Improving Recommendation Fairness via Data Augmentation

Project Proposal

Team 29
Kaki Hephzi Sunanda
210150018
h.sunanda@iitg.ac.in

I. Introduction

Collaborative filtering-based recommendation models learn from historical user behavior to provide accurate predictions. However, these systems often suffer from fairness problems. The performance of recommender systems differs between user groups defined by its sensitive attributes. This unfairness can lead to biased recommendations that favor certain demographics.

Many approaches have been proposed to alleviate recommendation fairness by optimizing specific fairness metrics, changing model architectures or changing the distribution of the dataset. However, these methods either require predefined fairness objectives or require significant changes in the existing recommendation models. The inherent imbalance in training data also adds to the unfairness in recommendations. This paper [1] studies on improving the recommendation fairness via data augmentation.

The paper introduces a framework called Fairness-aware Data Augmentation (FDA) to improve fairness in recommendation systems. FDA augments to the imbalanced training data, by generating synthetic (fake) user-item interactions, aiming to achieve a more balanced data. FDA does not need predefined fairness metric and can be integrated into the backbone of various embedding-based recommendation models. FDA generates synthetic data through a bi-level optimization process, to better reflect real-world user-item interaction scenarios. This framework achieves a trade-off between recommendation fairness and accuracy on two real world datasets: MovieLens and LastFM.

II. Objectives

The objective of this project is to use Fairness-aware Data Augmentation (FDA) framework proposed in the paper to further study fairness in recommender systems. The objectives include:

1. Study the framework's effectiveness on non-binary sensitive attributes, such as age or ethnicity, to understand how the data augmentation process impacts fairness across different demographics.
2. Evaluate the effectiveness of the FDA framework on datasets with multiple sensitive attributes to assess its ability to address complex forms of unfairness.
3. Understand implicit unfairness within datasets by analysing the distribution of user-item interactions and asses model's performance across different user groups. Evaluate how the FDA framework can be alleviate this implicit bias and ensure a fairer recommendation.

By achieving these objectives, the project aims to enhance the current understanding of fairness in recommendation systems.

III. Methodology

The methodology for this project involves implementing and extending the Fairness-aware Data Augmentation (FDA) framework to improve fairness in recommender systems. The process is divided into several key steps:

A. *Data Collection and Preprocessing*

1) Dataset: This project will use publicly available datasets, such as MovieLens and LastFM, as used in the original paper. Additionally, other datasets containing diverse sensitive attributes (e.g., age, ethnicity) will be incorporated to assess the proposed framework.

2) Preprocessing: Data preprocessing will involve filtering interaction records and encoding user and item features. The sensitive attributes will be extracted to create user groups for fairness evaluation. The implicit feedback data will be converted into a binary interaction matrix to support the FDA framework.

B. *Implementing the FDA Framework*

The original FDA framework will be implemented using the baseline models: Bayesian Personalized Ranking (BPR) and Graph Convolutional Collaborative Filtering (GCCF). The performance of FDA on these models will be used as a benchmark.

The FDA framework will generate synthetic user-item interactions by employing the hypotheses proposed in the paper. Specifically:

- **Hypothesis 1**: Fake positive interactions will be generated to balance the positive data distribution between user groups.
- **Hypothesis 2**: Fake negative interactions will be generated to balance the negative data distribution between user groups.

Small noise is added to item embeddings to maintain a distribution similar to the original data, which helps preserve the recommendation model's accuracy while enhancing fairness.

C. Evaluating the FDA Framework

The framework will be used to handle multiple sensitive attributes (e.g., age, ethnicity) by modifying the hypotheses to create balanced data for multiple groups. The data augmentation process will be adjusted to ensure that interactions between different groups are present.

The FDA framework's performance will be evaluated on datasets containing more than one sensitive attribute to assess its ability to address complex forms of unfairness.

The distribution of user-item interactions will also be analyzed, and the model's performance across different user groups will be evaluated to identify implicit biases within the datasets. The effectiveness of the existing FDA framework in alleviating these implicit biases will then be assessed.

D. Evaluation Metrics

Fairness Metrics: To assess the fairness of the recommendation models, group fairness metrics such as Demographic Parity (DP) and Equality of Opportunity (EO) will be used. Additionally, individual fairness metrics will be explored to provide more analysis of fairness at the user level.

Accuracy Metrics: To evaluate recommendation accuracy, metrics like Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG) will be used. The trade-off between fairness and accuracy will be examined across different experimental settings.

E. Experimental Setup

Bi-Level Optimization: The framework will be trained using a bi-level optimization process, where the inner optimization generates the fake data for augmentation, and the outer optimization updates the recommendation model parameters.

Hyperparameter Tuning: Parameters such as embedding size, learning rate, and the ratio of fake data (Max_{Mask}) will be fine-tuned for optimal performance. Experiments will be conducted to find the balance between synthetic data quantity and model accuracy.

By following this methodology, the project aims to implement the original FDA framework, explore the objectives, and assess its performance in improving fairness in recommender systems.

IV. Conclusion

Fairness in recommender systems has become an essential issue, as current models often increase bias present in historical user data, leading to different outcomes for different demographic groups. This project aims to build on the Fairness-aware Data Augmentation (FDA) framework proposed in recent research to address these biases. By implementing the original FDA framework and exploring objectives, such as incorporating non-binary sensitive attributes, evaluating on multiple sensitive attributes, and identifying and alleviating implicit unfairness in data and recommendations, this project aims to provide a broader perspective on the recommendation process.

Paper

[1] Chen, Lei, Le Wu, Kun Zhang, Richang Hong, Defu Lian, Zhiqiang Zhang, Jun Zhou, and Meng Wang. "Improving recommendation fairness via data augmentation." In Proceedings of the ACM Web Conference 2023, pp. 1012-1020. 2023.