**Instructions to use the algorithm**

***Data preparation***

A predictive model based on sociodemographic and clinical data (6 predictors) is available for predicting treatment response. To use this model, a dataset containing the appropriate set of predictors must be created. The dataset can be in .sav, .csv, or .xlsx format. For reference, a simulated dataset (“simulated\_data.csv”) demonstrating the required structure is available on GitHub (<https://github.com/beccadelf/Exec_functioning_treatment_response>).

The variables will need to be presented with the following name:

* is\_woman (0 = male, 1 = female)
* Age
* Abschluss\_Gymnasium (0 = other, 1 = high education)
* FAS (Fear of spiders questionnaire, total score)
* BDI\_II (Beck depression inventory II, total score)
* STAI\_T (State-trait anxiety inventory, trait subscale)

Before using the model, the dataset needs to undergo the following preprocessing steps: first, missing values need to be recoded as 77777 and 99999 for dichotomous and dimensional variables respectively. Second, they need to be imputed using e.g. the mode for dichotomous, and an iterative procedure (e.g., based on Bayesian Ridge Regression) for dimensional variables. Third, dimensional variables need to be standardized, while dichotomous variables should be coded as 0 and 1.

Steps 2-3 are handled in the provided Python script (see below) and recoding of missing values has already been performed in the simulated data.

***Running the Python script***

An example Python script is provided on GitHub to fully automate the preprocessing and prediction. You’ll also find the simulated dataset and labels for testing.

*Important:* Ensure that the lib/ folder containing custom Python functions is in the same directory as the script.

1. *Recreate the Python environment:*

For conda users:

* + Download the environment.yml file from GitHub
  + Run the following command in your terminal (adjust path as needed):

conda env create -f /path/to/environment.yml

* + Run the following command in your terminal (adjust path as needed):
  + Activate the environment:

conda activate treatment\_prediction\_env

For pip users:

* + Download the requirements.txt file from GitHub
  + In your terminal, navigate to the directory containing requirements.txt
  + (Optional but recommended) Create a virtual environment:

python -m venv treatment\_prediction\_env

* + Activate the virtual environment:

On Windows: treatment\_prediction\_env\Scripts\activate

On macOS/Linux: source treatment\_prediction\_env/bin/activate

* + Install the required dependencies:

pip install -r requirements.txt

1. *Run the script:*
   1. From the command line:
   * Open a terminal (e.g., Powershell on Windows, Terminal on macOS/Linux)
   * Run:

python your\_script.py \

--PATH\_INPUT\_DATA path\_to\_your\_data\_folder \

--OUTPUT\_PATH path\_for\_results \

--MODE inference\_only \

--X\_FILE simulated\_data.csv \

--Y\_FILE simulated\_labels.csv

* 1. From an IDE (e.g., Spyder):
* Open your preferred IDE in the treatment\_prediction\_env environment (see step 1)
* Set the script arguments as follows:

‘--PATH\_INPUT\_DATA’, ‘path\_to\_your\_features\_&\_labels’,

'--OUTPUT\_PATH', ‘path\_for\_results’,

'--MODE', ‘inference\_only’,

'--X\_FILE', ‘features\_filename’,

'--Y\_FILE', ‘labels\_filename’