

README for MATLAB function to classify sedentary behaviour with activPAL data (runBehaviourPredictionActivPal)

COPYRIGHT (c) 2021, Roman P Kuster, roman.kuster@alumni.ethz.ch. All rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted for academic purposes provided the conditions listed below are met. Use for commercial purposes is prohibited.

- Redistributions of source code must retain the above copyright notice, this list of conditions, and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions, and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

MATLAB

This function was created with MATLAB v9.6 (2019a), and needs access to the following toolboxes: Signal Processing Toolbox, Statistics and Machine Learning Toolbox, Financial Toolbox, Econometrics Toolbox. If you do not have access to MATLAB and these toolboxes, contact roman.kuster@alumni.ethz.ch

Preparation

Download the MATLAB and Excel files to your local hard disk (e.g. create the folder C:\User\me\documents\myMATLABskripts\activPALSedentaryBehaviourPrediction). Make sure all required files are stored in the same folder. The required files are (in alphabetical order):

- ActivityPredictionAlgorithm.mat (trained algorithm to predict sedentary behaviour)
- extractEpisodesToPredictAP.m (function to extract sitting episodes ≤ 1 minute)
- getFeaturesAndPredictAP.m (function to predict in-/activity in sitting)
- limitDataToProtocolAP.m (function to limit recording, requires optional input argument)
- loadDATXfileAP.m (function to load .datx or .csv raw data files)
- loadPALfileAP.m (function to load the .pal files)
- Protocol_Template.xlsx (Excel file to specify recording time for each subject)
- runBehaviourPredictionActivPal.m (main function calling all other functions)

Locate and organise the folder of your activPAL data to be used. The prediction requires the .pal file (containing the posture events) and the .datx file (containing the raw data). It is highly recommended to make a local copy of the folder on your hard disk. Similarly organise your data as described below for the folder myStudy located in C:\User\me\documents

Main Folder:

- C:\User\me\documents\myStudy

Subject folders in myStudy (one folder per subject):

- C:\User\me\documents\myStudy\AP0001
- C:\User\me\documents\myStudy\AP0002
- C:\User\me\documents\myStudy\AP0004
- ...

Files in each subject folder (only .pal and .datx are required)

- AP0001 02Nov17 12-00am for 7d 0m.pal
- AP0001 02Nov17 12-00am for 7d 0m.datx
- AP0001 02Nov17 12-00am for 7d 0m Events.csv
- ...

If you want to limit the behaviour prediction to a start and stop time for each subject (see optional input arguments how to do so), move the Protocol_Template.xlsx to the main folder (C:\User\me\documents\myStudy), and add the information for each subject. Skip this step if you do not want to limit the behaviour prediction to a start and stop time. In the Protocol_Template.xlsx:

- SubjectID must be equal to the name of the subject folders, e.g. AP0001, AP0002, AP0004
- StartDay contains the start day in format dd.mm.yyyy
- StartTime contains the time on the start day at which the sensor was attached in format HH:MM:SS
- StopDay contains the stop day in format dd.mm.yyyy
- StopTime contains the time on the stop day at which the sensor was removed in format HH:MM:SS

Make sure that there is only one .xlsx document in your main folder containing the case-sensitive keyword *Protocol*, and that this file is not open when executing the function. The SubjectID must be equal to the name of the subject folder, else the protocol data is ignored, and the entire recording is used. Make sure each subject has only one column, and let the cell empty if no date or time is available. If no StartDay is available, the StartTime is ignored. If no StartTime is available, the recording is assumed to have started at 00:00 on the specified StartDay. If no StopDay is available, the StopTime is ignored. If no StopTime is available, the recording is assumed to have ended at 24:00 on the specified StopDay. Do not enter any additional information in the Protocol_Template.xlsx file.

Download the MATLAB toolbox activpal_utils published by R. Broadley. Go to https://github.com/R-Broadley/activpal_utils-matlab/wiki, download the file, and follow the instructions to install. When skipping this step, the behaviour prediction requires the .csv raw data files (need to be created manually on instruction from the function) instead of the .datx files, which takes a significantly longer time.

Sedentary Behaviour Prediction

Open MATLAB and set its current directory to the folder the function files are stored (e.g. C:\User\me\documents\myMATLABskripts\activPALsedentaryBehaviourPrediction). Below are three alternatives on how to set the current directory of MATLAB:

- a) Type the following command in the Command Window (with the location of your folder):
`cd('C:\User\me\documents\myMATLABskripts\activPALsedentaryBehaviourPrediction')`
- b) Go to MATLABs Current Folder Window, and navigate to the folder manually
- c) Copy paste the folder C:\User\me\documents\myMATLABskripts\activPALsedentaryBehaviourPrediction into the folder information of MATLAB (below the main menu).

