Monotonic Fairness in Recommendation via Neural Additive Models with Contrastive Learning

Yuchen Guo[1,*], Zhenqi Zhao[2] and Menghan Wang[1]

[1]eBay
[2]Coupang

Abstract

Research on recommendation fairness has risen up in recent years. As recommender systems are typically customer-centric design, customer-side fairness is extensively studied. However, some less obvious fairness issues hidden on the provider side have not received comparable attention. One remaining question is whether product competitiveness in recommender systems (e.g., recommendation scores) exhibits monotonic variation with respect to protected sensitive attributes on the provider side, such as item prices in E-commerce or ad bids in sponsored search. It is inherently unfair for sellers if the recommendation score of their listed products decreases as they lower their prices. Our investigation reveals that such instances of unfairness are not uncommon in recommender systems. In this paper, we define this phenomenon as an individual monotonic fairness issue, and propose a novel, fairness-aware framework to address it. Our approach leverages monotonic neural additive models, theoretically ensuring monotonicity, and incorporates contrastive learning to enhance fairness through augmented samples. Additionally, we introduce specific evaluation metrics to quantify fairness. Extensive experiments on real-world datasets demonstrate that our method significantly improves monotonic fairness while still maintaining a high level of personalization compared to state-of-the-art recommendation algorithms. The source codes are available at https://github.com/yuchguo1007/MNAM-CL.

Keywords

Fairness-aware Recommendation, Neural Additive Model, Contrastive Learning

1. Introduction

Over recent decades, recommender systems have rapidly evolved and become integral to modern web and mobile applications like eBay, Netflix, and Spotify. These online platforms serve as intermediaries between content providers (e.g., eBay sellers, Netflix filmmakers, Spotify artists) and customers by offering recommendation services. Traditional personalized recommendations strive to enhance customer satisfaction by suggesting products to best match customers' interest, relying on historical user interactions. However, such data-driven design inevitably introduces unfairness, either on the customer side or the provider side. With the awakening of the unfairness in recommender systems, which is broadly defined as harmful disparity in user experience [1, 2, 3, 4, 5, 6], research on recommendation fairness has surged. Previous studies attempted to optimize two competing goals simultaneously: maximizing recommendation accuracy and minimizing the prediction discrimination of different subgroups (e.g., gender, age).

Unlike previous fairness studies, our investigation emphasizes the existence of *monotonic unfairness* in state-of-the-art recommendation algorithms. Briefly, we term it *monotonic fairness* when the recommendation score changes favorably for providers with respect to protected attributes, assuming all other attributes are held constant. We have observed *monotonic unfairness* in real cases, where sellers lowered the price of their items, leading to a decrease in the corresponding model score instead. Similarly, in other recommendation scenarios, an advertiser increases the bid rate, yet the winning rate of the advertisement does not improve; or in movie recommender systems, the mean star rating of a movie increases, but results in a decrease of its competitiveness. As shown in Figure 1, *monotonic*

ECOM'25: SIGIR Workshop on eCommerce, Jul 17, 2025, Padua, Italy
*Corresponding author.
yuchguo@ebay.com (Y. Guo); zhzhao2@coupang.com (Z. Zhao); menghawang@ebay.com (M. Wang)

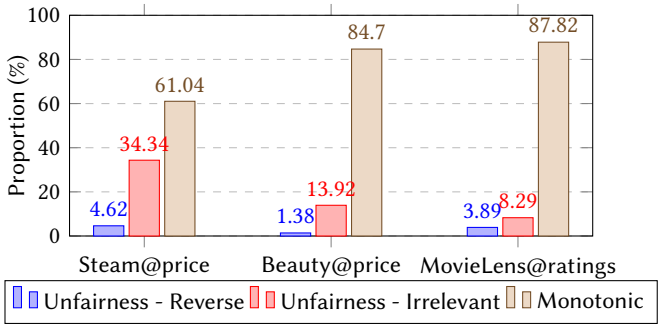

Figure 1: Provider-side monotonic unfairness is observed on three different datasets. Recommendation scores are generated by the two-tower model [7].

unfairness is observed in three different public datasets when applying the two-tower model [7] for recommendation.

