

Assignment 3

Machine Learning for the Quantified Self

Final assignment covering all material

Ali el Hassouni and Mark Hoogendoorn

Deadline: 20/06/2021, 23:59

1 THE ASSIGNMENT

For this final assignment you can come up with a plan you would like to study yourself. There are four options you can select from with respect to the data you use. Obviously the datasets have to cover human sensory data:

1. focus on a single dataset you did not study before but is available to you (e.g. on the Internet).
2. focus on collecting dataset yourself of a single user.
3. focus on a dataset concerning multiple users you did not study before but is available to you (e.g. on the Internet).
4. focus on collecting a dataset of multiple users (you can make it a joint effort between teams)

To start, make an overview of the data you plan to collect (what sensors, what frequency, what time period, etc.) and also what target you want to set for your machine learning problem (e.g. predict the activity). Discuss this with your TA.

Once discussed and approved, start collecting or downloading the data and follow the full machine learning for the quantified self cycle as we have seen in the book **except for the**

reinforcement learning techniques. Select the appropriate techniques and have good reasons for doing so. Write everything down clearly. Draw inspiration from how things the steps have been described in the book.

2 SUBMISSION REQUIREMENTS

You can work in groups of three students. You should write a report describing the steps you have performed and the results you obtained. The report should follow the Springer Lecture Notes in Computer Science template (you can easily find it if you Google it, there are both LaTeX and Word templates) and should be at most 8 pages. You can submit via Canvas.

3 PRESENTATION

For the final lecture a number of groups that have done outstanding work will be asked to present their work. For those groups, discuss the setting you have selected and present highlights of the results.

4 GRADING

The following criteria are used for grading:

- dataset description (5%)
- techniques selected (25%)
- quality of evaluation (20%)
- rationale provided (25%)
- overall writing style (15%)
- creativity (10%)