Data overview, participant characteristics, feature selection, and AQ1

Author: Nele Albers **Date**: December 2024

This file is meant to guide you through reproducing our Python-based analyses for step 5 from our pipeline and our results reported in the "Results"-section.

Types of analyses

Data overview and participant characteristics

Refer to the file "participant_characteristics_and_data_overview.ipynb" to reproduce:

- the mean effort per activity (cluster) from the "Step 5: Training the model"-section and Table S10 in the Appendix,
- the mean effort per preparatory activity cluster and combination of values for the three selected user-inquired features from Figure S3 in the Appendix,
- the number of samples per activity (cluster) and combination of values for the three selected user-inquired features from Figure S4 in the Appendix,
- the dropout response per session from the "Step 5: Training the model"-section, and
- the participant characteristics (e.g., age, gender, smoking frequency) from Table S8 in the Appendix.

Feature selection

Refer to the file "feature_selection.ipynb" to reproduce our selection of three user-inquired state features.

Analysis for AQ1

Refer to the file "aq1_building_expert_competencies.ipynb" to reproduce:

- Figure 5.4 from the chapter,
- the Cohen's h values from our results for AQ1,
- Figure S5 from the Appendix, and
- the examples of optimal activities in the eight possible starting states from our setup for AQ1.

The computations in the file "aq1_building_expert_competencies.ipynb" rely on reward functions, transition functions, value functions, and optimal policies that we have previously computed via the file "compute_policies.py" and that are stored in the folder "Intermediate_Results."

Steps to reproduce analyses

The reproduction of our code is based on Docker and Jupyter Notebook. Take the following steps:

1. Make sure that you have Docker installed. You can check whether you do by running `docker -v`.
2. Now choose from the following two options:
   - In the directory of this README-file, build the Docker image via `docker build . -t gbna4/usefulness2024_python`.

- Pull the Docker image from Dockerhub via `docker pull gbna4/usefulness2024_python` .

3. Run the Docker container via `docker run -p 8888:8888 -e JUPYTER_ENABLE_LAB=yes -v <this_working_directory>:/home/jovyan/work gbna4/usefulness2024_python` , where `<this_working_directory>` is the path to the directory that this README-file is in.
4. Go to one of the links presented in the terminal upon running the Docker container to access Jupyter Notebook.
5. Open the "work"-folder in Jupyter Notebook.
6. Open one of the notebooks to reproduce the corresponding analyses.

Explanation of files and folders

This directory contains the following files and folders:

- Data: contains the RL samples and contributions of the preparatory activities to the expert competencies. To reproduce the former, run the file "merge_data.py." To reproduce the latter, refer to the folder "Step_4_Creating_The_Model."
- Figures: contains Figure 5.4 from the chapter, and Figure S3, Figure S4, and Figure S5 from the Appendix as created by our code.
- Intermediate_Results: contains previously computed reward functions, transition functions, value functions, and optimal policies, obtained by running the file "compute_policies.py."
- aq1_building_expert_competencies.ipynb: notebook to reproduce our analysis for AQ1.
- compute_dynamics_feat_sel.py: functions for computing dynamics and performing feature selection.
- compute_policies.py: code to compute reward functions, transition functions, value functions, and optimal policies for our analysis for AQ1.
- Dockerfile: Dockerfile to build the Docker image yourself.
- feture_selection.ipynb: notebook to reproduce our selection of three user-inquired state features.
- optimal_policy_computations.py: functions to compute different kinds of optimal policies for our analysis for AQ1.
- participant_characteristics_and_data_overview.ipynb: notebook to reproduce our data overview and participant characteristics.
- README.md/README.pdf: this README-file.