Specifically, these unfairness situations can be classified into two types: *irrelevant*, where changes in protected attributes do not affect the recommendation score, and *reverse*, where the recommendation score changes contrary to the providers' expectations with respect to protected attributes. The underlying reason of these phenomena in recommender systems is the indiscriminate treatment of protected and unprotected attributes within complex nonlinear transformations of the recommendation model, causing changes to protected attributes to be easily obliterated.

To tackle the innegligible problem, we propose a novel method that promotes monotonic fairness without compromising performance. Firstly, we provide a clear and intuitive definition of monotonic fairness, which can be evaluated through pairwise ranking accuracy. Futhermore, we classify the unfairness scenarios into two types: *reverse* and *irrelevant*. The *reverse* instances are addressed by isolating protected attributes from the population and constructing a certified monotonic neural additive model. To address the more challenging *irrelevant* cases, we employ self-supervised contrastive learning to enhance model training via augmented samples.

The contributions of this paper are summarized as follows:

- We propose and formulate the problem of monotonic fairness in recommender systems and introduce new metrics to measure it.
- Through theoretical analysis, we design MNAM-CL (Monotonic Neural Additive Model with Contrastive Learning), which is implemented using neural additive models with contrastive learning. To our best knowledge, it is the first work to address the monotonic unfairness for deep recommendations in a principled way.
- We evaluate MNAM-CL on three different real-world datasets, demonstrating its ability to achieve better monotonic fairness while maintaining state-of-the-art performance compared to other competitive algorithms, as well as showing this method could be easily applied to different existing deep models for recommendation.

2. Related Work

2.1. Provider Fairness Discovery

Provider fairness, or supplier fairness, refers to not discriminating against individual provider or groups on sensitive attributes [3, 8, 9]. On one hand, unfairness in recommender system is usually caused by various forms of biases [10, 6], which mainly comes from the original training data. On the other hand, increasingly complex structure of embedding based deep models did bring huge improvements for data-driven recommender systems, but this also exacerbates the risk of amplifying the bias, which finally leads to unfairness on sensitive attributes, as well as harming the benefit of the minority. Recent work

has explored the development of deep models with fairness-aware regulations to achieve fairness [5, 2]. These kind of fairness should be ensured regardless training data or models. Therefore, a two-pronged approach of data augmentation and model structure improvement is a reasonable and effective solution to this unfairness.

2.2. Neural Additive Model

Neural Additive Models (NAMs) make restrictions on the structure of neural networks, which yields a family of models that are inherently interpretable while suffering little loss in prediction accuracy when applied to tabular data. Methodologically, NAMs belong to a larger model family called Generalized Additive Models (GAMs) [11].

NAMs learn a linear combination of networks that each attend to a single input feature: each in the traditional GAM formulationis parametrized by a neural network [12, 13]. These networks are trained jointly using backpropagation and can learn arbitrarily complex shape functions. With the continuous development of AI technology, NAMs play a very important role in fields where interpretable and explainable models are required, such as healthcare and finance [14]. Interpreting NAMs is easy as the impact of a feature on the prediction does not rely on the other features and can be understood by visualizing its corresponding shape function. NAMs are more easily extendable than existing GAMs due to their differentiability and composability.

2.3. Contrastive Learning

Contrastive learning has emerged as a powerful paradigm in unsupervised and self-supervised learning techniques [15, 16, 17] by significantly reducing the performance gap between supervised and unsupervised learning. At its core, contrastive learning aims to learn similar representations for semantically similar instances and dissimilar representations for distinct ones. It accomplishes this by continuous optimizing target contrastive learning loss, through a variety of corresponding data augmentation methods [18]. Especially, by leveraging large amounts of unlabeled data, it opens up new avenues for model training in scenarios where labeled data is scarce or expensive to obtain. The effectiveness of this approach has been showcased in numerous applications, such as image, speech recognition, and natural language processing [19, 20, 21].

3. Monotonic Fairness

In this section, we introduce the definition of monotonic fairness and its measurement methodology.

