DATA DICTIONARY

How data was recorded:

By using the sensors(Gyroscope and accelerometer) in a smartphone, they have captured '3-axial linear acceleration'(*tAcc-XYZ*) from accelerometer and '3-axial angular velocity' (*tGyro-XYZ*) from Gyroscope with several variations.

1. prefix 't' in those metrics denotes time.
2. suffix 'XYZ' represents 3-axial signals in X , Y, and Z directions

Feature names

1) These sensor signals are preprocessed by applying noise filters and then sampled in fixed-width windows(sliding windows) of 2.56 seconds each with 50% overlap. ie., each window has 128 readings.
2) From Each window, a feature vector was obtianed by calculating variables from the time and frequency domain.
3) The accelertion signal was saperated into Body and Gravity acceleration signals(___tBodyAcc-XYZ___ and ___tGravityAcc-XYZ___) using some low pass filter with corner frequecy of 0.3Hz.
4) After that, the body linear acceleration and angular velocity were derived in time to obtian _jerk signals_ (___tBodyAccJerk-XYZ___ and ___tBodyGyroJerk-XYZ___).
5) The magnitude of these 3-dimensional signals were calculated using the Euclidian norm. This magnitudes are represented as features with names like _tBodyAccMag_, _tGravityAccMag_, _tBodyAccJerkMag_, _tBodyGyroMag_ and _tBodyGyroJerkMag_.
6) Finally, We've got frequency domain signals from some of the available signals by applying a FFT (Fast Fourier Transform). These signals obtained were labeled with ___prefix 'f'___ just like original signals with ___prefix 't'___. These signals are labeled as ___fBodyAcc-XYZ___, ___fBodyGyroMag___ etc.,.
7) These are the signals that we got so far.
   + 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

8) We can esitmate some set of variables from the above signals. ie., We will estimate the following properties on each and every signal that we recoreded so far.

+ ___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()___: Autorregresion 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 to vectors.

9) We can obtain some other vectors by taking the average of signals in a single window sample. These are used on the angle() variable'
   + gravityMean
   + tBodyAccMean
   + tBodyAccJerkMean
   + tBodyGyroMean
   + tBodyGyroJerkMean

Y_Labels(Encoded)
+ In the dataset, Y_labels are represented as numbers from 1 to 6 as their identifiers.

- WALKING as __1__
- WALKING_UPSTAIRS as __2__
- WALKING_DOWNSTAIRS as __3__
- SITTING as __4__
- STANDING as __5__
- LAYING as __6__

Data

All the data is present in 'UCI_HAR_dataset/' folder in present working directory.
- Feature names are present in 'UCI_HAR_dataset/features.txt'
- ___Train Data___
  - 'UCI_HAR_dataset/train/X_train.txt'
  - 'UCI_HAR_dataset/train/subject_train.txt'
  - 'UCI_HAR_dataset/train/y_train.txt'
- ___Test Data___
  - 'UCI_HAR_dataset/test/X_test.txt'
  - 'UCI_HAR_dataset/test/subject_test.txt'
  - 'UCI_HAR_dataset/test/y_test.txt'