

Codebook

Lui

July 19, 2015

The file Run_Analysis.R contains the main function Run_Analysis(). The function writes the resulting dataset to a file called GettingCleaningDataProject.txt and returns it to the caller as well. For more details look the code or the readme.md file.

This is a portion of result of a call to Run_Analysis():

```
## Activity Subject tBodyAcc.mean...X tBodyAcc.mean...Y
## 1 Walking      1      0.2773308      -0.01738382
## 2 Walking      2      0.2764266      -0.01859492
## 3 Walking      3      0.2755675      -0.01717678
## 4 Walking      4      0.2785820      -0.01483995
## 5 Walking      5      0.2778423      -0.01728503
```

```
dataset <- Run_Analysis()
```

The following table shows the columns of the resulting dataset compared to the original. Each column in the resulting dataset is the average of the original calculated per activity per subject.

##	Original.Dataset.columns	Dataset.Columns
## 1	<NA>	Activity
## 2	<NA>	Subject
## 3	tBodyAcc-mean()-X	tBodyAcc.mean...X
## 4	tBodyAcc-mean()-Y	tBodyAcc.mean...Y
## 5	tBodyAcc-mean()-Z	tBodyAcc.mean...Z
## 6	tBodyAcc-std()-X	tBodyAcc.std...X
## 7	tBodyAcc-std()-Y	tBodyAcc.std...Y
## 8	tBodyAcc-std()-Z	tBodyAcc.std...Z
## 9	tGravityAcc-mean()-X	tGravityAcc.mean...X
## 10	tGravityAcc-mean()-Y	tGravityAcc.mean...Y
## 11	tGravityAcc-mean()-Z	tGravityAcc.mean...Z
## 12	tGravityAcc-std()-X	tGravityAcc.std...X
## 13	tGravityAcc-std()-Y	tGravityAcc.std...Y
## 14	tGravityAcc-std()-Z	tGravityAcc.std...Z
## 15	tBodyAccJerk-mean()-X	tBodyAccJerk.mean...X
## 16	tBodyAccJerk-mean()-Y	tBodyAccJerk.mean...Y
## 17	tBodyAccJerk-mean()-Z	tBodyAccJerk.mean...Z
## 18	tBodyAccJerk-std()-X	tBodyAccJerk.std...X
## 19	tBodyAccJerk-std()-Y	tBodyAccJerk.std...Y
## 20	tBodyAccJerk-std()-Z	tBodyAccJerk.std...Z
## 21	tBodyGyro-mean()-X	tBodyGyro.mean...X
## 22	tBodyGyro-mean()-Y	tBodyGyro.mean...Y
## 23	tBodyGyro-mean()-Z	tBodyGyro.mean...Z
## 24	tBodyGyro-std()-X	tBodyGyro.std...X
## 25	tBodyGyro-std()-Y	tBodyGyro.std...Y
## 26	tBodyGyro-std()-Z	tBodyGyro.std...Z
## 27	tBodyGyroJerk-mean()-X	tBodyGyroJerk.mean...X
## 28	tBodyGyroJerk-mean()-Y	tBodyGyroJerk.mean...Y

