ZuCo – ET DataLoader for Python

A small guide about how to load the ZuCo Eye-Tracking data fast and conveniently into Python (according to the user's preferences). This will only load the ET features (we will soon extend this to also loading the EEG features).

First, the user needs to specifiy the sub-directory (i.e., folder) where the .mat files are stored in the arguments of the "get_matfiles" function in utils_ZuCo.py. Python will then automatically find the path with the corresponding sub-directory (e.g., "results_zuco").

Second, for object instantiation the user needs to specify the task folder (e.g., 'task1') where all files for the respective task are stored (this must be a folder in the sub-directory).

All incomplete rows (incomplete data due to technical issues specified in the README_ZuCo.pdf) are automatically skipped for the respective subjects.

Here is an example:

```
from utils_ZuCo import *

# instantiate data transformer object for task 1 (corresponds to the folder where results for task 1 in specified subdir are stored) on sentence level with min-max scaling

datatransform_t1 = DataTransformer('task1', level='sentence', scaling='min-max', fillna='zeros')

# NOTE: data for each sbj will be stored in a pd.DataFrame (i.e., this is a list of 12 pd.DataFrames)

sbjs_t1 = [datatransform_t1(i) for i in range(12)]

# show the first couple of rows for subject 1 for task 1

sbjs_t1[0].head()

# convert DataFrame to .csv file and save it to path

sbjs_t1[0].to_csv(path)
```

DataTransformer can be instantiated with the following options (see __init__ for more information):

- Level options: 'word' or 'sentence'
- Scaling options: 'min-max', 'mean-norm', 'standard', 'raw' (no normalization)
- Fill NaNs / Handle missing values: 'zeros', 'mean' (per feature), 'min' (per feature)

Libraries that need to be installed (see utils_ZuCo.py for more information):

- NumPy
- Pandas
- SciPy