Popularity Debiasing While Preserving Performance in Collaborative Filtering

Abstract

Popularity bias, an important, long-standing issue in recommendation systems, refers to the tendency to overly recommend popular items, neglecting diversity. Various methods have been proposed to address popularity bias. One effective approach is the compensation model, which aims to balance recommendations by considering factors beyond popularity. The compensation model defines a framework where recommendations are adjusted based on additional criteria such as relevance, diversity, or novelty. However, while the compensation model expands coverage, it often sacrifices metrics like precision and recall, which measure the effectiveness of recommendations. This paper suggests that items with similar popularity should receive varied compensation scores based on their unique attributes, thereby enhancing recommendation diversity. Moreover, this research introduces coherence as a potential substitute for popularity within the compensation model. By replacing popularity with coherence, the study demonstrates that recommendation performance can be maintained while mitigating the adverse effects of popularity bias.

1 Introduction

Popularity bias poses a significant challenge in recommendation systems, where the tendency to prioritize popular items can limit the system's ability to provide diverse and personalized recommendations. This bias arises from algorithms favoring items with higher user engagement metrics, such as views or purchases, without adequately considering the individual preferences of users. As a result, users may be presented with recommendations that overlook less mainstream yet potentially relevant options, leading to reduced user satisfaction and engagement. Addressing popularity bias is crucial for enhancing the effectiveness and user experience of recommendation systems. This paper explores strategies to mitigate popularity bias, focusing particularly on the compensation model as a potential solution. By examining how alternative metrics, such as coherence, can supplement or replace popularity in recommendation algorithms, this study aims to demonstrate new metric to maintain recommendation performance while fostering diversity and relevance in recommendations.

2 Method

In this section, we explain the existing method and introduce a new metric called coherence. Subsequently, we replace the existing metric, popularity, with coherence and compare their performance.

2.1 User Based Collaborative Filtering

User-Based Collaborative Filtering (CF) is a recommendation algorithm that leverages user similarity to provide personalized recommendations. The core idea is that users with similar preferences will have similar tastes in items. The predicted rating $\hat{r}_{ui}$ is given by:

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_{v \in N(u)} sim(u,v)(r_{vi} - \bar{r}_v)}{\sum_{v \in N(u)} |sim(u,v)|}$$

where $\bar{r}_v$ is the rating of neighbor v, r_{vi} is the rating of neighbor v for item i and N(u) is the set of neighbors to user u.

2.2 Existing Method

One of the two primary models that assign reward scores initially aimed to eliminate popularity bias by multiplying the actual score by a factor that decreases according to popularity and then taking a weighted sum with the actual score.

$$r_{ui} = \alpha \cdot \hat{r}_{ui} + (1-\alpha) \cdot \frac{\hat{r}_{ui}}{\log(pop(i))}$$

where:

- α is a weighting factor that balances the predicted rating and the popularity-adjusted term.
- $\hat{r}_{ui}$ is the predicted rating for user u and item i.
- $pop(i)$ refers to the popularity of the item i, often measured by the number of ratings or interactions.

Although subsequent research utilized more complex scoring methods,

$$C_{u,i} = \frac{1}{pop(i)} \cdot (\beta \cdot \hat{R}_{u,i})$$

and the adjusted predicted rating $\hat{R}^*_{u,i}$ is given by:

$$\hat{R}^*_{u,i} = \hat{R}_{u,i} + \alpha \cdot C_{u,i} \cdot \frac{\|C_u\|}{\|R_u\|}$$

, where α and β are constant, the principle remained the same.

The common point in both studies is that they applied debiasing weights based on popularity regardless of the item. This approach ultimately led to a small enhancement in the coverage of the recommendation system. To address this issue, we introduce a new metric called Coherence in the next section.

