ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9.
Available from http://www.i6doc.com/en/.

FairBayRank: A Fair Personalized Bayesian Ranker

Armielle Noulapeu Ngaffo, Julien Albert, Benoît Frénay and Gilles Perrouin

University of Namur - NaDI - Faculty of Computer Science - PReCISE
Rue Grandgagnage 21, B-5000 Namur - Belgium

Abstract. Recommender systems are data-driven models that successfully provide users with personalized rankings of items (movies, books...). Meanwhile, for user minority groups, those systems can be unfair in predicting users' expectations due to biased data. Consequently, fairness remains an open challenge in the ranking prediction task. To address this issue, we propose in this paper FairBayRank, a fair Bayesian personalized ranking algorithm that deals with both fairness and ranking performance requirements. FairBayRank evaluation on real-world datasets shows that it efficiently alleviates unfairness issues while ensuring high prediction performances.

1 Introduction

The increasing volume of multimedia content has motivated the development of recommender systems (RS) to align relevant products (shoes, books, movies, etc.) with users' preferences [1]. In the literature, RS are organized in ranking-based and rating-based models. Ranking-based algorithms predict the ranked list of most relevant items for the final user, while rating-based models predict the absolute relevance score of an item and not its rank [1]. Since their main objective is to provide a ranked list of the most relevant items, many recommender models are optimized with ranking algorithms [2]. Those data-oriented systems perform users' taste predictions based on consumers' past behavior and therefore become unfair in case of biased data. Indeed, user minority groups can be provided with unfair predictions due to their under-representation in the whole dataset [3]. In this paper, we present the FairBayRank model which performs a fair personalized Bayesian ranking to alleviate fairness concerns while maintaining a high-ranking prediction quality. Section 2 presents fairness-related work in RS and Section 3 formalises this problem. Section 4 presents our model, and experiments are discussed in Section 5. Section 6 wraps the paper with conclusions and perspectives.

2 Fairness in Recommendation

Fairness in recommender systems (RS) ensures that users' preferences predictions are performed without considering the values of certain attributes. One qualifies those attributes as *protected* or *sensitive*. They include gender, age, race, grade, etc. Fairness in RS is studied from the user side to fairly satisfy users' requirements or the item side to fairly promote items without discrimination [3]. Several approaches [1] addressing fairness are data-oriented, re-ranking-based, or ranking-based. Data-oriented methods require a significant preprocessing of the data load with an important risk of performance decrease during subsequent data re-sampling steps. Re-ranking-based

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

methods fairly adjust recommendation results from models and consequently perform a fairness-unaware prediction. Ranking-based methods are most often implemented by using regularization [4, 5], adversarial learning [2], or reinforcement learning techniques [6]. Regularization-based methods are more flexible and easily extensible [1]. Zhu *et al.* propose a fair ranking-based method implemented through an indirect regularization by isolating sensitive user attributes from latent feature matrices [4]. Wan *et al.* add fairness metrics to the loss function to perform fair recommendations [7]. Thanks to their flexibility [1], ranking-based methods are the most-used fair recommendation techniques. However, state-of-the-art methods focus on tackling fairness issues in rating prediction and tend to ignore the more challenging issues of fairness in ranking [1]. To fill this gap, our FairBayRank approach is a fair Bayesian personalized ranking that *directly performs a fairness-aware ranking*. The two next sections present the problem formulation, Bayesian Personalized Ranking, and our proposal.

3 Problem Formulation and Bayesian Personalized Ranking

We define the set of users $\mathcal{U}$ interacting with a set of items $\mathcal{I}$. The evaluation score x_{ui} expresses the relevance of item i to user u. $\mathcal{I}^+$ is the set of positive items, i.e., the most relevant items to user u. In opposition, $\mathcal{I}^- = \mathcal{I} \setminus \mathcal{I}^+$ is the set of negative items. $\mathcal{D}$ is the dataset of triplets of (u, i, j) where $u \in \mathcal{U}, i \in \mathcal{I}^+, j \in \mathcal{I}^-$. The problem is to determine the optimal criterion that ensures that users will be provided with an accurately ranked list of the most relevant items without discrimination on user groups.