3.1. Definition

DEFINITION 1 (MONOTONIC FAIRNESS)**.** Generally, assume we can partition a multi-dimensional vector $x \in \mathbb{R}^d$ into $x = (p, q) \in \mathbb{R}^{d-m} \times \mathbb{R}^m$, such that a function $f(x)$ for $x = (p, q)$ is *monotonic* on q if this inequality holds[1]:

$$f(p, q) < f(p, q'), \forall p, \forall q < q', \tag{1}$$

where $q < q'$ denotes the inequality for all the elements (i.e., $q_i \leq q'_i$ for all $1 \leq i \leq m$, where q_i denotes the i-th element of q). The formula above shows that f is monotonic on q. For a differentiable function f, Equation (1) is equivalent to:

$$\min_{i \in [1,m]} \frac{\partial f(p, q)}{\partial q_i} \geq 0. \tag{2}$$

Assume that p and q refer to the unprotected and protected attributes in recommender systems, Equation (2) indicates the individual monotonic fairness of each protected attribute.

[1]Assume that all monotonic constraints are increasing; the monotonically non-increasing case can be considered analogously.

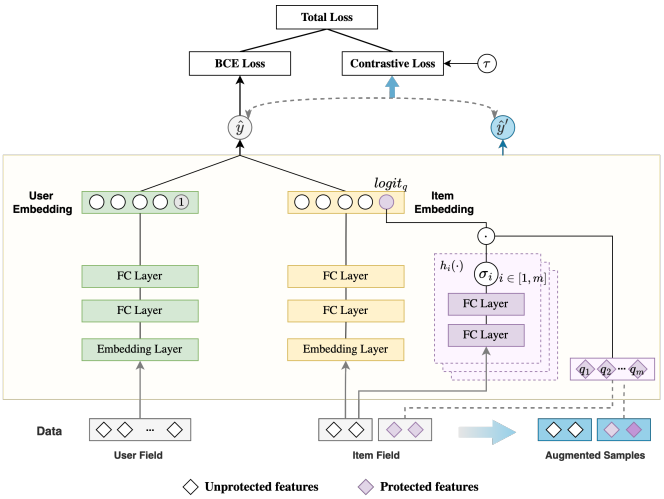

Figure 2: Overall architecture of MNAM-CL. Each of the protected attributes q_i gets an expert network $h_i(\cdot)$ as the weight, the input of which are unprotected features. $\hat{y}$ and $\hat{y}'$ are model scores from the regular sample and augmented sample, respectively. τ is the weight factor for contrastive loss.

3.2. Metric

According to Definition 1, the best match metric to measure monotonic fairness is pairwise ranking accuracy [22, 23, 24, 25], where the idea is to calculate the accuracy of a system ranking a pair of items correctly conditioned on the true outcome. Formally, the metric is defined as:

$$\begin{aligned} \text{PairAcc} &= P(f(x) < f(x') | x \in \mathcal{X}, x' \in \mathcal{X}) \\ &= P(f(p,q) < f(p,q') | \forall p, \forall q < q') \end{aligned} \tag{3}$$

Intuitively, this metric means that given an item from dataset $\mathcal{X}$, the probability of model score keeps monotonic compared with itself when the protected feature q varies. When referring to cases not meeting the criteria (i.e., unfair cases), they can be divided into two types:

$$\text{Unfairness} = \begin{cases} P(f(p,q) > f(p,q') | \forall p, \forall q < q'), & \textit{reverse} \\ P(f(p,q) = f(p,q') | \forall p, \forall q < q'), & \textit{irrelevant} \end{cases}. \tag{4}$$

As shown in Equation (4), *reverse* and *irrelevant* represent different levels of unfairness, both of which are the targets to be eliminated. For *irrlevant* cases, the tolerable precision of the equal sign is set to 1e-6.

4. Proposed Model

In this section we describe the proposed framework in details. The architecture of MNAM-CL, as illustrated in Figure 2, is based on the classic two-tower model. It addresses *reverse* unfairness instances by isolating protected attributes from the population and constructing a certified monotonic neural additive model. Furthermore, to tackle the more challenging *irrelevant* instances, where subtle changes in protected attributes fail to influence final outcomes, we employ an additional approach. Augmented samples are generated through a self-supervised manner by simulating real changes to protected attributes, serving as extra data for fairness tasks to assist model training.

