Human Activity Detection Project

* **Data Description:**

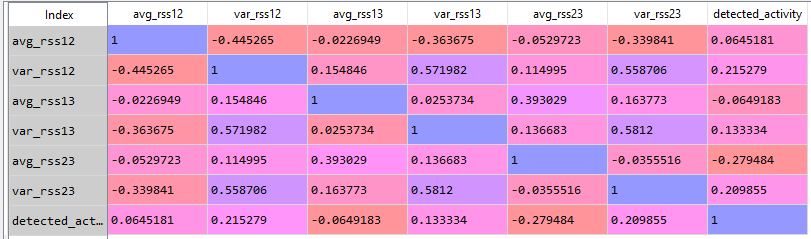
This dataset, obtained from UCI repository website, contains temporal data from a Wireless Sensor Network worn by an actor performing the activities: bending, cycling, lying down, sitting, standing, walking. For each activity 15 temporal sequences of input RSS data are present and contains the following information: avg\_rss12, var\_rss12, avg\_rss13, var\_rss13, avg\_rss23, var\_rss23 where avg and var are the mean and variance values over 250 ms of data, respectively. Target data is provided as the detected activity. The dataset contains 480 sequences, for a total number of 42240 instances.

* **Data preprocessing:**

1. Missing data: no missing value.
2. Encoding: we used the raking method for the “detected activity column” as below:

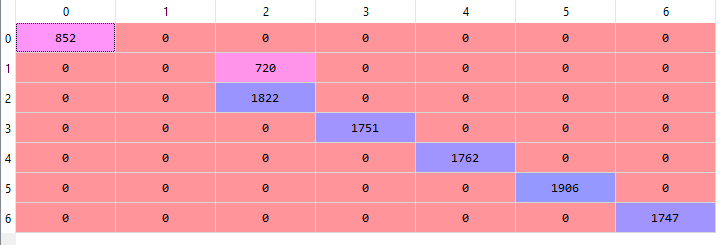
({'bending\_1':1,'bending\_2':2,'cycling':3, 'lying': 4, 'sitting':5, 'standing':6, 'walking':7})

1. Dropping unnecessary feature: we dropped “time” column, because it does not add any value to the data and the data is independent in time
2. Scaling: not required since the data are almost in the same range of values
3. Correlation analysis: using the spearman coefficient, there is a relatively weak correlation.

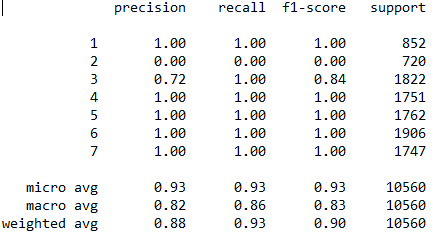


* **RFE implementation:** the selectee feature is only var\_rrs23 as the dominant feature.
* **Results:**

1. Confusion metrics: Showing an imbalanced data that all target in ‘bending\_2’ is predicted as ‘cycling’



1. Classification Report



3. Accuracy score = 0.9318181818181818