Under development

Phenylketonuria landscape analysis

Pipeline:

1. Apply Gaussian model and extract features
2. Perform QC (rmse - var)
3. Detect non-responders
4. Subpopulation detection

Run pipeline

install packages

```
 numpy
pandas
scipy
matplotlib
```

To run the pipeline use the script `gm.py`, which requres experimant files as `*.xlsm`. Example 1) run the pipeline for one experimant

```
python gm.py --input_dir Data/experiments_all/20240305_EXP13.xlsm
```

Example 2) run the pipeline for all experimants in a directory (*e.g.* `Data/experiments_all`)

```
python gm.py --input_dir Data/experiments_all
```

By default, the directory `gm_output/features/` will be created and the resulting feature set will be stored as a *.csv* file. To plot and save the landscapes, you need to add the option `--save_plot`. This will create the directory `gm_output/images/` and save all the landscapes there. It is worth noting that using `--save_plot` may significantly increase the running time.

Other code options:

- `--save_image_dir` : Directory to save images
- `--save_feature_dir` : Directory to save feature csv file
- `--tag` : Name tag for the feature *.csv* file
- `--qc_thr_rmse` : QC threshold for RMSE; `default=[0.2, 0.25]`
- `--qc_thr_n_peaks` : QC threshold for N Peacks; `default=[5, 8]`
- `--qc_thr_variation` : QC threshold for Variations; `default=[0.1, 0.25]`