assignment while we are available online for questions.**
**Join via the regular Zoom link for lecture.**

## Concept

This is the first of three assignments that build on top of each other, where you work with a more complex set of time series data. The aim is to arrive at a model that can fulfil a forecasting task (FT) on the data.

| | | |
|---|---|---|
| **FT I** | Preprocessing | Mo, 23.11.2020, 14:15 |
| **FT I** | Preprocessing Presentation | We, 25.11.2020, 10:15 |
| **FT II** | Feature Selection | Mo, 14.12.2020, 14:15 |
| **FT II** | Feature Selection Presentation | We, 16.12.2020, 10:15 |
| **FT III** | Model Selection | We, 27.01.2021, 10:15 |
| **FT III** | Model Selection Presentation | We, 27.01.2021, 12:15 |

For each assignment we *do not* give you a fixed prescription of what to do. We rather want you reflect upon the subjects presented so far in the lectures and come up with an *basic* analysis approach for time series data. It is mandatory for each group give a brief presentation and hand in your results (see below).

## Data Acquisition

Download the time series data files (*OLAT* → *Resources* → *Assignment*). Create a Jupyter notebook that, started from the same directory, imports the downloaded data.

## Preprocessing

For this first assignment, take a look at the lecture chapter "Preprocessing". Expand the above Jupyter notebook and address the following topics:

- Missing Values, Scaling, Outliers
- Data encoding and Signal processing
- Visualisation of the data

Write your considerations for each of these topics regarding the dataset – maybe supported by plots – into the notebook and *preferably* create code, that does the actual preprocessing steps.

## Presentation

For the next exercise session (the following Wednesday) create a presentation of your results:

- The presentation should be *3 to 4 pages short and take at the most 5 minutes.*
- Briefly introduce your group.
- Outline the overall approach.
- Present your result: plots, stats, descriptions, ...

**Upload your results before the exercise session.** For this, use the 'participant folder' of your group in OLAT. This upload should include:

- Your presentation, preferably as a PDF file.
- The Jupyter notebook you used for the preprocessing.
- The actual preprocessed data as a result of running the notebook.