

Title: Application of Graph Neural Networks to Mitigate Popularity Bias in Content Recommendations

Abstract: In this project, we explore the use of Graph Neural Networks (GNNs) to address the issue of popularity bias in content recommendations. Popularity bias issue leads to the over-representation of popular items, thereby reducing the visibility of less popular but potentially relevant content. Such skewed item exposure not only undermines user satisfaction by limiting diversity and coverage, but also hinders the discovery of niche items that could better align with individual user preferences. By considering the ability of GNNs to capture complex user-item interaction patterns, we aim to develop a graph-based approach that provides fair exposure to all items regardless of their popularity. We formulate this as an optimization problem balancing traditional accuracy metrics with fairness- and diversity-oriented objectives, ensuring that improvements in mitigating popularity bias do not come at the cost of overall recommendation quality.

Datasets: The computational experiment will be conducted on three datasets to ensure robustness, scalability and reproducibility. All datasets are open-source and come preprocessed or easily configurable for ready-to-model usage.

1. **MovieLens 32M:** a widely used movie-rating dataset with approximately 32 million user-item interactions. This is a typical long-tail dataset that can provide a strong baseline for our approach. Available at: GroupLens.
2. **OTTO:** a dataset originating from a Kaggle challenge, it represents user sessions and interactions with online retail products, including views, cart additions, and purchases. Available at: Kaggle.
3. **Amazon Books:** a dataset that provides detailed information on books, including titles, authors, genres, ratings, reviews and purchase statistics. It is frequently used in analyzing the phenomenon of popularity bias. Available at: Amazon Review Data.
4. **Last.FM:** this dataset contains social networking, tagging, and music artist listening information from Last.FM online music system. Available at: GroupLens.

References: Papers with a fast intro and the basic solution to compare.

1. **The formulation of the problem:** «A Survey on Popularity Bias in Recommender Systems» [1]
2. **A baseline and new results:** «LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation» [2].
3. **Some fast introduction:** «XSimGCL: Towards Extremely Simple Graph Contrastive Learning for Recommendation» [3].

Basic solution: LightGCN model that demonstrates excellent results in the accuracy of recommendations, but poorly takes into account the popularity bias. The code of the authors' solution is available at the link: <https://github.com/kuandeng/LightGCN>.

Authors: Kirill Shevchenko, Nikita Zelinskiy

References

- [1] Anastasiia Klimashevskaya, Dietmar Jannach, Mehdi Elahi, and Christoph Trattner. A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction*, 34(5):1777–1834, July 2024.
- [2] Tinglin Huang, Yuxiao Dong, Ming Ding, Zhen Yang, Wenzheng Feng, Xinyu Wang, and Jie Tang. Mixgcf: An improved training method for graph neural network-based recommender systems. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery; Data Mining*, KDD '21. ACM, August 2021.
- [3] Junliang Yu, Xin Xia, Tong Chen, Lizhen Cui, Nguyen Quoc Viet Hung, and Hongzhi Yin. Xsimgcl: Towards extremely simple graph contrastive learning for recommendation. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–14, 2023.