Data Source:

About the data

Human Activity Recognition Using Smartphones Dataset

Version 1.0

==================================================================

Jorge L. Reyes-Ortiz, Davide Anguita, Alessandro Ghio, Luca Oneto.

Smartlab - Non Linear Complex Systems Laboratory

DITEN - Universit‡ degli Studi di Genova.

Via Opera Pia 11A, I-16145, Genoa, Italy.

activityrecognition@smartlab.ws

www.smartlab.ws

The list of files available:

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

-1’features.txt': List of all features.

-2 'activity\_labels.txt': Links the class labels with their activity name.

-3 'train/X\_train.txt': Training set.

-4 'train/y\_train.txt': Training labels.

-5’test/X\_test.txt': Test set.

-6 test/y\_test.txt': Test labels.

The following files are available for the train and test data. Their descriptions are equivalent.

-7 'train/subject\_train.txt': Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.

Feature Selection

=================

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

mad(): Median absolute deviation

max(): Largest value in array

min(): Smallest value in array

sma(): Signal magnitude area

energy(): Energy measure. Sum of the squares divided by the number of values.

iqr(): Interquartile range

entropy(): Signal entropy

arCoeff(): Autorregresion coefficients with Burg order equal to 4

correlation(): correlation coefficient between two signals

maxInds(): index of the frequency component with largest magnitude

meanFreq(): Weighted average of the frequency components to obtain a mean frequency

skewness(): skewness of the frequency domain signal

kurtosis(): kurtosis of the frequency domain signal

bandsEnergy(): Energy of a frequency interval within the 64 bins of the FFT of each window.

angle(): Angle between to vectors.

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

gravityMean

tBodyAccMean

tBodyAccJerkMean

tBodyGyroMean

tBodyGyroJerkMean

The complete list of variables of each feature vector is available in 'features.txt'

TidyData Table

This table used 7 of the tables in the original data source.

1. The test and train data were compiled by adding row specific record information and column names.
2. We then built an index to filter and select a subset of columns
3. Reformatted the column names to make them more succint
4. Merged additional activity labels and the train and test sample data together.
5. The data table had to ‘melted’ and ‘cast’ to summarize it.
6. I chose the summarization that created a unique record for the Subject- Activity pair. I also had a column for flagging train or test data - in preparation for model build.
7. Changing the case of the column names was required before the melt/ cast functions could be used.

| Columns | Measure Category | Format |
| --- | --- | --- |
| Sample | Train or Test sample Source | Character |
| ActivityType | 1-6 | Integer |
| SubjectActivity | 1-30 | Subjects Number in study |
| tBodyAcc.Mean.X | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAcc.Mean.Y | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAcc.Mean.Z | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAcc.Std.X | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAcc.Std.Y | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAcc.Std.Z | At time Raw Accelerator Signals (Hz) | Numeric |
| tGravityAcc.Mean.X | Raw Gravity Accelerator at time | Numeric |
| tGravityAcc.Mean.Y | Raw Gravity Accelerator at time | Numeric |
| tGravityAcc.Mean.Z | Raw Gravity Accelerator at time | Numeric |
| tGravityAcc.Std.X | Raw Gravity Accelerator at time | Numeric |
| tGravityAcc.Std.Y | Raw Gravity Accelerator at time | Numeric |
| tGravityAcc.Std.Z | Raw Gravity Accelerator at time | Numeric |
| tBodyAccJerk.Mean.X | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerk.Mean.Y | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerk.Mean.Z | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerk.Std.X | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerk.Std.Y | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerk.Std.Z | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyGyro.Mean.X | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyro.Mean.Y | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyro.Mean.Z | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyro.Std.X | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyro.Std.Y | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyro.Std.Z | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Mean.X | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Mean.Y | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Mean.Z | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Std.X | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Std.Y | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerk.Std.Z | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyAccMag.Mean | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccMag.Std | At time Raw Accelerator Signals (Hz) | Numeric |
| tGravityAccMag.Mean | At time Raw Accelerator Signals (Hz) | Numeric |
| tGravityAccMag.Std | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerkMag.Mean | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyAccJerkMag.Std | At time Raw Accelerator Signals (Hz) | Numeric |
| tBodyGyroMag.Mean | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroMag.Std | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerkMag.Mean | At time Raw Gyroscope Signals (Hz) | Numeric |
| tBodyGyroJerkMag.Std | At time Raw Gyroscope Signals (Hz) | Numeric |
| fBodyAcc.Mean.X | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAcc.Mean.Y | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAcc.Mean.Z | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAcc.Std.X | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAcc.Std.Y | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAcc.Std.Z | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Mean.X | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Mean.Y | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Mean.Z | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Std.X | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Std.Y | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccJerk.Std.Z | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyGyro.Mean.X | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyGyro.Mean.Y | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyGyro.Mean.Z | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyGyro.Std.X | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyGyro.Std.Y | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyGyro.Std.Z | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyAccMag.Mean | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyAccMag.Std | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyBodyAccJerkMag.Mean | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyBodyAccJerkMag.Std | FF transformed Accelerator Signal (Hz) | Numeric |
| fBodyBodyGyroMag.Mean | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyBodyGyroMag.Std | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyBodyGyroJerkMag.Mean | FF transformed Gyroscope Signal (Hz) | Numeric |
| fBodyBodyGyroJerkMag.Std | FF transformed Gyroscope Signal (Hz) | Numeric |
| Activity | 6 levels of activity tested | Character |

Table TidyAvg

This Table takes the mean of the above table by Subject and Activity.

The data represents a summary table of all the above, with the exception of ‘Sample’ and ‘Activity Type’.