4.1. Certified Monotonic Neural Additive Models

Neural additive models (NAMs) [12] consist of a linear combination of neural networks that each attend to a single input feature, making it possible for learning arbitrarily complex relationships between

Algorithm 1: Procedures of Data Augmentation (Pairwise)

Input: A training dataset $\mathscr{D} = \{\mathscr{X}, \mathscr{Y}\}$.
Output: Augmented dataset $\mathscr{D}' = \{(\mathscr{X}, \mathscr{X}'), \mathscr{Y}'\}$.
for $i \leftarrow 1$ **to** m **do**
  $\alpha \leftarrow$ lower boundary of q_i;
  $\beta \leftarrow$ upper boundary of q_i;
  for $x_j \in \mathscr{X}$ **do**
    $q'_{ij} = random(\alpha, \beta)$;
    $x'_{ij} = (p, q_1, \cdots, q'_i, \cdots, q_m)_j$;
    $y'_{ij} = \mathbb{1}(q_{ij} < q'_{ij})$;
    Generate an augmented sample $((x_j, x'_{ij}), y'_{ij})$;
  end
end

their input feature and the output. Drawing inspiration from NAMs, we develop a Monotonic Neural Additive Model (MNAM), as illustrated in Figure 2. Each of the protected attributes gets an expert network as the weight, the input of which are unprotected features. Thus MNAM is mathematically formulated as follows:

$$f(x) = g(p) + \sum_{i=1}^{m} h_i(p) \cdot \Theta_i(q_i), \tag{5}$$

where $g(\cdot)$ is the score function for unprotected features, $\Theta_i(\cdot)$ is the normalization function for i-th protected feature, and $h_i(\cdot)$ is the weight function, respectively. The partial derivative of $f(\cdot)$ with respect to q is given by the following formula:

$$\min_{i \in [1,m]} \frac{\partial f(p,q)}{\partial q_i} = \min_{i \in [1,m]} \frac{\partial h_i(p) \cdot \Theta_i(q_i)}{\partial q_i} = h_i(p) \cdot \Theta'_i(q_i). \tag{6}$$

Under the condition that $h_i(\cdot) > 0$ and $\Theta_i(\cdot)$ is differentiable, the above derivation theoretically guarantees monotonicity between q and f as long as $\Theta'_i(q_i) \geq 0$. In our work, $h_i(\cdot)$ is a multi-layer fully connected network with sigmoid as the final activation function, which satisfies the monotonic constraint while introducing nonlinearity.

4.2. Self-supervised Contrastive Learning

Contrastive learning aims to learn generalizable and transferable representations from unlabeled data using contrastive pairs [20]. In our study, we focus on protected attributes, augmenting original training data in a self-supervised manner.

4.2.1. Data Augmentation

The augmentation module plays a crucial role in contrastive learning, as indicated by previous research[17, 16, 15]. Based on the hypothesis of the monotonicity relationship between protected attributes and model scores, it naturally guides the following data augmentation approach, as shown in Algorithm 1. Specifically, we generated additional augmented samples by randomly adjusting the value of protected attributes within certain constraints while keeping the remaining unprotected features unchanged. The augmentation simply simulates changes to the protected attributes on provider side, actively or passively. According to the assumption of monotonicity, it becomes straightforward to determine the model score relationship between the original samples and the generated samples.

4.2.2. Loss Calculation

With the above data augmentation manner, the choice of loss function is significant. As shown in Figure 2, MNAM-CL consists of two kinds of losses: primary loss and fairness loss. For the primary task we

Table 1
Datasets Statistics.

Dataset	#User	#Item	#Rating	Positive Ratio	Protected Attribute
MovieLens-1M	6,040	3,706	1×10^6	0.5751	ratings
Steam	26,542	11,423	9×10^5	0.4719	price
Amazon Beauty	1.56×10^6	1.58×10^5	1×10^6	0.6380	ratings, price

use pointwise cross entropy as the loss function:

$$L_{primary} = -\sum_{j=1}^{n} y_j \log(\hat{y}_j) + (1 - y_j) \log(1 - \hat{y}_j), \tag{7}$$

where y_j and $\hat{y}_j$ are the true label and the sigmoid probability of j-th sample, respectively. As for the fairness task, we use BPR loss [26] to measure the mutual information between the original samples and augmented samples.

