

This project is an improvement on the HW3 recommendation system.
It incorporates a hybrid model with collaborative filtering and gradient boosting.

To run, put everything in a directory together
(comp.py, train_review.json, user.json, business.json, test_review.json)
and run comp.py via command line. Default values will take care of the params
OR manually set input files with argparse:
comp.py --train_file <file> --user_file <file> --bus_file <file> --test_file <file>
--output_file <file>

Methods for improving original system:

- 1) transform item/user based system into CF of either type (very similar to original assignment)
- 2) add XGBoostRegressor for model-based, and use numpy for preprocessing
- 3) calculate BOTH model predictions
- 4) normalize each prediction by maximum of neighbors, then add for final prediction
- 5) run through manual filter for tweaking rmse125
- 6) only use top 40 for CF neighbor values
- 7) tune hyperparameters
- 8) manually round extreme values similarly to in original system (improves rmse125)