This work uses the Bayesian Personalized Ranking (BPR) model [8] as a baseline that we modify to enforce fairness. Indeed, the baseline BPR is fed by triplets (u, i, j) without considering any sensitive attribute. Consequently, it performs a fairness-unaware ranking prediction across user groups. To remedy this, FairBayRank (see Section 4) is trained with additional parameters referring to the different user groups.

Instead of predicting evaluation scores, BPR predicts the probability $P(i > j; u)$ that a positive item i has a higher rank compared to a negative item j. We determine the Bayesian pairwise ranking optimal criterion by maximizing the posterior function $P(\mathbf{\Theta}|\mathcal{I}^+ > \mathcal{I}^-; \mathcal{U}) \propto \prod_{(u,i,j)\in\mathcal{D}} P(i > j; u)_{\mathbf{\Theta}} * P(\mathbf{\Theta})$ where $\mathbf{\Theta}$ are model parameters; $P(i > j; u)_{\mathbf{\Theta}} \equiv \delta_{\mathbf{\Theta}}(x_{uij})$ with $x_{uij} = x_{ui} - x_{uj}$, and $\delta(\mathbf{x}) = \frac{1}{1+e^{-\mathbf{x}}}$ defined by the logistic sigmoid function. The prior function $P(\mathbf{\Theta})$ is a normal distribution. The log-posterior probability becomes

$$\begin{aligned} \log P(\mathbf{\Theta}|\mathcal{I}^+ > \mathcal{I}^-; \mathcal{U}) &= \sum_{(u,i,j)\in D} \log(P(i > j; u)_{\mathbf{\Theta}}) + \log(P(\mathbf{\Theta})) \\ &= \sum_{(u,i,j)\in\mathcal{D}} \log(\delta_{\mathbf{\Theta}}(x_{uij})) - \lambda_{\mathbf{\Theta}} \left\|\mathbf{\Theta}\right\|^2, \end{aligned} \tag{1}$$

where $\lambda_{\mathbf{\Theta}}$ is a hyperparameter of the model on which the ranking method is applied. The optimal Bayesian ranking is obtained by minimizing the BPR loss function

$$L_{\text{BPR}} = -\sum_{(u,i,j)\in\mathcal{D}} \log(\delta_{\mathbf{\Theta}}(x_{uij})) + \lambda_{\mathbf{\Theta}} \left\|\mathbf{\Theta}\right\|^2. \tag{2}$$

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

4 FairBayRank Recommendation Method

FairBayRank is a ranking-based method that intrinsically tackles fairness concerns during the ranking prediction process. To mitigate unfairness, the FairBayRank loss function is evaluated by extending the Bayesian personalized ranking loss with a customized unfairness term. As a regularization-based fair ranking method, the FairBayRank loss function presents the advantage to be flexible and easily extensible. Figure 1 describes the FairBayRank prediction process. Since the proposed method is interested in addressing user-side fairness concerns, for a user u, items are divided into positive items i and negative items j. Positive items i have greater relevance for user u contrary to negative items j that are lower-relevant for u. For user-side fairness purposes, sensitive attributes are extracted to clearly identify minority and majority user groups.

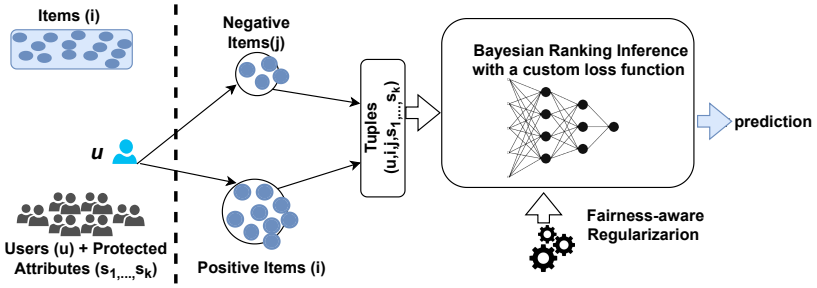

Figure 1: The FairBayRank system model.

