CODE BOOK – HUMAN ACTIVITY RECOGNITION USING SMARTPHONES DATASET

Study Design

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, 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.

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, which has gravitational and body motion components, was separated using a Butterworth low-pass filter into body acceleration and gravity. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of features was obtained by calculating variables from the time and frequency domain.

Data Dictionary

subject.id 2 Integer class

Subject Identifier

1..30 Unique identifier for each experiment participant in a group of 30 volunteers within an age bracket of 19-48 years.

activity.label 18 Character class

Activity Label Each volunteer performed 6 activities wearing the smartphone on the waist:

1. WALKING
2. WALKING_UPSTAIRS
3. WALKING_DOWNSTAIRS
4. SITTING
5. STANDING
6. LAYING

Features. Tokens are combined to form features (variables of the tidy1 and tidy2 datasets) of numeric class.

Token	Description
acc	Acceleration type signal taken from an accelerometer.
body	Signal based on the body motion of an experiment participant, the 1st of 2

components derived from the time-based, acceleration signal received by the smartphone's accelerometer.

freq	Measurement based on a frequency-domain, taken as a Fast Fourier Transform of the time-based signals.
gravity	Signal based on gravity, the force that attracts a body towards the center of the earth. Gravity is the 2nd of the 2 measurement components derived from the time-based, acceleration signal received by the smartphone's accelerometer.
gyro	Angular velocity signal from a gyroscope.
jerk	Jerk-type signals derived from time-based, body linear acceleration and angular velocity.
mag	Magnitude of 3-dimensional, time-based signals calculated using the Euclidean norm.
mean	Mean value estimated from the signals.
std	Standard deviation estimated from the signals.
time	Indicates a time-domain type signal.
x	Denotes axial signals in the X-direction.
y	Denotes axial signals in the Y-direction.
z	Denotes axial signals in the Z-direction.

Notes:
======
- Features are normalized and bounded within [-1,1].
- In the tidy1 dataset, feature variables are limited to only the measurements on the mean and standard deviation for each measurement.
- Each feature vector is a row or measurement on the tidy1 dataset.
- The tidy1 dataset has 10,299 rows or measurements and 68 columns or variables (subject.id, activity.label and 66 feature variables).
- The tidy2 dataset has 180 summary rows and the same variables as the tidy1 dataset.
- The tidy2 dataset (derived from the tidy1 dataset) has values that are the average of each feature variable for each activity.label and each subject.id.