This is a documentation for the Heterogeneity Dataset for Human Activity Recognition (HHAR) from Smartphones and Smartwatches from the public repository: https://archive.ics.uci.edu/ml/datasets/Heterogeneity+Activity+Recognition+Data+Set or the personal Website: http://cs.au.dk/~allans/heterogenity/.

The Heterogeneity Dataset for Human Activity Recognition from Smartphone and Smartwatches is a dataset devised to benchmark human activity recognition algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc) containing sensor heterogeneities.The files in this archive contain all the samples from the activity recognition experiment. The dataset contains the readings of two motion sensors commonly found in smartphones' recorded while users executed activities scripted in no specific order carrying smartwatches and smartphones.The data is split into 4 files in total divided by device (phone or watch) and sensor (gyroscope and accelerometer). The files for phones are: Phones\_accelerometer.csv, Phones\_gyroscope.csv for the accelerometer and gyroscope respectively, and for the Watch\_accelerometer.csv, Watch\_gyroscope.csv for the accelerometer and gyroscope as well.Activities: ‘Biking’, ‘Sitting’, ‘Standing’, ‘Walking’, ‘Stair Up’ and ‘Stair down’.Sensors: Two embedded sensors, i.e., Accelerometer and Gyroscope sampled at the highest frequency possible by the deviceDevices: 4 smartwatches (2 LG watches, 2 Samsung Galaxy Gears)8 smartphones (2 Samsung Galaxy S3 mini, 2 Samsung Galaxy S3, 2 LG Nexus 4, 2 Samsung Galaxy S+)Recordings: 9 users currently named: a,b,c,d,e,f,g,h,i consistently across all files.The data set is structured in the following way:------------- Accelerometer Samples ------------All the csv files have the same structure of following columns:'Index', 'Arrival\_Time', 'Creation\_Time', 'x', 'y', 'z', 'User', 'Model', 'Device', 'gt'And the columns have the following values:Index: is the row number.Arrival\_Time: The time the measurement arrived to the sensing applicationCreation\_Time The timestamp the OS attaches to the sampleX,y,z The values provided by the sensor for the three axes, X,y,zUser: The user this sample originates from, the users are named a to i.Model: The phone/watch model this sample originates fromDevice: The specific device this sample is from. They are prefixed with the model name and then the number, e.g., nexus4\_1 or nexus4\_2.Gt: The activity the user was performing: bike sit, stand, walk, stairsup, stairsdown and nullEach accelerometer sample if represented as a single row in the file and with all columns having repeated values.

For the semantics of the different timestamp versions, see the publication [1]. Also due to issues with sampling some users have few collected samples for specific activities, e.g., User h and activity sit in the Phones\_accelerometer.csv.------------- Groundtruths --------------------The null class is defined as null in the gt (groundtruth) column, whereas the rest of the classes are called bike sit, stand, walk, stairsup, stairsdown. ------------- Devices and models --------------------------The names and models of the devices used in the HAR data set are:LG-Nexus 4 'nexus4\_1' 'nexus4\_2'Saumsung Galaxy S3 's3\_1' 's3\_2’Samsung Galaxy S3 min: 's3mini\_1' 's3mini\_2'Samsung Galaxy S+: 'samsungold\_1' 'samsungold\_2'The names of the same devices used in the ‘still experiment’ are the following:‘it-116', 'it-133', 'it-108', 'it-103','it-123','3Renault-AH', 'no-name/LG-Nexus4','G-Watch'[1] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen "Smart Devices are Different: Assessing and Mitigating Mobile Sensing Heterogeneities for Activity Recognition" In Proc. 13th ACM Conference on Embedded Networked Sensor Systems (SenSys 2015), Seoul, Korea, 2015. http://dx.doi.org/10.1145/2809695.2809718