Let $\mathcal{G}$ be the subgroup of minority users that are underrepresented, and therefore discriminated based on a sensitive attribute s; and $\neg\mathcal{G}$ the majority user subgroup. $\mathcal{D}'$ is the dataset of tuples of $(u, i, j, g, \neg g)$ where $u \in \mathcal{U}, i \in \mathcal{I}^+, j \in \mathcal{I}^-$, g and $\neg g$ are binary and express the group membership to $\mathcal{G}$ or $\neg\mathcal{G}$. FairBayRank performs a matrix factorization using a fair ranking loss function optimized by stochastic gradient descent (SGD). This matrix factorization (MF) redefines the evaluation score as $\hat{x}_{ui} = \sum_f \mathbf{p_{uf}}.\mathbf{q_{if}}$ where f is the number of latent factors, $\mathbf{p_{uf}}$ is the user latent factor vector, and $\mathbf{q_{if}}$ is the item latent factor vector. The optimal Bayesian ranking is obtained by maximizing the posterior function $P(\mathbf{\Theta}|\mathcal{I}^+ > \mathcal{I}^-; \mathcal{U}, \mathcal{G}, \neg\mathcal{G}) \equiv \prod_{(u,i,j,g,\neg g)\in\mathcal{D}'} P(i > j; u, g, \neg g)_{\mathbf{\Theta}} * P(\mathbf{\Theta})$ where $P(i > j; u, g, \neg g)_{\mathbf{\Theta}} \equiv \delta_{\mathbf{\Theta}}(\hat{x}_{uij})$ with $\hat{x}_{uij} = \hat{x}_{ui} - \hat{x}_{uj}$. L'_{BPR} is computed as follows:

$$L'_{\text{BPR}} = -\sum_{(u,i,j,g,\neg g)\in\mathcal{D}'} \log(\delta_{\mathbf{\Theta}}(\hat{x}_{uij})) + \lambda_{\mathbf{\Theta}} \|\mathbf{\Theta}\|^2 . \quad (3)$$

$\mathbf{\Theta}$ represents MF model parameters and describes the user latent factor vector $\mathbf{p_{uf}}$, the positive item latent factor vector $\mathbf{q_{if}}$ and the negative item latent factor vector $\mathbf{q_{jf}}$ which are updated during the SGD execution until the algorithm converges. To build $L_{\text{FairBayRank}}$, an unfairness loss L_{unfair} is added to penalize the L'_{BPR} loss function when unfairness occurs across different groups. L_{unfair} is defined as the gap between loss averages across user groups $\mathcal{G}$ and $\neg\mathcal{G}$. The idea is to balance model performances across different user groups while optimizing the ranking for each user group as much

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

as possible. L_{unfair} is defined as

$$L_{\text{unfair}} = \sum_{(u,i,j,g,\neg g)\in\mathcal{D}'} \left((E_{\mathcal{G}}[\log(\delta_{\mathbf{\Theta}}(\hat{x}_{uij}))]]\right)^2 - \left(E_{\neg\mathcal{G}}[\log(\delta_{\mathbf{\Theta}}(\hat{x}_{uij}))]]^2\right)^2 \quad (4)$$

and the FairBayRank loss function of the fair ranking system becomes

$$L_{\text{FairBayRank}} = L'_{\text{BPR}} + \gamma * L_{\text{unfair}}, \quad (5)$$

where γ is a regularization constant that should be appropriately set to calibrate the regularization impact. The next section presents our experiments.