Execute the function by typing `runBehaviourPredictionActivPal` in your MATLAB Command Window. An explorer window pops up where you have to select the folder in which your data is stored. Select the main folder containing all the subject folders (e.g. C:\User\me\documents\myStudy). The function first checks the input files (.pal and .datx), and will then run through all the subjects. Each step is documented in the Command Window. In case the `activpal_utils` is not able to load a particular .datx file (or `activpal_utils` is not installed), the Command Window will instruct you how to generate the raw data .csv file using `activPAL3` and the processing continues with the next subject. Follow the instructions to generate the .csv, and re-run the function to get the sedentary behaviour prediction for this subject. The processing should take, depending on your computer and the recording duration, around 30 seconds to 5 minutes per subject.

The output data of the function is stored in a .csv file in each subject folder, with `EventsSedentary` appended to the file name (e.g. "AP0001 02Nov17 12-00am for 7d 0m EventsSedentary.csv"). The output file is in the same format as the original `activPAL` event file, but instead of the `activPAL Activity Code` there is a `Behaviour Code` with:

- -1 for sedentary behaviour
- 0 for active sitting
- 1 for standing (equal to `activPAL Activity Code`)
- 2 for stepping (equal to `activPAL Activity Code`)
- 99 for sitting in excluded long bouts (see optional input arguments below)

An additional output file named "Summarised Behaviour Prediction_date-and-time.xls" containing the data of all subjects processed in one run through the function is stored in the main folder (the one you selected in the beginning, e.g. C:\User\me\documents\myStudy). The time the file was created is appended to prevent overwriting previous rounds. You find the data of each subject in an own spreadsheet.

Specify optional input arguments if 1) you want to use time information to limit the behaviour prediction (protocol), 2) you want to exclude long sitting episodes from the behaviour prediction, 3) you want to exclude subjects already having the output file, 4) you want to erase all non-valid minutes in the output file, 5) you want to suppress the generation of a summary file for all subjects, 5) you want to summarize adjacent `activPAL` episodes of the same code into one episode (i.e. stepping). By default, all optional input arguments are deactivated.

1) Limit the behaviour prediction to a particular start and stop time for each subject as stored in the `Protocol_Template.xlsx` file. To run the function with this optional input, copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('UseProtocolFileToLimitTime','Yes')
```

The command window will tell you whether the recording could be limited with the protocol data or not.

2) To exclude long sitting episodes (e.g. those ≥ 24 hours) from the behaviour prediction, copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('ExcludeExtralongEpisodesInHours', 24)
```

Change the number to the desired episode length in hours. This optional input is particularly helpful to speed up the processing if no protocol data is available. Episodes equal or longer as the specified duration will be coded with 99.

3) To process only the subjects not yet having an output file, copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('ExcludeSubjectsHavingOutputCSV', 'Yes')
```

Note that each time you run the function a new summary file will be created in the main folder containing only the processed subject data. Make sure that you do not specify this input argument if you changed the information in the protocol.

4) To delete non-valid minutes (as specified in the Protocol of with the optional input of ExcludeExtralongEpisodesInHours) in the output file, copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('ExcludeNonValidTimeInOutput', 'Yes')
```

5) To suppress the generation of the summary file, copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('SuppressSummaryFile', 'Yes')
```

6) To sum adjacent episodes of the same activPAL code (i.e. stepping) into one episode, copy the following command in the Command Window: `runBehaviourPredictionActivPal('SumAdjacentEpisodes', 'Yes')`

You can combine the optional input arguments in whatever combination, e.g. by copy the following command in the Command Window:

```
runBehaviourPredictionActivPal('ExcludeSubjectsHavingOutputCSV', 'Yes',  
'ExcludeExtralongEpisodesInHours', 12, 'SuppressSummaryFile', 'Yes') to process only the  
subjects not having an output file, exclude sitting episodes  $\geq 12$  hours, and suppress the generation of a summary file.
```

Troubleshooting

The function was tested on various activPAL data sets recorded with different sensor and firmware versions. As it is very unlikely that all errors and bugs were fixed, we ask you to report any errors that occurred to roman.kuster@alumni.ethz.ch. Contact us if you need any assistance, as well as for questions and suggestions to improve.