

Getting and Cleaning Data

Sizwe Mashao

13 May 2017

Introduction

This codebook is submitted as part of the required documents for the Course on Getting and Cleaning Data.

The data

The data comes from experiments that 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. More information on the data and the data itself can be found at the following link: <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

Variables

The cleaned data set contains the following variables

- [1] "activity_ID" "subject_ID" "timeBodyAccelerator-mean()-X"
- [4] "timeBodyAccelerator-mean()-Y" "timeBodyAccelerator-mean()-Z" "timeBodyAccelerator-std()-X"
- [7] "timeBodyAccelerator-std()-Y" "timeBodyAccelerator-std()-Z" "timeGravityAccelerator-mean()-X"
- [10] "timeGravityAccelerator-mean()-Y" "timeGravityAccelerator-mean()-Z" "timeGravityAccelerator-std()-X"
- [13] "timeGravityAccelerator-std()-Y" "timeGravityAccelerator-std()-Z" "timeBodyAcceleratorJerk-mean()-X"
- [16] "timeBodyAcceleratorJerk-mean()-Y" "timeBodyAcceleratorJerk-mean()-Z" "timeBodyAcceleratorJerk-std()-X"
- [19] "timeBodyAcceleratorJerk-std()-Y" "timeBodyAcceleratorJerk-std()-Z" "timeBodyGyroscope-mean()-X"
- [22] "timeBodyGyroscope-mean()-Y" "timeBodyGyroscope-mean()-Z" "timeBodyGyroscope-std()-X"
- [25] "timeBodyGyroscope-std()-Y" "timeBodyGyroscope-std()-Z" "timeBodyGyroscopeJerk-mean()-X"
- [28] "timeBodyGyroscopeJerk-mean()-Y" "timeBodyGyroscopeJerk-mean()-Z" "timeBodyGyroscopeJerk-std()-X"
- [31] "timeBodyGyroscopeJerk-std()-Y" "timeBodyGyroscopeJerk-std()-Z" "timeBodyAcceleratorMagnitude-mean()"

[34] "timeBodyAcceleratorMagnitude-std()" "timeGravityAcceleratorMagnitude-mean()" "timeGravityAcceleratorMagnitude-std()"

[37] "timeBodyAcceleratorJerkMagnitude-mean()" "timeBodyAcceleratorJerkMagnitude-std()" "timeBodyGyroscopeMagnitude-mean()"

[40] "timeBodyGyroscopeMagnitude-std()" "timeBodyGyroscopeJerkMagnitude-mean()" "timeBodyGyroscopeJerkMagnitude-std()"

[43] "frequencyBodyAccelerator-mean()-X" "frequencyBodyAccelerator-mean()-Y" "frequencyBodyAccelerator-mean()-Z"

[46] "frequencyBodyAccelerator-std()-X" "frequencyBodyAccelerator-std()-Y" "frequencyBodyAccelerator-std()-Z"

[49] "frequencyBodyAccelerator-meanFreq()-X" "frequencyBodyAccelerator-meanFreq()-Y" "frequencyBodyAccelerator-meanFreq()-Z"

[52] "frequencyBodyAcceleratorJerk-mean()-X" "frequencyBodyAcceleratorJerk-mean()-Y" "frequencyBodyAcceleratorJerk-mean()-Z"

[55] "frequencyBodyAcceleratorJerk-std()-X" "frequencyBodyAcceleratorJerk-std()-Y" "frequencyBodyAcceleratorJerk-std()-Z"

[58] "frequencyBodyAcceleratorJerk-meanFreq()-X" "frequencyBodyAcceleratorJerk-meanFreq()-Y" "frequencyBodyAcceleratorJerk-meanFreq()-Z"

[61] "frequencyBodyGyroscope-mean()-X" "frequencyBodyGyroscope-mean()-Y" "frequencyBodyGyroscope-mean()-Z"

[64] "frequencyBodyGyroscope-std()-X" "frequencyBodyGyroscope-std()-Y" "frequencyBodyGyroscope-std()-Z"

[67] "frequencyBodyGyroscope-meanFreq()-X" "frequencyBodyGyroscope-meanFreq()-Y" "frequencyBodyGyroscope-meanFreq()-Z"

[70] "frequencyBodyAcceleratorMagnitude-mean()" "frequencyBodyAcceleratorMagnitude-std()" "frequencyBodyAcceleratorMagnitude-meanFreq()"

[73] "frequencyBodyBodyAcceleratorJerkMagnitude-mean()" "frequencyBodyBodyAcceleratorJerkMagnitude-std()" "frequencyBodyBodyAcceleratorJerkMagnitude-meanFreq()" [76] "frequencyBodyBodyGyroscopeMagnitude-mean()" "frequencyBodyBodyGyroscopeMagnitude-std()" "frequencyBodyBodyGyroscopeMagnitude-meanFreq()"

[79] "frequencyBodyBodyGyroscopeJerkMagnitude-mean()" "frequencyBodyBodyGyroscopeJerkMagnitude-std()" "frequencyBodyBodyGyroscopeJerkMagnitude-meanFreq()"

>