5 Experimentations

We conducted our experiments on two real-world datasets: Movielens [1] and LastFM [2]. MovieLens contains one hundred thousand interactions between users and movies resulting in ratings to evaluate the movie relevance per user. The LastFM dataset contains 360k tuples made of users, artists, and songs. From the LastFM dataset, we randomly sample one hundred thousand users' interactions on played artists' songs. The number of plays is recorded as an evaluation score of the relevance of an artist to a user. For each dataset, we identify the gender attribute as sensitive noticing that female users are underrepresented. Females and males also have been separately extracted to appreciate prediction results group-wise compared to the overall performance. For both aforementioned datasets, evaluation scores are scaled from 1 to 5 and a threshold τ has been set to 2. For a user u, evaluation scores strictly upper than τ expressed items considered as positive and then relevant to u. Other items are therefore considered as negative or irrelevant to u. The FairBayRank model is trained with tuples that include group memberships to mitigate unfairness across different user groups during the training phase. For each user, positive items are divided into training and test data. Negative items are all used in the tuple-making process. Indeed, tuples are made from users, positive items, negative items, and group membership (female or male). The customized loss is optimized by using Adam with a learning rate of 0.001 and batch size of 256. Another sensitive user attribute has been used to study how FairBayRank mitigates fairness concerns over more than two groups. Precisely, we extracted four user subgroups based on the profession, and fairness has been evaluated across them.

We use the area under the curve (AUC) to assess the ranking prediction accuracy. For an instance that is a tuple made of a positive item and a negative one, the AUC measures the accuracy of the model to correctly classify a positive item as relevant for the user. In addition, the Normalized Discounted Cumulative Gain (NDCG) and the Mean Average Precision (MAP) assess the ranking prediction quality.

The proposed method aims to ensure a fair ranking prediction quality between minority and majority user groups while maintaining high overall performance. We thus compare FairBayRank to the baseline Bayesian Personalized Ranking (BPR) [8] in different scenarios. In one scenario, BPR is trained with balanced data, and in another scenario, BPR is trained with the original unbalanced data.

[1] https://www.tensorflow.org/datasets/catalog/movieLens
[2] https://www.upf.edu/web/mtg/lastfm360k

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

Figure 2 shows FairBayRank performances through AUC, NDCG, and MAP metrics. Subfigures 2a,2b,2c show that despite unbalanced data, FairBayRank better alleviates unfairness across female and male user groups of Movielens while ensuring high prediction performances. Fairness mitigation across different user groups is moreover observed with the LastFM dataset (see subfigures 2j,2k,2l).

