Required time for the task 1 approx. 2-3 hours

Data:

Below, you can find the link to the dataset which can be downloaded as a zip-file:

 http://zhen-wang.appspot.com/rice/files/uwave/uWaveGestureLibrary.zip

It is a time series classification task (gesture detection). Some details/intuition about the data are illustrated below (taken from uWave: Accelerometer-based personalized gesture recognition and its applications).
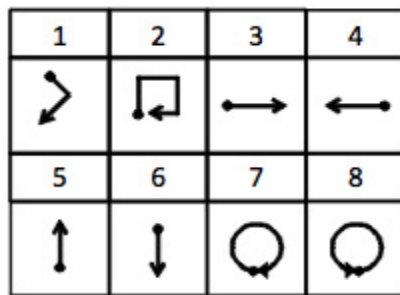

Figure 3: Gesture vocabulary adopted from [KKM+06]. The dot denotes the start and the arrow the end

The dot denotes the start and the arrow the end of the gesture.

If you download and unpack the data, you will get a couple of .rar files.

Here the meaning:

- On the top level, each .rar file includes the gesture samples collected from one user on one day. The .rar files are named as U$userIndex ($dayIndex).rar, where $userIndex is the index of the participant from 1 to 8, and $dayIndex is the index of the day from 1 to 7.

- Inside each .rar file, there are .txt files recording the time series of acceleration of each gesture. The .txt files are named as [somePrefix]$gestureIndex-$repeatIndex.txt, where $gestureIndex is the index of the gesture as in the 8-gesture vocabulary, and $repeatIndex is the index of the repetition of the same gesture pattern from 1 to 10.

- In each .txt file, the first column is the x-axis acceleration, the second y-axis acceleration, and the third z-axis acceleration. The unit of the acceleration data is G, or acceleration of gravity.

Important: Overall, we are interested in the way you build your **ML pipeline**.

1) Given this data, your task is to perform gesture detection. That is,

given a time-series, classify it in one of the 8 classes. More specific:

A. We want you to use python for all aspects of the "data science workflow: data preprocessing/feature extraction/ML-training/ML_test". (a few exceptions below)
B. Please implement and train a logistic regression model by using open source libraries such as sklearn or pytorch
C. If you want to use other ML-models (neural networks, support vector machines), feel free to use existing libraries. (Please, do not invest too much time here)
D. You can (should) use libraries for visualizing the results. (whatever you feel makes sense to visualize)