STUDY DESIGN

Human Activity Recognition Using Smartphones Dataset

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. See 'features_info.txt' for more details.

All can be downloaded from:
https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip

Reference for data set:
[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

INFORMATION ON THE SUMMARY DATA CHOICES MADE

Data sets from this zipped file that are needed in the working directory to run the run_analysis.R script:
features.txt
subject_test.txt
subject_train.txt
X_test.txt
X_train.txt
y_test.txt
y_train.txt

The run_analysis.R script takes the above 7 data sets and merges them so that each row is labelled with the subject and activity. In addition, the columns are given variable names, which have been made all lower case and have had "_" and "()" removed. This makes the column names fit into a tidy data set as these were the parameters given in the course for

variable names in tidy data sets (all lower case, no _, no spaces). The variable names are understandable to anyone familiar with the study.

The original data set is subset to only contain data on mean and standard deviation as described in the assignment. In this case, any variable name that contained "mean" or "std" anywhere in the given variable name was included, since the assignment directions did not specify further. In this case, if some of these variables are not needed, it would be easy to remove them.

The script also takes the refined data set and creates a second, independent tidy data set with the average of each variable for each activity and each subject. In this case, the newly created data set is a long data set where there is only one column, which is a mean, and each row represents a subject, activity, and one of the variables, as marked.

The final data set for parts 1-4 of the assignment will be written to a file called DescData.txt. The final data set for part 5 of the assignment (the second, independent tidy data set with the average of each variable for each activity and each subject) will be written to a file called tidy.txt.

CODE BOOK

81 Variables:
"subjects"   30 subjects, labelled 1-30
"activities"   each subject performed 6 activities: walking, walking upstairs, walking downstairs, sitting, standing, and laying
The following 79 measurements were taken:
 "tbodyaccmeanx"
"tbodyaccmeany"
 "tbodyaccmeanz"
"tgravityaccmeanx"
 "tgravityaccmeany"
"tgravityaccmeanz"
 "tbodyaccjerkmeanx"
"tbodyaccjerkmeany"
"tbodyaccjerkmeanz"
"tbodygyromeanx"
"tbodygyromeany"
"tbodygyromeanz"
"tbodygyrojerkmeanx"
"tbodygyrojerkmeany"
"tbodygyrojerkmeanz"
"tbodyaccmagmean"
"tgravityaccmagmean"
"tbodyaccjerkmagmean"
"tbodygyromagmean"

"tbodygyrojerkmagmean"
"fbodyaccmeanx"
"fbodyaccmeany"
"fbodyaccmeanz"
"fbodyaccmeanfreqx"
"fbodyaccmeanfreqy"
"fbodyaccmeanfreqz"
"fbodyaccjerkmeanx"
"fbodyaccjerkmeany"
"fbodyaccjerkmeanz"
"fbodyaccjerkmeanfreqx"
"fbodyaccjerkmeanfreqy"
"fbodyaccjerkmeanfreqz"
"fbodygyromeanx"
"fbodygyromeany"
"fbodygyromeanz"
"fbodygyromeanfreqx"
"fbodygyromeanfreqy"
"fbodygyromeanfreqz"
"fbodyaccmagmean"
"fbodyaccmagmeanfreq"
"fbodybodyaccjerkmagmean"
"fbodybodyaccjerkmagmeanfreq"
"fbodybodygyromagmean"
"fbodybodygyromagmeanfreq"
"fbodybodygyrojerkmagmean"
"fbodybodygyrojerkmagmeanfreq"
"tbodyaccstdx"
"tbodyaccstdy"
"tbodyaccstdz"
"tgravityaccstdx"
"tgravityaccstdy"
"tgravityaccstdz"
"tbodyaccjerkstdx"
"tbodyaccjerkstdy"
"tbodyaccjerkstdz"
"tbodygyrostdx"
"tbodygyrostdy"
"tbodygyrostdz"
"tbodygyrojerkstdx"
"tbodygyrojerkstdy"
"tbodygyrojerkstdz"
"tbodyaccmagstd"
"tgravityaccmagstd"
"tbodyaccjerkmagstd"
"tbodygyromagstd"

"tbodygyrojerkmagstd"
"fbodyaccstdx"
"fbodyaccstdy"
"fbodyaccstdz"
"fbodyaccjerkstdx"
"fbodyaccjerkstdy"
"fbodyaccjerkstdz"
"fbodygyrostdx"
"fbodygyrostdy"
"fbodygyrostdz"
"fbodyaccmagstd"
"fbodybodyaccjerkmagstd"
"fbodybodygyromagstd"
"fbodybodygyrojerkmagstd"

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

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

- 'train/Inertial Signals/total_acc_x_train.txt': The acceleration signal from the smartphone accelerometer X axis in standard gravity units 'g'. Every row shows a 128 element vector. The same description applies for the 'total_acc_x_train.txt' and 'total_acc_z_train.txt' files for the Y and Z axis.

- 'train/Inertial Signals/body_acc_x_train.txt': The body acceleration signal obtained by subtracting the gravity from the total acceleration.

- 'train/Inertial Signals/body_gyro_x_train.txt': The angular velocity vector measured by the gyroscope for each window sample. The units are radians/second.