**Codebook – Average Human Activity Tracking Dataset**

**Date: June 2020**

**Project Description**

A project to analyse the human activity tracking experiment output and create a tidy dataset of the results.

The underlying 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, we 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.

**Data Processing**

**Notes / Collection of the Raw Data**

Raw data for this analysis was the data collected as part of the activity tracking trial and included the following datasets

- Measurement data from the trial with Triaxial acceleration from the accelerometer and the estimated body acceleration (561 observation types)

- 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.

**Processed Data / Tidy Data File**

**Datafile Creation**

The tidy datafile can be recreated using the following steps:

1. Download the repository / folder structure
2. Open Runanalysis.R file in R
3. Install relevant packages if not already installed. These are dplyr and rstudioapi
4. Run the Runanalysis.R script

**Data Processing Overview**

The Runanalysis.R script uses the trial data as an input to join together data from a test cohort and a training cohort of subjects.

This data is then combined together with labels for the features of the data and the types of activities undertaken.

Finally the script summarises multiple observations for each subject / activity pair by taking the mean of observations and outputs the resulting tidy dataset. For more information on the script please see the Readme.txt file

**Description of the Tidy Data set**

Dimensions: 180 rows with 88 variables

Data types: Data contains two columns holding the unique subject / activity pair identifier, with the remaining 86 being numeric variables containing the mean output from the activity monitors.

**Classification Variables**

**subject.id** - Identification number of the Subject undertaking the tests

Values from 1:30 representing the 30 subjects between age of 19-48 taking part in the trial

Class: integer

**activity** – Name of the type of activity the subject was undertaking – Activites include WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING. Data was converted from activity IDs 1-6 to activity names using the activity\_labels information in the raw data. Class: factor

**Measurement Variables**

Mean of observations for each subject and activity were taken for the 86 different measurements obtained for each subject from the trials. Data format for the following 86 columns is numeric and output has been normalised and are bounded between [-1, 1].

The measurement variables included were those which name mean or standard deviation within the feature vector label.

Feature variable titles were processed to remove special characters and include only lowercase variable names

**Time Data Variables**

**tbodyacc.mean.x** – mean of observed time data of body accelerometer measurements in the x direction

**tbodyacc.mean.y**– mean of observed time data of body accelerometer measurements in the y direction

**tbodyacc.mean.z**– mean of observed time data of body accelerometer measurements in the z direction

**tbodyacc.std.x**– standard deviation of observed time data of body accelerometer measurements in the x direction

**tbodyacc.std.y**– standard deviation of observed time data of body accelerometer measurements in the y direction

**tbodyacc.std.z**– standard deviation of observed time data of body accelerometer measurements in the z direction

**tgravityacc.mean.x**– mean of observed time data of gravity accelerometer measurements in the x direction

**tgravityacc.mean.y**– mean of observed time data of gravity accelerometer measurements in the y direction

**tgravityacc.mean.z**– mean of observed time data of gravity accelerometer measurements in the z direction

**tgravityacc.std.x**– standard deviation of observed time data of gravity accelerometer measurements in the x direction

**tgravityacc.std.y**– standard deviation of observed time data of gravity accelerometer measurements in the y direction

**tgravityacc.std.z**– standard deviation of observed time data of gravity accelerometer measurements in the z direction

**tbodyaccjerk.mean.x**– mean of observed time data of body jerk accelerometer measurements in the x direction

**tbodyaccjerk.mean.y**– mean of observed time data of body jerk accelerometer measurements in the y direction

**tbodyaccjerk.mean.z**– mean of observed time data of body jerk accelerometer measurements in the z direction

**tbodyaccjerk.std.x**– standard deviation of observed time data of body jerk accelerometer measurements in the x direction

**tbodyaccjerk.std.y**– standard deviation of observed time data of body jerk accelerometer measurements in the y direction

**tbodyaccjerk.std.z**– standard deviation of observed time data of body jerk accelerometer measurements in the z direction

**tbodygyro.mean.x**– mean of observed time data of body gyroscope measurements in the x direction

**tbodygyro.mean.y**– mean of observed time data of body gyroscope measurements in the y direction

**tbodygyro.mean.z**– mean of observed time data of body gyroscope measurements in the z direction

**tbodygyro.std.x**– standard deviation of observed time data of body gyroscope measurements in the x direction

**tbodygyro.std.y**– standard deviation of observed time data of body gyroscope measurements in the y direction

**tbodygyro.std.z**– standard deviation of observed time data of body gyroscope measurements in the z direction

**tbodygyrojerk.mean.x** - mean of observed time data of body jerk gyroscope measurements in the x direction

**tbodygyrojerk.mean.y**- mean of observed time data of body jerk gyroscope measurements in the y direction

**tbodygyrojerk.mean.z**- mean of observed time data of body jerk gyroscope measurements in the z direction

**tbodygyrojerk.std.x-** standard deviation of observed time data of body jerk gyroscope measurements in the x direction

**tbodygyrojerk.std.y**- standard deviation of observed time data of body jerk gyroscope measurements in the y direction

**tbodygyrojerk.std.z**- standard deviation of observed time data of body jerk gyroscope measurements in the z direction

**tbodyaccmag.mean** - mean of observed time data of for the magnitude of body data from the accelerometer

**tbodyaccmag.std**- standard deviation of observed time data of for the magnitude of body data from the accelerometer

**tgravityaccmag.mean**- mean of observed time data of for the magnitude of gravity data from the accelerometer

**tgravityaccmag.std**- standard deviation of observed time data of for the magnitude of gravity data from the accelerometer

**tbodyaccjerkmag.mean** - mean of observed time data of for the magnitude of body jerk data from the accelerometer

**tbodyaccjerkmag.std** – standard deviation of observed time data of for the magnitude of body jerk data from the accelerometer

