

Practical II: Processing accelerometer data

In this practical, you will process your data collected from accelerometer.

1. Setup and installation

Navigate to your `wearables` folder to run the following in your terminal.

- First, create a folder to store your data

```
$ mkdir data
```

- Next, we need to get some scripts for the pre-processing of your accelerometer data.

```
$ cd scripts
$ git clone https://github.com/activityMonitoring/biobankAccelerometerAnalysis.git
$ cd biobankAccelerometerAnalysis
$ bash utilities/downloadDataModels.sh
$ pip3 install --user .
$ javac -cp java/JTransforms-3.1-with-dependencies.jar java/*.java
```

If you did not have JDK installed, you would not be able to run the last command. In this case, you should follow these instructions to install it.

- Download `jdk-13.0.2-osx-x64_bin.dmg` (macOS, 173MB) [here](#). Make sure you accept the License Agreement. This will take ~10 minutes.
- Run the installer. Check that you've installed it properly by typing this in your terminal and you should see the java version installed.

```
$ javac --version
```

2. Extracting data

We can now extract the raw data from your devices.

Accelerometer

- Plug in your accelerometer to your computer. Go to the external drive where it is located and copy the `.CWA` drive over to your `wearables/data` folder and rename the file as `myAcc.cwa`.
- Currently the raw data stored in the CWA file does not allow easy extraction of the 3-axis acceleration data, so we need to convert this to a more suitable format such as CSV for further manipulation. Run the following (you're still expected to be in the `biobankAccelerometerAnalysis` folder.)

```
$ python3 accProcess.py ../../data/myAcc.cwa --rawOutput True --activityClassification False --deleteIntermediateFiles
```

3: Preprocessing

In this practical, we will focus on the preprocessing steps performed on your accelerometer data for activity recognition. This is an often time-consuming step yet very important step before your data is useful for further analysis. You are provided with a half-completed jupyter notebook to work with.

After this you will our built pipeline and perform activity recognition on your accelerometer data with machine learning.

Building features from your accelerometer traces

Your raw accelerometer traces is now in `wearables/data/myAcc.csv` . It contains 4 columns: time, x, y and z accelerometer readings.

Go through the jupyter notebook `Features.ipynb` in your `prac2` folder.

Biobank Pipeline

While you have had a taste of raw sensor data preprocessing through the jupyter notebook, many other operations (such as readings imputation, non-wear time detection ..) can be done on your raw data to further perfect your dataset for downstream machine learning tasks.

So instead of using the dataset you just created, `my_data.csv` , we will now go back to your raw data again - the CWA file. We will run some scripts which form the standard analysis pipeline on accelerometer data in the Biobank studies, this involves preprocessing operations as well as running a machine learning model that had been trained on the Biobank data for activity recognition.

Note that these operations also involves mapping the finegrained annotations you had into larger classes.

Run this from your `wearables/scripts/biobankAccelerometerAnalysis/` directory.

```
$ python accProcess.py ../../data/myAcc.cwa --skipFiltering True
```

To visualise the time series and activity classification output:

```
$ python accPlot.py ../../data/myAcc-timeSeries.csv.gz ../../data/plot.png
```