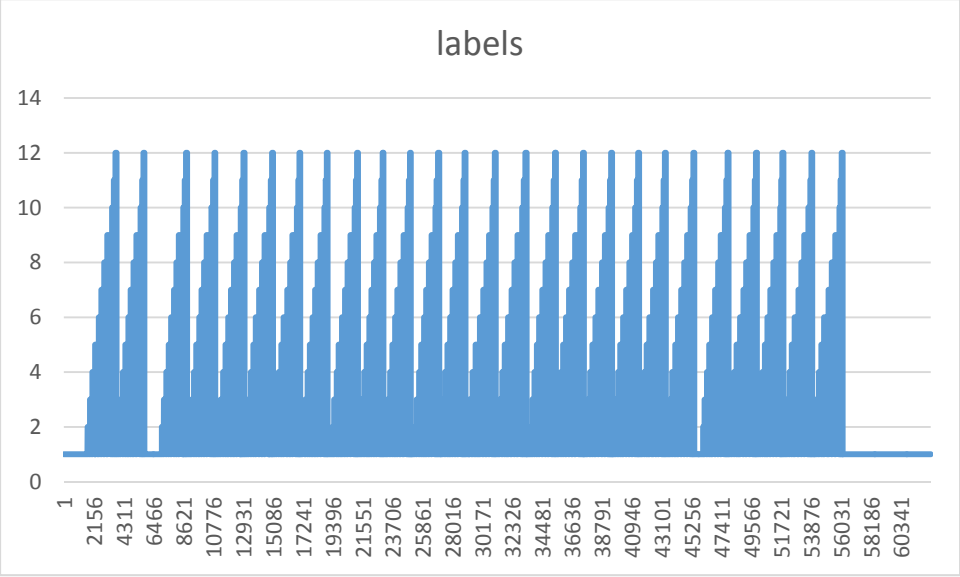
Start with PASDAC

Step1. Install python2 >= 2.7, or python3 >=3.4, numpy and pandas. The easiest way for the majority of users to install pandas is to install it as part of the Anaconda distribution, a cross platform distribution for data analysis and scientific computing. This is the recommended installation method for most users.

Step2. Understand PASDAC code: you can start from exp.py, this script shows the top level framework: PASDAC runs the following 5 functions: import settings, prepare fold data (fold is an important concept when you do cross validation, one fold corresponds to one sample from each activity, and helps you when performing cross validation evaluation), run evaluation and save results.

The program can run under multiple settings, which you can edit through, 'mysettings.py' and 'class_settings.py' files. In 'class_settings.py' you can find all the variables that define the behavior of the experiment. Attributes 'PATH_DATA' sets the folder with the raw data files. We provide 2 datasets for you to play/test with under the sub-folders: '/Data' and '/Data2R'. '/Data2R' contains data for 2 repetitions of each activity (used to rapidly test your code), and '/Data' contains data for 26 repetitions. A Repetition is an important concept and refers to how many times the participant has performed all the activities in a loop.

If you open '/Data/ subject1_gesture_labels.csv' file and plot the labels (which represent categories of activities), the figure shows the repetitions. In the following figure it shows there are 26 repetitions. Getting familiar with fold and repetition will help you a lot to understand PASDAC.

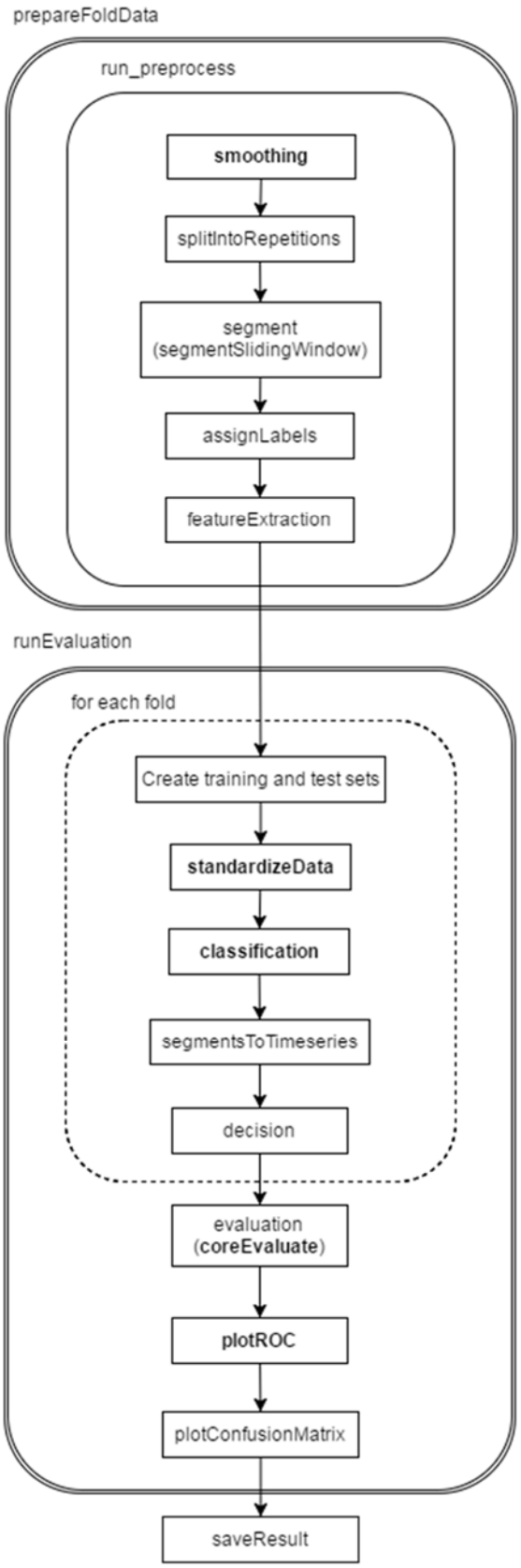


Notice that when you change the dataset, you must also change line 32 in ‘class_settings.py’ file. The number of folds should be the same as the number of repetitions:

Line 32: self.FOLDS = 2 #26 for /Data, 2 for /Data2R

The last command in script ‘mysettings.py’ sets the evaluation method, either pi or pd. Pi will build/evaluate a participant independent and pd will build/evaluate a participant dependent models.

The following flow chart simply shows the framework of PASDAC. You are welcome to modify the components or add new components in it (like feature selection, which is not released in this version). As some parts are held for your assignment, the metrics evaluation is left for you to generate. Our later assignments will focus on the parts with bold words.


prepareFoldData
run_preprocess
smoothing
splitIntoRepetitions
segment
(segmentSlidingWindow)
assignLabels
featureExtraction
runEvaluation
for each fold
Create training and test sets
standardizeData
classification
segmentsToTimeseries
decision
evaluation
(coreEvaluate)
plotROC
plotConfusionMatrix
saveResult