**tbodygyromag.mean**- mean of observed time data of for the magnitude of body data from the gyroscope

**tbodygyromag.std**- standard deviation of observed time data of for the magnitude of body data from the gyroscope

**tbodygyrojerkmag.mean**- mean of observed time data of for the magnitude of body jerk data from the gyroscope

**tbodygyrojerkmag.std**- standard deviation of observed time data of for the magnitude of body jerk data from the gyroscope

**Frequency Data Variables**

**fbodyacc.mean.x** – mean of observed frequency data of body accelerometer measurements in the x direction

**fbodyacc.mean.y** – mean of observed frequency data of body accelerometer measurements in the y direction

**fbodyacc.mean.z**– mean of observed frequency data of body accelerometer measurements in the z direction

**fbodyacc.std.x**– standard deviation of observed frequency data of body accelerometer measurements in the x direction

**fbodyacc.std.y**– standard deviation of observed frequency data of body accelerometer measurements in the y direction

**fbodyacc.std.z**– standard deviation of observed frequency data of body accelerometer measurements in the z direction

**fbodyacc.meanfreq.x**– weighted average frequency of observed data of body accelerometer measurements in the x direction

**fbodyacc.meanfreq.y** - weighted average frequency of observed data of body accelerometer measurements in the y direction

**fbodyacc.meanfreq.z**- weighted average frequency of observed data of body accelerometer measurements in the z direction

**fbodyaccjerk.mean.x**– mean of observed frequency data of body jerk accelerometer measurements in the x direction

**fbodyaccjerk.mean.y**– mean of observed frequency data of body jerk accelerometer measurements in the y direction **fbodyaccjerk.mean.z**– mean of observed frequency data of body jerk accelerometer measurements in the z direction

**fbodyaccjerk.std.x**– standard deviation of observed frequency data of body jerk accelerometer measurements in the x direction

**fbodyaccjerk.std.y**– standard deviation of observed frequency data of body jerk accelerometer measurements in the y direction

**fbodyaccjerk.std.z**– standard deviation of observed frequency data of body jerk accelerometer measurements in the z direction

**fbodyaccjerk.meanfreq.x**– weighted average frequency of observed data of body jerk accelerometer measurements in the x direction

**fbodyaccjerk.meanfreq.y**– weighted average frequency of observed data of body jerk accelerometer measurements in the y direction

**fbodyaccjerk.meanfreq.z**– weighted average frequency of observed data of body jerk accelerometer measurements in the z direction

**fbodygyro.mean.x**– mean of observed frequency data of body gyrometer measurements in the x direction

**fbodygyro.mean.y**– mean of observed frequency data of body gyrometer measurements in the y direction

**fbodygyro.mean.z**– mean of observed frequency data of body gyrometer measurements in the z direction

**fbodygyro.std.x**– standard deviation of observed frequency data of body gyrometer measurements in the x direction

**fbodygyro.std.y**– standard deviation of observed frequency data of body gyrometer measurements in the y direction

**fbodygyro.std.z**– standard deviation of observed frequency data of body gyrometer measurements in the z direction

**fbodygyro.meanfreq.x** – weighted average frequency of observed data of body gyrometer measurements in the x direction

**fbodygyro.meanfreq.y** – weighted average frequency of observed data of body gyrometer measurements in the y direction

**fbodygyro.meanfreq.z**– weighted average frequency of observed data of body gyrometer measurements in the z direction

**fbodyaccmag.mean**– mean of observed frequency data of the magnitude of body accelerometer measurements

**fbodyaccmag.std**– standard deviation of observed frequency data of the magnitude of body accelerometer measurements

**fbodyaccmag.meanfreq**– weighted average frequency of observed frequency data of the magnitude of body accelerometer measurements

**fbodybodyaccjerkmag.mean**– mean of observed frequency data of the magnitude of body jerk accelerometer measurements

**fbodybodyaccjerkmag.std**– standard deviation of observed frequency data of the magnitude of body jerk accelerometer measurements

**fbodybodyaccjerkmag.meanfreq**– weighted average frequency of observed frequency data of the magnitude of body jerk accelerometer measurements

**fbodybodygyromag.mean**– mean of observed frequency data of the magnitude of body gyrometer measurements

**fbodybodygyromag.std**– standard deviation of observed frequency data of the magnitude of body gyrometer measurements

**fbodybodygyromag.meanfreq**– weighted average frequency of observed frequency data of the magnitude of body gyrometer measurements

**fbodybodygyrojerkmag.mean**– mean of observed frequency data of the magnitude of body jerk gyrometer measurements

**fbodybodygyrojerkmag.std**– standard deviation of observed frequency data of the magnitude of body jerk gyrometer measurements

**fbodybodygyrojerkmag.meanfreq**– weighted average frequency of observed frequency data of the magnitude of body jerk gyrometer measurements

**Angle Data**

**angletbodyaccmean.gravity** – mean of the angle between vectors using the body accelerometer measurements and gravity measurements

**angletbodyaccjerkmean.gravitymean** – mean of the angle between vectors using the body jerk accelerometer measurements and mean gravity measurements

**angletbodygyromean.gravitymean** – mean of the angle between vectors using the body gyrometer measurements and mean gravity measurements

**angletbodygyrojerkmean.gravitymean** – mean of the angle between vectors using the body jerk gyrometer measurements and mean gravity measurements

**anglex.gravitymean** – mean of the angle between vectors in the x-direction and the mean gravity measurements

**angley.gravitymean**– mean of the angle between vectors in the y-direction and the mean gravity measurements

**anglez.gravitymean**– mean of the angle between vectors in the z-direction and the mean gravity measurements

Source Data:

[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