Let $\Delta\hat{y}_{ij} = \hat{y}'_{ij} - \hat{y}_j$, then we have

$$L_{fairness} = \sum_{j=1}^{n} \sum_{i=1}^{m} y'_{ij} \log \sigma(\Delta\hat{y}_{ij}) + (1 - y'_{ij}) \log \sigma(1 - \Delta\hat{y}_{ij}), \tag{8}$$

where $\hat{y}'_{ij}$ is the predicted score of augmented sample x'_{ij}, and σ is sigmoid function. Finally the total loss function becomes a linear combination of Equation (7) and Equation (8). We also add $L2$-regularization terms to avoid overfitting:

$$L_{total} = L_{primary} + \tau * L_{fairness} + regularization, \tag{9}$$

where τ is a temperature hyperparameter. MNAM-CL basically follows the conventional stochastic gradient descent (SGD) training routine.

5. Experiments

In this section, we conduct experiments on three public datasets and answer the following Research Questions (RQs):

- **RQ1:** How to define sensitive attributes for monotonic fairness?
- **RQ2:** Is monotonic unfairness prevalent in standard models in recommendation?
- **RQ3:** Could MNAM completely eliminate monotonic unfairness?

5.1. Datasets and Experimental Settings

5.1.1. Dataset

We evaluate fairness and recommendation performance on three datasets: MovieLens[2], Steam[3] and Beauty[4]. Table 1 provides a summary of the dataset statistics. Note the original ratings of some datasets are explicit integer ratings range from 1 to 5, and we transform them into binary labels by threshold 3 to construct a binary classification model. Additionally, we mark protected attributes for each dataset, selecting one attribute for both MovieLens (i.e., ratings) and Steam (i.e., price), and multiple attributes for Beauty (i.e., ratings and price). It is worth mentioning that after splitting the validation set from raw dataset at a ratio of 20%, we follow the same procedures as Algorithm 1 to generate a monotonic fairness validation set.

[2]https://grouplens.org/datasets/movielens/
[3]https://cseweb.ucsd.edu/~jmcauley/datasets.html#steam_data
[4]http://snap.stanford.edu/data/amazon/productGraph/categoryFiles/

5.1.2. Competitors

We compare MNAM-CL with several baselines: (1) **Wide&Deep** [27]. This model combines both linear models and deep models to improve memorization and generalization capabilities. (2) **NeuMF** [28]. A neural network-based collaborative filtering method utilizing binary cross-entropy loss. (3) **NAM** [12] An explainable model that learns a separate subnetwork for each input feature and combines their outputs through an additive operation. (4) **T2** [7]. A widely-used Two-Tower model in recommender systems that employs two separate deep neural networks to learn user and item embeddings independently. MANM-CL in this paper belongs to the family of generalized two-tower models.

5.1.3. Evaluation Protocols

We use the classical ROC-AUC and NDCG@10 to measure model accuracy. As defined in section 3.2, we select *reverse* rate and *irrelevant* rate to measure monotonic fairness, where smaller values denotes better fairness performance. For parameter settings, τ is set to 0.1 as the weight of $L_{fairness}$ in loss function, while L2 regularization coefficient is set to 1e-6.

Table 2
Comparisons of different models on three datasets.

Models	MovieLens-1M		@ratings		Steam		@price	
	ROC-AUC	NDCG@10	*reverse*	*irrelevant*	ROC-AUC	NDCG@10	*reverse*	*irrelevant*
Wide&Deep	0.7133	0.4082	0.2802	0.4437	0.8915	0.3539	0.2589	0.5158
NeuMF	0.7038	0.4044	0.0008	0.1557	0.8947	0.3546	0.0027	0.5789
NAM	0.7156	**0.4207**	0.0031	0.0723	**0.9068**	0.3568	0.0189	0.0003
T2	**0.7179**	0.4122	0.0389	0.0829	0.8907	0.3511	0.0462	0.3434
MNAM	0.7128	0.4090	**0.0**	7.9e-5	0.8962	0.3565	**0.0**	6.2e-5
MNAM-CL	0.7133	0.4093	**0.0**	**4.9e-5**	0.9036	**0.3569**	**0.0**	**5.1e-5**

