

Intermediate Presentation Group 17 Fabian Hartmann Sebastian Hepp Yifan Mayr Matthias Moik

Current standings

Available resources:

- python files, config-files, datasets and output files are available
- versions of used packages, used CPU, GPU and size of RAM

Reproduced results (currently only on MovieLens 1M):

- reproduced recall off by -0.32%
- reproduced NDCG off by -0.11%
- difference in epochs needed when reproducing: -7 epochs or -5.15%
- possible cause: different random seed for random sampling

Current standings

Experiment design flaw:

- No validation set is used for epoch selection and determining early stoppage, instead done on test set which induces bias to the final evaluation.
- Results paper: recall: 0.2778, NDCG: 0.264
- Our results: recall: 0.134, NDCG: 0.251
- Reason: Reduced train-set and no bias

Future work

Still to do:

- Run tests on the Amazon dataset
- Reproduce results showing the impact of different parameters

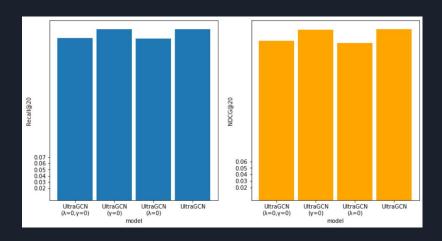
Difficulties:

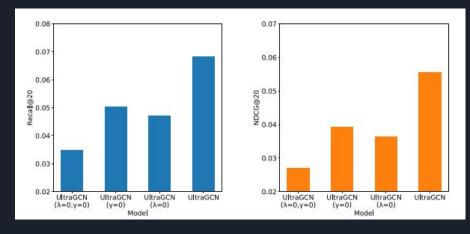
 Amazon dataset was used for all the tests, which needs a lot of RAM (more than 32GB)

Future work

Influence of parameters:

• Parameter influence on MovieLens 1M (ours) vs. Amazon dataset (paper)





Thanks for your attention!