## 29	tBodyGyroJerk-mean()-Z	tBodyGyroJerk.mean...Z
## 30	tBodyGyroJerk-std()-X	tBodyGyroJerk.std...X
## 31	tBodyGyroJerk-std()-Y	tBodyGyroJerk.std...Y
## 32	tBodyGyroJerk-std()-Z	tBodyGyroJerk.std...Z
## 33	tBodyAccMag-mean()	tBodyAccMag.mean..
## 34	tBodyAccMag-std()	tBodyAccMag.std..
## 35	tGravityAccMag-mean()	tGravityAccMag.mean..
## 36	tGravityAccMag-std()	tGravityAccMag.std..
## 37	tBodyAccJerkMag-mean()	tBodyAccJerkMag.mean..
## 38	tBodyAccJerkMag-std()	tBodyAccJerkMag.std..
## 39	tBodyGyroMag-mean()	tBodyGyroMag.mean..
## 40	tBodyGyroMag-std()	tBodyGyroMag.std..
## 41	tBodyGyroJerkMag-mean()	tBodyGyroJerkMag.mean..
## 42	tBodyGyroJerkMag-std()	tBodyGyroJerkMag.std..
## 43	fBodyAcc-mean()-X	fBodyAcc.mean...X
## 44	fBodyAcc-mean()-Y	fBodyAcc.mean...Y
## 45	fBodyAcc-mean()-Z	fBodyAcc.mean...Z
## 46	fBodyAcc-std()-X	fBodyAcc.std...X
## 47	fBodyAcc-std()-Y	fBodyAcc.std...Y
## 48	fBodyAcc-std()-Z	fBodyAcc.std...Z
## 49	fBodyAcc-meanFreq()-X	fBodyAcc.meanFreq...X
## 50	fBodyAcc-meanFreq()-Y	fBodyAcc.meanFreq...Y
## 51	fBodyAcc-meanFreq()-Z	fBodyAcc.meanFreq...Z
## 52	fBodyAccJerk-mean()-X	fBodyAccJerk.mean...X
## 53	fBodyAccJerk-mean()-Y	fBodyAccJerk.mean...Y
## 54	fBodyAccJerk-mean()-Z	fBodyAccJerk.mean...Z
## 55	fBodyAccJerk-std()-X	fBodyAccJerk.std...X
## 56	fBodyAccJerk-std()-Y	fBodyAccJerk.std...Y
## 57	fBodyAccJerk-std()-Z	fBodyAccJerk.std...Z
## 58	fBodyAccJerk-meanFreq()-X	fBodyAccJerk.meanFreq...X
## 59	fBodyAccJerk-meanFreq()-Y	fBodyAccJerk.meanFreq...Y
## 60	fBodyAccJerk-meanFreq()-Z	fBodyAccJerk.meanFreq...Z
## 61	fBodyGyro-mean()-X	fBodyGyro.mean...X
## 62	fBodyGyro-mean()-Y	fBodyGyro.mean...Y
## 63	fBodyGyro-mean()-Z	fBodyGyro.mean...Z
## 64	fBodyGyro-std()-X	fBodyGyro.std...X
## 65	fBodyGyro-std()-Y	fBodyGyro.std...Y
## 66	fBodyGyro-std()-Z	fBodyGyro.std...Z
## 67	fBodyGyro-meanFreq()-X	fBodyGyro.meanFreq...X
## 68	fBodyGyro-meanFreq()-Y	fBodyGyro.meanFreq...Y
## 69	fBodyGyro-meanFreq()-Z	fBodyGyro.meanFreq...Z
## 70	fBodyAccMag-mean()	fBodyAccMag.mean..
## 71	fBodyAccMag-std()	fBodyAccMag.std..
## 72	fBodyAccMag-meanFreq()	fBodyAccMag.meanFreq..
## 73	fBodyBodyAccJerkMag-mean()	fBodyBodyAccJerkMag.mean..
## 74	fBodyBodyAccJerkMag-std()	fBodyBodyAccJerkMag.std..
## 75	fBodyBodyAccJerkMag-meanFreq()	fBodyBodyAccJerkMag.meanFreq..
## 76	fBodyBodyGyroMag-mean()	fBodyBodyGyroMag.mean..
## 77	fBodyBodyGyroMag-std()	fBodyBodyGyroMag.std..
## 78	fBodyBodyGyroMag-meanFreq()	fBodyBodyGyroMag.meanFreq..
## 79	fBodyBodyGyroJerkMag-mean()	fBodyBodyGyroJerkMag.mean..
## 80	fBodyBodyGyroJerkMag-std()	fBodyBodyGyroJerkMag.std..
## 81	fBodyBodyGyroJerkMag-meanFreq()	fBodyBodyGyroJerkMag.meanFreq..

The original variables were calculated as follows (extract from the file “features_info.txt ”):

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:

Note: only std() and mean() value columns were considered from the original dataset for a total of 79 columns out of the original 561.

- 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(): Autorregression 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 two 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’