2.3 New Metric for Popularity Debiasing

To reduce popularity bias, I suggest new definition coherence that represents the user's similarity that used specific item. The coherence is given by:

$$C_i = median\left(\left\{\frac{\mathbf{u}_j \cdot \mathbf{u}_k}{\|\mathbf{u}_j\|\|\mathbf{u}_k\|} \mid j < k\right\}\right) \quad (1)$$

where:

- $\mathbf{u}_j$ and $\mathbf{u}_k$ are the rating vectors of users j and k who rated movie i.
- $\|\mathbf{u}\|$ denotes the Euclidean norm of vector $\mathbf{u}$.
- $\mathbf{u}_j \cdot \mathbf{u}_k$ denotes the dot product of vectors $\mathbf{u}_j$ and $\mathbf{u}_k$.
- $median(\cdot)$ denotes the median of the set.

It can achieve higher coverage by assigning different reward scores to items with the same popularity. By mimicking the form of $\frac{1}{\log(pop)}$, it is possible to gain accuracy along with the effect of removing popularity bias.

The final score can be calculated by:

$$r_{ui} = \alpha \cdot \hat{r}_{ui} + (1-\alpha) \cdot \hat{r}_{ui} \cdot C_i \quad (2)$$

where α is constant, so we can control the compensation score.

3 Experimental Results

3.1 Experiment Setup

The dataset for the experiment is Movielens-1M and evaluation metrics are precision, recall, coverage and popularity.

Precision

Precision measures the proportion of recommended items that are relevant. A high precision value indicates that the system accurately recommends items that the user is likely to prefer.

$$Precision = \frac{Number of relevant items recommended}{Number of recommended items}$$

Recall

Recall measures the proportion of relevant items that have been recommended. A high recall value indicates that the system does not miss many items that the user would prefer.

$$Recall = \frac{Number of relevant items recommended}{Number of relevant items}$$

Coverage

Coverage measures the proportion of the total items that have been recommended. A high coverage value indicates that the system recommends a diverse set of items, avoiding bias towards a small subset of items.

$$Coverage = \frac{Number of items recommended}{Total number of items}$$

Popularity

Popularity measures the average popularity of the recommended items. Popularity is typically measured by the total number of ratings an item has received. A low popularity value indicates that the system is recommending less popular items, showcasing its ability to explore and recommend a wider range of items.

$$Popularity = \frac{\sum_{i \in Recommended items} Popularity(i)}{Number of recommended items}$$

Model	Precision	Recall	Coverage	Popularity
Baseline	0.1966	0.1215	0.4198	7.1950
Model using popularity	0.1968	0.1217	0.4350	7.1752
Model using coherence	0.1937	0.1198	0.4674	0.7111

Table 1: Evaluation metrics for different models.

Conclusion

The evaluation was conducted by replacing the traditional method of dividing by popularity with multiplying by coherence. The table above illustrates how incorporating coherence can increase coverage while decreasing the influence of popularity bias (as seen in the Popularity metric), while maintaining or improving Precision and Recall compared to the baseline and popularity-based models. The future work will be applying coherence to complicated model and various of evaluation metrics.

Reference

- F. Maxwell Harper and Joseph A. Konstan. "The MovieLens Datasets: History and Context." *ACM Transactions on Interactive Intelligent Systems (TiiS)*, Volume 5, Issue 4, Article 19, December 2015. ACM, New York, NY, USA. DOI: 10.1145/2827872. Available online: https://dl.acm.org/doi/10.1145/2827872

- Ziwei Zhu, Yun He, Xing Zhao, James Caverlee. "Evolution of Popularity Bias: Empirical Study and Debiasing." 2022. ArXiv eprints, arXiv:2207.03372. https://arxiv.org/abs/2207.03372

- Ziwei Zhu, Yun He, Xing Zhao, Yin Zhang, Jianling Wang, James Caverlee. "Popularity-Opportunity Bias in Collaborative Filtering." In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM '21)*, Virtual Event, Israel, 2021. ACM, New York, NY, USA. DOI: 10.1145/3437963.3441820. Available online: https://doi.org/10.1145/3437963.3441820