1. Study Design

This section is based mainly on the text provided by the team who has undertaken the experiments entitled “Human Activity Recognition Using Smartphones Dataset”, designed by Jorge L. Reyes-Ortiz, Davide Anguita, Alessandro Ghio, Luca Oneto of Smartlab - Non Linear Complex Systems Laboratory in Italy. See Ref (1)

The purpose of the experiments is to investigate in a feature selection process to allow for the recognition of the activities performed by human using the sensors like accelerator and gyroscope of the Samsung smartphone.

According to the authors, the experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist.

Using its embedded accelerometer and gyroscope, the authors captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.

The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window).

The sensor acceleration signal 🡺 body acceleration and gravity. From each window, a vector of features was obtained by calculating variables from the time and frequency domain.

For each record it is provided:

- Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.

- Triaxial Angular velocity from the gyroscope.

- A 561-feature vector with time and frequency domain variables.

- Its activity label.

- An identifier of the subject who carried out the experiment.

The dataset includes the following files:

'README.txt'

'features\_info.txt': Shows information about the variables used on the feature vector.

'features.txt': List of all features.

'activity\_labels.txt': Links the class labels with their activity name.

'train/X\_train.txt': Training set.

'train/y\_train.txt': Training labels.

'test/X\_test.txt': Test set.

'test/y\_test.txt': Test labels.

'train/subject\_train.txt': Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.

'test/subject\_test.txt': Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.

[1] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

1. Code Book

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **File/data set** | **Position or Column Nb** | **Type of Value** | **Values** | **Examples/Explanation** |
| Activity name | activity\_labels.txt |  | Categorical | 1 WALKING |  |
|  |  |  |  | 2 WALKING\_UPSTAIRS |  |
|  |  |  |  | 3 WALKING\_DOWNSTAIRS |  |
|  |  |  |  | 4 SITTING |  |
|  |  |  |  | 5 STANDING |  |
|  |  |  |  | 6 LAYING |  |
| Performed Activity Label | y\_train, y\_test | by row | Numeric | 1 to 6 | Refer to Activity Name |
| Subject | subject\_train.txt | by row | Identifier | 1 to 30 | subject performing the associated activity |
|  | subject\_test.txt |  |  |  |  |
| Feature Variable | features.txt |  | characters | Combination of: | fBodyAccJerk-min()-Y |
|  |  |  |  | Time or Frequency domain |  |
|  |  |  |  | Signal |  |
|  |  |  |  | Statistical Function |  |
|  |  |  |  | Direction |  |
| Measurement per subject and per activity | x\_train, x\_test |  | row of values of 561 feature variables | Numeric between -1 and +1 |  |

Below is the illustration of linking all the data of different files:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Step 1 |  |  |  |  |  |  |  |  |
|  | Measurement No | Activity Label | Subject Id | Feature var 1 | Feature var 2 |  |  | Feature var … |
|  | 1 | 5 | 1 | 0.4034743 | -0.0150744 |  |  |  |
|  |  |  |  |  |  |  |  |  |
| After linking different data tables | | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | Measurement No | Activity Name | Subject Id | tBodyAcc-mean-X | tBodyAcc-mean-Y |  |  | Feature var … |
|  | 1 | STANDING | 1 | 0.4034743 | -0.0150744 |  |  |  |