

README

The script “run_analysis.R” assumes that it is located inside the folder ‘UCI HAR Dataset’ that is created when the data set is unzipped. Please move the script accordingly.

This script works by first loading the training and testing variables associated with x (measurements of phone sensor signals), y (Measurement labels corresponding to the measurements in x), and subjects (identification numbers corresponding to which subject performed which action). The training and testing sets for each of the three variables are combined.

The labels corresponding to the mean and standard deviation of sensor measurements are then selected out of the accumulated sensor data. The numerical indicators for the movement types are then changed to strings descriptive of the action being performed (e.g., ‘1’ becomes ‘Walking’, ‘2’ becomes ‘Walking upstairs’, and so on).

The subject data, sensor data, and movement type indicators are then combined to make a single data frame. The column names are changed to indicate the Subject, Activity, and measurement types (the latter correspond to the feature labels extracted from the “features.txt” file). Finally, a new data frame is created by grouping the data by subject and activity and calculating the mean of each sensor measurement. The resulting data frame is written to “dataset.txt”.