Nadine Imholz
Virgilio Strozzi
Hélène Durand-Smet

Exercise 2: Wearable Activity Monitoring

Step Counting

In our approach, we use the magnitude of the wristband accelerometer data because it is hard to isolate the orthogonal accelerations in the global frame. By applying a butter bandpass filter within the 1-3 Hz range, we effectively isolate frequencies characteristic of walking strides. Additionally, we utilize standard deviation thresholding on a centered moving average to differentiate between walking and other activities, as it is one of the best-performing algorithms for walk detection. Subsequently, employing Windowed Peak Detection (WPD) enables us to find potential steps within the walking segments.

Features for Machine Learning

For the remaining tasks, we use machine learning to predict the location of the smartwatch, the covered path, as well as the activities performed. We perform two types of features selection; the first one utilizes windows of 60 seconds of length, overlapping every 30 seconds, for every sample, where for every sensor signal (e.g. ax, ay, az, altitude, …) we derive relevant statistics such as mean, std, max frequency and the energy of the signal in that window. Moreover, for some of the sensors we also evaluate their correlation. We are left with a dataframe containing relevant features for every window for all samples. After that, for every sample we gather all these features with mean, quantile 25, quantile 75, std, generating a single row per sample containing a combination of all the previous features. The second type of feature selection, simply defines a maximum length a sample can have. Every sample with shorter length has its sensors' measurements padded with zeros. At that point, we utilize a window of 60 seconds of length overlapping every 30 seconds and we evaluate the mean of every sensor per window. We are then left with a single row per sample, with a fixed number of windows containing the mean measurements of every sample in that window.

Smartwatch Location

For this approach, we utilize the first features selection. From these features, we preprocess the data by removing labels that we think are unhelpful or have a low correlation with our target variable. In the end, we only keep the features related to the measurements of the gyroscope and the accelerometer. Then, we fit the data to the watch location label with XGBoost. To predict the smartwatch location, we prepare the test data in the same way as the training data and infer the location with the trained model.

Activity Recognition

We proceed similarly to the approach for smartwatch location by using the first approach of features selection. We remove labels that we think are not relevant for training, hence we only keep the gyroscope, the accelerometer data and the predicted number of steps per window. Then, we fit the data to the path index label with XGBoost and use this trained model for the prediction of the activity label on the test set preprocessed in the same way.

Nadine Imholz
Virgilio Strozzi
Hélène Durand-Smet

Path Classification

In this case we use the second approach of features selection since it produces better results. Again, we remove irrelevant labels and low correlation features (under some mean and std threshold), yielding to the utilization of only the magnetometer, the altitude and predicted steps per window. Lastly, we fit the data to the path label with XGBoost. Then, we use this same model for the test set preprocessed in the same way.