Human Activity Recognition Using Smartphones Dataset - Version 2.0

Description of raw observations:

The raw data set contained measurements on sensor signals (accelerometer and gyroscope) on 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. The obtained dataset was randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.

Each record provided in the provided datasets measure the following:

- 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.
- Features are normalized and bounded within [-1,1].
- Each feature vector is a row on the text file.

The raw data was further processed into the training and test sets containing calculated mean, standard deviations and frequency measures for each subject and for each activity. This processed dataset was used to arrive at the two tidy datasets described below:

Description of the two tidy datasets and steps taken to create them

MeasuredMeanStd:

This dataset contains an aggergate of all measured means and standard deviations from training and test sets. The following steps were taken to create this dataset:

- Merged the X_test and y_test data and used the features file to name the variables
- Merged the X_train and y_train and used the features file to name the variables
- Merged the the training and test datasets into one dataset - this file contains 563 variables and 10299 observations
- Re-coded the test data variable to reflect the actual activity names for easier reading
- Extracted all the variables with mean and standard deviation measures aong with the subject and activity variables to create the final tidy dataset - this dataset contains 68 variables and 10299 observations

Average of Mean and Std measures:

This dataset is created by transforming the variables from the previous iteration. This was done by creating the average of each variable for each activity and each subject. The final dataset is arranged by summarizing by subject for each activity for each mean and standard deviation variable

---

Citation:

1. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human 2. 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

3. http://tgmstat.wordpress.com/2013/10/31/reshape-and-aggregate-data-with-the-r-package-reshape2/

4. Coursera discussion forums

5. Stack Overflow

Measure description:

The features selected for this database come from the accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ. These time domain signals (prefix 't' to denote time) were captured at a constant rate of 50 Hz. Then they were filtered using a median filter and a 3rd order low pass Butterworth filter with a corner frequency of 20 Hz to remove noise. Similarly, the acceleration signal was then separated into body and gravity acceleration signals (tBodyAcc-XYZ and tGravityAcc-XYZ) using another low pass Butterworth filter with a corner frequency of 0.3 Hz.

Subsequently, the body linear acceleration and angular velocity were derived in time to obtain Jerk signals (tBodyAccJerk-XYZ and tBodyGyroJerk-XYZ). Also the magnitude of these three-dimensional signals were calculated using the Euclidean norm (tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag, tBodyGyroJerkMag).

Finally a Fast Fourier Transform (FFT) was applied to some of these signals producing fBodyAcc-XYZ, fBodyAccJerk-XYZ, fBodyGyro-XYZ, fBodyAccJerkMag, fBodyGyroMag, fBodyGyroJerkMag. (Note the 'f' to indicate frequency domain signals).

These signals were used to estimate variables of the feature vector for each pattern:

'-XYZ' is used to denote 3-axial signals in the X, Y and Z directions.

tBodyAcc-XYZ

tGravityAcc-XYZ

tBodyAccJerk-XYZ

tBodyGyro-XYZ

tBodyGyroJerk-XYZ

tBodyAccMag

tGravityAccMag

tBodyAccJerkMag

tBodyGyroMag

tBodyGyroJerkMag

fBodyAcc-XYZ

fBodyAccJerk-XYZ

fBodyGyro-XYZ

fBodyAccMag

fBodyAccJerkMag

fBodyGyroMag

fBodyGyroJerkMag

There were a set of 17 variables that were estimated from these signals. We included the following two of them for our study:

mean(): Mean value

std(): Standard deviation

The following additional vectors obtained by averaging the signals in a signal window sample. These are used on the angle() variable:

gravityMean

tBodyAccMean

tBodyAccJerkMean

tBodyGyroMean

tBodyGyroJerkMean

List of actual variables contained in the tidy data sets for the analyses:

* subject
* activity
* tBodyAcc-mean()-X
* tBodyAcc-mean()-Y
* tBodyAcc-mean()-Z
* tGravityAcc-mean()-X
* tGravityAcc-mean()-Y
* tGravityAcc-mean()-Z
* tBodyAccJerk-mean()-X
* tBodyAccJerk-mean()-Y
* tBodyAccJerk-mean()-Z
* tBodyGyro-mean()-X
* tBodyGyro-mean()-Y
* tBodyGyro-mean()-Z
* tBodyGyroJerk-mean()-X
* tBodyGyroJerk-mean()-Y
* tBodyGyroJerk-mean()-Z
* tBodyAccMag-mean()
* tGravityAccMag-mean()
* tBodyAccJerkMag-mean()
* tBodyGyroMag-mean()
* tBodyGyroJerkMag-mean()
* fBodyAcc-mean()-X
* fBodyAcc-mean()-Y
* fBodyAcc-mean()-Z
* fBodyAccJerk-mean()-X
* fBodyAccJerk-mean()-Y
* fBodyAccJerk-mean()-Z

* fBodyGyro-mean()-X
* fBodyGyro-mean()-Y
* fBodyGyro-mean()-Z
* fBodyAccMag-mean()
* fBodyBodyAccJerkMag-mean()
* fBodyBodyGyroMag-mean()
* fBodyBodyGyroJerkMag-mean()
* tBodyAcc-std()-X
* tBodyAcc-std()-Y
* tBodyAcc-std()-Z
* tGravityAcc-std()-X
* tGravityAcc-std()-Y
* tGravityAcc-std()-Z
* tBodyAccJerk-std()-X
* tBodyAccJerk-std()-Y
* tBodyAccJerk-std()-Z
* tBodyGyro-std()-X
* tBodyGyro-std()-Y
* tBodyGyro-std()-Z
* tBodyGyroJerk-std()-X
* tBodyGyroJerk-std()-Y
* tBodyGyroJerk-std()-Z
* tBodyAccMag-std()
* tGravityAccMag-std()
* tBodyAccJerkMag-std()
* tBodyGyroMag-std()
* tBodyGyroJerkMag-std()
* fBodyAcc-std()-X
* fBodyAcc-std()-Y

* fBodyAcc-std()-Z
* fBodyAccJerk-std()-X
* fBodyAccJerk-std()-Y
* fBodyAccJerk-std()-Z
* fBodyGyro-std()-X
* fBodyGyro-std()-Y
* fBodyGyro-std()-Z
* fBodyAccMag-std()
* fBodyBodyAccJerkMag-std()
* fBodyBodyGyroMag-std()
* fBodyBodyGyroJerkMag-std()