The codebook

YT

4/11/2020

The codebook

Project: Human Activity Recognition Using Smartphones Data Set

It describes the variables, the data, and transformations that were performed to clean up the data.

1. Raw Data

1. Test files

1. subject_test.txt - Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.
2. X_test.txt - Test data set
3. y_test.txt - Test activity labels
4. Inertial signals with 9 files

2. Train files

1. subject_train.txt - Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.
2. X_train.txt - Train data set
3. y_train.txt - Train activity labels
4. Inertial signals with 9 files

UNITS: The acceleration signal from the smartphone accelerometer X axis in standard gravity units 'g'.

The body acceleration signal obtained by subtracting the gravity from the total acceleration.

The units for angular velocity vector measured by the gyroscope for each window sample are radians/second.

3. General files

'README.txt' General information about the project. 'features_info.txt': Shows information about the variables used on the feature vector. 'features.txt': List of all features. 'activity_labels.txt': Links the class labels with their activity name.

2. Summary of transformations

The step-by-step process with comments is provided in the run_analysis.R script. 1. Inertial signals (18) files were excluded, since they were not a part of the task. 2. Loaded the remaining 6 files into R (3 for the test data set and 3 for the train data set). 3. They were combined into one single data set with renamed columns using features.txt. 4. Only the measurements on the mean and standard deviation for each measurement were extracted. A total of 66 measurements. meanFreq measurements were not included. 5. Values in the column with activity codes (1-6) were renamed with descriptive names from activity_labels.txt. 6. All columns were renamed with descriptive names. 7. Final tidy data set was grouped by subject (1 through 30) and then activity and means of all remaining columns were found.

3. Processed Data

Tidy data sets (180x68) consists of:

[1] "subject"
[2] "activity"
[3] "timeBodyAcceleration_mean_X"
[4] "timeBodyAcceleration_mean_Y"
[5] "timeBodyAcceleration_mean_Z"
[6] "timeBodyAcceleration_std_X"
[7] "timeBodyAcceleration_std_Y"
[8] "timeBodyAcceleration_std_Z"
[9] "tGravityAcceleration_mean_X"
[10] "tGravityAcceleration_mean_Y"
[11] "tGravityAcceleration_mean_Z"
[12] "tGravityAcceleration_std_X"
[13] "tGravityAcceleration_std_Y"
[14] "tGravityAcceleration_std_Z"
[15] "timeBodyAccelerationJerk_mean_X"
[16] "timeBodyAccelerationJerk_mean_Y"
[17] "timeBodyAccelerationJerk_mean_Z"
[18] "timeBodyAccelerationJerk_std_X"
[19] "timeBodyAccelerationJerk_std_Y"
[20] "timeBodyAccelerationJerk_std_Z"
[21] "timeBodyGyroscope_mean_X"
[22] "timeBodyGyroscope_mean_Y"
[23] "timeBodyGyroscope_mean_Z"
[24] "timeBodyGyroscope_std_X"
[25] "timeBodyGyroscope_std_Y"
[26] "timeBodyGyroscope_std_Z"
[27] "timeBodyGyroscopeJerk_mean_X"
[28] "timeBodyGyroscopeJerk_mean_Y"
[29] "timeBodyGyroscopeJerk_mean_Z"
[30] "timeBodyGyroscopeJerk_std_X"
[31] "timeBodyGyroscopeJerk_std_Y"
[32] "timeBodyGyroscopeJerk_std_Z"
[33] "timeBodyAccelerationMagnitude_mean"
[34] "timeBodyAccelerationMagnitude_std"
[35] "tGravityAccelerationMagnitude_mean"
[36] "tGravityAccelerationMagnitude_std"
[37] "timeBodyAccelerationJerkMagnitude_mean"
[38] "timeBodyAccelerationJerkMagnitude_std"
[39] "timeBodyGyroscopeMagnitude_mean"

[40] “timeBodyGyroscopeMagnitude_std”
[41] “timeBodyGyroscopeJerkMagnitude_mean”
[42] “timeBodyGyroscopeJerkMagnitude_std”
[43] “frequencyBodyAcceleration_mean_X”
[44] “frequencyBodyAcceleration_mean_Y”
[45] “frequencyBodyAcceleration_mean_Z”
[46] “frequencyBodyAcceleration_std_X”
[47] “frequencyBodyAcceleration_std_Y”
[48] “frequencyBodyAcceleration_std_Z”
[49] “frequencyBodyAccelerationJerk_mean_X”
[50] “frequencyBodyAccelerationJerk_mean_Y”
[51] “frequencyBodyAccelerationJerk_mean_Z”
[52] “frequencyBodyAccelerationJerk_std_X”
[53] “frequencyBodyAccelerationJerk_std_Y”
[54] “frequencyBodyAccelerationJerk_std_Z”
[55] “frequencyBodyGyroscope_mean_X”
[56] “frequencyBodyGyroscope_mean_Y”
[57] “frequencyBodyGyroscope_mean_Z”
[58] “frequencyBodyGyroscope_std_X”
[59] “frequencyBodyGyroscope_std_Y”
[60] “frequencyBodyGyroscope_std_Z”
[61] “frequencyBodyAccelerationMagnitude_mean”
[62] “frequencyBodyAccelerationMagnitude_std”
[63] “frequencyBodyAccelerationJerkMagnitude_mean” [64] “frequencyBodyAccelerationJerkMagnitude_std” [65] “frequencyBodyGyroscopeMagnitude_mean”
[66] “frequencyBodyGyroscopeMagnitude_std”
[67] “frequencyBodyGyroscopeJerkMagnitude_mean”
[68] “frequencyBodyGyroscopeJerkMagnitude_std”