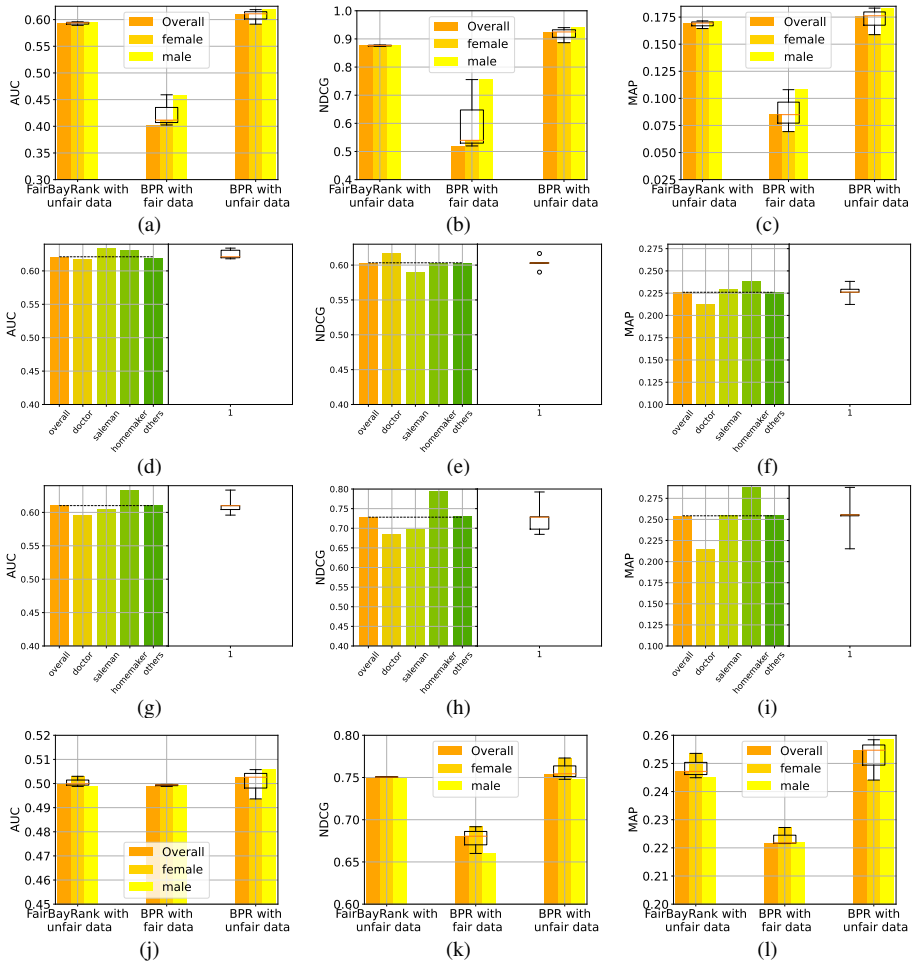

Figure 2: FairBayRank testing performances on MovieLens and LastFM datasets. Columns correspond to the AUC, NDCG, and MAP, respectively. The first and last rows show results with two user groups for MovieLens and LastFm. The second and third rows compare results with four user groups for MovieLens using FairBayRank and BPR.

It can be observed that data balancing as a fairness handling technique is detri-

ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9. Available from http://www.i6doc.com/en/.

mental because of the loss of prediction performance. This is due to the deletion of data required to balance different user groups. A reasonable performance loss is observed on FairBayRank which tries to fulfill both fairness and prediction performance constraints. Subfigures 2d,2e,2f,2g,2h,2i compared vertically side-by-side, show that FairBayRank efficiently mitigates unfairness across more than two user groups. It is a particularly challenging use case since fairness concerns are most often tackled over only two groups [1]. Additionally, when FairBayRank cannot provide equal performance for all groups, it tends to privilege the minority group, as we can observe in Figures 2j and 2l.

6 Conclusion and Perspectives

We proposed FairBayRank, a fair and personalized ranking approach to address fairness concerns in recommendation applications. We improved a Bayesian Pairwise ranking model by regularizing the baseline loss function with a specific fairness-aware term. The regularization term penalizes the loss in case of unfairness occurring across different user groups. The proposed FairBayRank model has been trained with tuples that exceptionally consider user group membership in order to effectively handle fairness concerns. The fairness constraint across groups relies on one attribute (gender or occupation) at once. We showed that FairBayRank significantly alleviates fairness concerns across different user groups while ensuring reasonable model performances. Moreover, the proposed method efficiently tackles fairness even across more than two user groups while maintaining high overall prediction performances. We plan to extend our work on multi-attribute fairness constraints to model intricate unfairness cases underlying data.

References

[1] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. A survey on the fairness of recommender systems. *ACM Trans. Inf. Syst.*, 41(3), 2023.

[2] Xiangnan He, Zhankui He, Xiaoyu Du, and Tat-Seng Chua. Adversarial personalized ranking for recommendation. In *Proc. ACM SIGIR*, pages 355–364, 2018.

[3] Evaggelia Pitoura, Kostas Stefanidis, and Georgia Koutrika. Fairness in rankings and recommendations: An overview. *The VLDB Journal*, 31(3):431–458, 2021.

[4] Ziwei Zhu, Xia Hu, and James Caverlee. Fairness-aware tensor-based recommendation. In *Proc. CIKM*, pages 1153–1162, 2018.

[5] Meike Zehlike, Ke Yang, and Julia Stoyanovich. Fairness in ranking, part ii: Learning-to-rank and recommender systems. *ACM Comput. Surv.*, 55(6), dec 2022.

[6] Weiwen Liu, Feng Liu, Ruiming Tang, Ben Liao, Guangyong Chen, and Pheng Ann Heng. Balancing between accuracy and fairness for interactive recommendation with reinforcement learning. In *Advances in Knowledge Discovery and Data Mining*, pages 155–167, 2020.

[7] Mengting Wan, Jianmo Ni, Rishabh Misra, and Julian McAuley. Addressing marketing bias in product recommendations. In *Proc. WSDM*, pages 618–626, 2020.

[8] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bayesian personalized ranking from implicit feedback. In *Uncertainty in Artificial Intelligence*, pages 452–461, 2014.