

Setting Physical Activity Goals with a Virtual Coach: Vicarious Experiences, Personalization and Acceptance

Prediction Model

Author: Nele Albers **Date:** September 2022

This file is meant to guide you through the steps needed to reproduce our steps for obtaining the prediction model reported in Table 1 in the paper.

Authored by Nele Albers, Beyza Hizli, Bouke L. Scheltinga, Eline Meijer, and Willem-Paul Brinkman.

Requirements

You need to have Docker installed.

Steps to Reproduce Analysis

The main file to reproduce Table 1 is the file "model.Rmd." The reproduction of our results is based on Docker and R Studio. Take the following steps:

1. Make sure you have Docker installed. You can check if you do by running `docker -v`.
2. Navigate to the folder this README-file is in.
3. Now you have 2 options:
 - Build the Docker image via `docker build . -t gbna4/goals_model_r`, or
 - Pull the Docker image from Dockerhub via `docker pull gbna4/goals_model_r`.
4. Run the Docker container via `docker run -d -p 8787:8787 -v <path_to_this_directory>:/home/rstudio/analysis -e PASSWORD=<some_password> gbna4/goals_model_r`.
5. Go to localhost:8787.
6. Login with username 'rstudio' and the password chosen in step 4.
7. Navigate to the "analysis"-folder in R Studio. If the "analysis"-folder is empty, check whether you have used the correct path in the command for running the Docker container. Also, sharing the folder content may not work if you use a drive other than the C drive.
8. Now you can run the analysis using the "model.Rmd"-workbook. To knit a pdf-file from the Rmd-file, you can use the "Knit"-button in R Studio.

Moreover, the files "preprocess_for_model.ipynb" and "collaborative_filtering.ipynb" contain some processing of data that is then used in the file "model.Rmd."

Specifically:

- The file "collaborative_filtering.ipynb" is used to perform collaborative filtering to fill in similarity ratings from people who rated the examples, as not every person rated every example. This is needed to obtain clusters of examples based on the similarity ratings. Refer to the file "model.Rmd" to see why we compute such clusters.

- The file "preprocess_for_model.ipynb" preprocesses the data to be fed into the model. This includes computing people's distances based on user variables.

You can reproduce our results from these two files using Docker and Jupyter Notebook. Take these 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 -f Dockerfile-python . -t gbna4/goals_model_python`.
 - In the directory of this README-file, pull the Docker image from Dockerhub via `docker pull gbna4/goals_model_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/goals_model_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. If the "work"-folder is empty, check whether you have used the correct path in the command for running the Docker container. Also, sharing the folder content may not work if you use a drive other than the C drive.
6. Now you can run code in the two workbooks mentioned above.

Explanation of Files

This directory contains the following files and folders:

- Data: Folder to store the pre-processed data needed for the computations,
- Figures: Contains a figure showing the similarity rating-based clusters we computed,
- centers.csv and clusters.csv: Output of the similarity rating-based clustering conducted in the file "model.Rmd,"
- collaborative_filtering.ipynb: Collaborative filtering to fill in similarity ratings from people who rated the examples,
- Dockerfile: Dockerfile for R,
- Dockerfile-python: Dockerfile for Python,
- model.pdf: Output of the steps to obtain the prediction model reported in Table 1 as run by us,
- model.Rmd: Code to reproduce our steps to obtain the prediction model reported in Table 1,
- motivation_final.csv: Output of the file "preprocess_for_model.ipynb" that processes the data to be fed into the prediction model,
- preprocess_for_model.ipynb: Workbook to process the data to be fed into the prediction model,
- README.md/README.pdf: This README-file,
- references.bib: References used in the .Rmd-file, and
- similarity_predictions.csv: Output of the collaborative filtering to be used in the file "preprocess_for_model.ipynb."