The impact of human feedback in a chatbot-based smoking cessation intervention: An empirical study into psychological, economic and ethical factors

Author: Nele Albers

Date: January 2025

This Readme-file is meant to give you an overview of the data and analysis code underlying the chapter "The impact of human feedback in a chatbot-based smoking cessation intervention: An empirical study into psychological, economic and ethical factors" of the PhD thesis by Nele Albers.

Authored by Nele Albers, Francisco S. Melo, Mark A. Neerincx, Olya Kudina, and Willem-Paul Brinkman.

Types of information

Below we describe the types of information we provide in this repository.

Data

The Data-folder contains all data underlying our analyses. Refer to the corresponding Readme-file for more information.

Preprocessing

We provide files to reproduce our preprocessing steps:

- The file merge_data.py shows how we created non-abstracted transition samples for our analysis for RQ1 as well as our state feature selection.
- The file create_abstracted_rl_samples.py shows how we created abstracted state feature values for the transition samples.
- The file preprocess_data_for_study_comparison.py shows how we combined the data from our study with the data from two previous studies to report the mean effort spent in the "Discussion"-section.

For each preprocessing file, there is a corresponding .txt-file with a log of running the file.

To run the file "preprocess_data_for_study_comparison.py", you first need to:

- download the data from the first study from this DOI: https://doi.org/10.4121/21553055. Specifically, download the file "Data/conversational_sessions_anonym.xlsx" and save it as "Data/Data_Previous_Studies/_data_study_0_conversational_sessions_anonym.xlsx."
- download the data from the second study from this DOI: https://doi.org/10.4121/9c4d9c35-3330-4536-ab8d-d5bb237c277d. Specifically, you need to download the file "Step_5_Training_The_Model_Results/Data/sessionsdata_anonym.csv" and save it as "Data/Data_Previous_Studies/_data_study_1_sessionsdata_anonym.csv."

RQ1

Navigate to the folder "RQ1" and follow the instructions in the Readme-file therein to reproduce our analyses for RQ1.

Data overview, RQ2, and RQ3

Navigate to the folder "DataOverview_RQ2_RQ3" and follow the instructions in the Readme-file therein to reproduce our data overview (e.g., mean effort spent on activities, participant characteristics) as well as the analyses for RQ2 and RQ3.

Explanation of files and folders

This directory contains the following files and folders:

- Data: Anonymized data.
- DataOverview_RQ2_RQ3: To reproduce our data overview, analysis for RQ2, and analysis for RQ3.
- RQ1: To reproduce our analysis for RQ1.
- all_states: All states, including those of people with no state data from session 2.
- all_states_sessions: Sessions corresponding to the states in all_states.
- create_abstracted_rl_samples.py: To create abstracted transition samples.
- create_abstracted_rl_samples_output.txt: Output of the file above.
- map_activities_between_studies.xlsx: To map activity indices between the three studies whose effort scores we compare.
- merge_data.py: To create transition samples.
- merge_data_output.txt: Output of the file above.
- preprocess_data_for_study_comparison.py: To combine the data from the three studies whose effort scores we compare.
- preprocess_data_for_study_comparison_output.txt: Output of the file above.
- README.md/README.pdf: This Readme-file.