Codebook - Aggregation of Human Activity Recognition Using Smartphones

Mike Quashne

1/2/2020

Description

The dataset labeled "dataset.txt" is a table of the mean of summary statistics for smartphone accelerometer observations for a given Subject (smartphone) during a given Activity (Walking, Sitting, etc.).

The source files observations from the accelerometers of 30 Samsung Galaxy S smartphones during six different motion activities. A description of the steps used to build dataset.txt can be found in README.md

Columns included in this dataset

The following columns are included:

- Subject — Integer Factor, 1 to 30. The number assigned to a specific smartphone from which observations were taken.

- Activity — String Factor. The motion activity being performed at the time of the observation. Motion activities are:

    - WALKING
    - WALKING_UPSTAIRS
    - WALKING_DOWNSTAIRS
    - SITTING
    - STANDING
    - LAYING

- Numbered Summary Statistics (remaining 66 columns) — Double, -1 to 1. Each column contains an average of the named variable for a given Subject and Activity. These variables represent a summary statistic (mean or standard deviation) related to standardized and normalized measurements from a smartphone accelerometer at the time of observation. The variables included as columns in this dataset are:

    - 1 tBodyAcc-mean()-X
    - 2 tBodyAcc-mean()-Y
    - 3 tBodyAcc-mean()-Z
    - 41 tGravityAcc-mean()-X
    - 42 tGravityAcc-mean()-Y
    - 43 tGravityAcc-mean()-Z
    - 81 tBodyAccJerk-mean()-X
    - 82 tBodyAccJerk-mean()-Y
    - 83 tBodyAccJerk-mean()-Z
    - 121 tBodyGyro-mean()-X
    - 122 tBodyGyro-mean()-Y

- 123 tBodyGyro-mean()-Z
- 161 tBodyGyroJerk-mean()-X
- 162 tBodyGyroJerk-mean()-Y
- 163 tBodyGyroJerk-mean()-Z
- 201 tBodyAccMag-mean()
- 214 tGravityAccMag-mean()
- 227 tBodyAccJerkMag-mean()
- 240 tBodyGyroMag-mean()
- 253 tBodyGyroJerkMag-mean()
- 266 fBodyAcc-mean()-X
- 267 fBodyAcc-mean()-Y
- 268 fBodyAcc-mean()-Z
- 345 fBodyAccJerk-mean()-X
- 346 fBodyAccJerk-mean()-Y
- 347 fBodyAccJerk-mean()-Z
- 424 fBodyGyro-mean()-X
- 425 fBodyGyro-mean()-Y
- 426 fBodyGyro-mean()-Z
- 503 fBodyAccMag-mean()
- 516 fBodyBodyAccJerkMag-mean()
- 529 fBodyBodyGyroMag-mean()
- 542 fBodyBodyGyroJerkMag-mean()
- 4 tBodyAcc-std()-X
- 5 tBodyAcc-std()-Y
- 6 tBodyAcc-std()-Z
- 44 tGravityAcc-std()-X
- 45 tGravityAcc-std()-Y
- 46 tGravityAcc-std()-Z
- 84 tBodyAccJerk-std()-X
- 85 tBodyAccJerk-std()-Y
- 86 tBodyAccJerk-std()-Z
- 124 tBodyGyro-std()-X
- 125 tBodyGyro-std()-Y
- 126 tBodyGyro-std()-Z
- 164 tBodyGyroJerk-std()-X
- 165 tBodyGyroJerk-std()-Y
- 166 tBodyGyroJerk-std()-Z
- 202 tBodyAccMag-std()
- 215 tGravityAccMag-std()
- 228 tBodyAccJerkMag-std()
- 241 tBodyGyroMag-std()
- 254 tBodyGyroJerkMag-std()
- 269 fBodyAcc-std()-X
- 270 fBodyAcc-std()-Y
- 271 fBodyAcc-std()-Z
- 348 fBodyAccJerk-std()-X
- 349 fBodyAccJerk-std()-Y
- 350 fBodyAccJerk-std()-Z
- 427 fBodyGyro-std()-X
- 428 fBodyGyro-std()-Y
- 429 fBodyGyro-std()-Z
- 504 fBodyAccMag-std()
- 517 fBodyBodyAccJerkMag-std()
- 530 fBodyBodyGyroMag-std()

– 543 fBodyBodyGyroJerkMag-std()

Data Description from source files

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

The set of variables that were estimated from these signals are:

- mean(): Mean value
- std(): Standard deviation