Models	Beauty		@ratings		@price	
	ROC-AUC	NDCG@10	*reverse*	*irrelevant*	*reverse*	*irrelevant*
Wide&Deep	0.6120	0.2706	0.3163	0.3674	0.3272	0.3464
NeuMF	0.6164	0.2787	0.0132	0.0755	**0.0**	0.1980
NAM	**0.6260**	**0.2791**	**0.0**	0.0020	0.1414	0.0066
T2	0.6257	0.2764	0.0031	0.0872	0.0138	0.1392
MNAM	0.6143	0.2788	**0.0**	0.0001	**0.0**	0.0003
MNAM-CL	0.6140	0.2788	**0.0**	**2.1e-5**	**0.0**	**0.0002**

5.2. Scope of Protected Sensitive Attributes (RQ1)

Unlike inherent attributes of human beings, such as gender, age and race, which segment users into subgroups, sensitive attributes related to monotonic fairness are more individual and quantifiable dimensions. Ensuring fairness along these dimensions in recommender systems is critical for platforms, as neglecting them can lead to a lose-lose scenario for both providers and platforms. For instance, price is a critical factor influencing user purchase. Consider a seller who lowers the price of his item. If the platform's recommender system does not protect price attribute, the item's model score might fail to increase as expected. This misalignment could result in lost transaction opportunities—a loss for both the seller and the platform. In summary, any attribute that may disrupt the online platform ecosystem if it fails to meet the monotonic fairness criteria in Definition 1, falls within the scope of sensitive attributes discussed in this paper. These include, but are not limited to, item prices for sellers, bids for advertisers, and movie ratings for filmmakers.

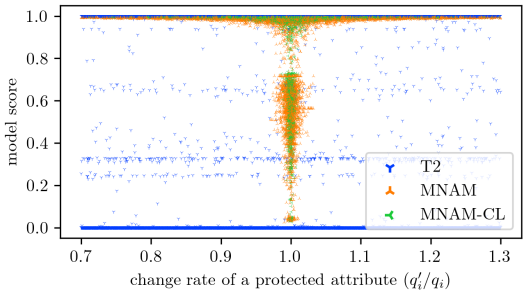

Figure 3: Unfairness distribution: Amazon Beauty@price. Each dot represents a bad case, where the model score fails to change in the expected direction after the protected attribute shifts from q_i to q_i'

5.3. Monotonic Fairness Evaluation (RQ2)

Before the real research begins, it is essential to clarify the current status of monotonic fairness in representative recommender systems. As previously defined, monotonic unfairness on protected attributes in recommendation is divided into *reverse* and *irrelevant*. We evaluate various baseline models on the same datasets to observe their fairness performance. Table 2 demonstrates that these baseline models still suffer from monotonic unfairness of varying degrees, despite good ranking metrics. Such results emphasize the importance of addressing monotonic unfairness from both the provider and system perspectives.

5.4. Effectiveness of MNAM (RQ3)

Theoretically, MNAM alone could eliminate all *reverse* cases, which is verified by experiments, as shown in Table 2. Nonetheless, MNAM still exhibits a certain number of irrelevant cases. According to Equation (5), $h_i(p)$ is trained as the weight factor for q_i, influencing the contribution of protected attributes to the final score. Therefore the *irrelevant* cases occur when $h_i(p) \cdot |\Theta_i(q_i') - \Theta_i(q_i)| \leq$ 1e-6. The reason why MNAM is powerless in reducing *irrelevant* cases is the absence of supervised constraints. In summary, MNAM cannot completely eliminate monotonic unfairness, indicating a need for improvement in reducing *irrelevant* cases.

We conduct additional experiments for investigating effectiveness of each component in MNAM-CL. Figure 3 presents the unfairness distribution of Beauty@price, illustrating that the base model (T2) has a relatively high occurrence of unfairness indiscriminately. In comparison, MNAM tends to generate unfairness only when the score is near the upper bound 1.0 or the change rate is small ($q_i'/q_i \approx 1.0$). This clearly demonstrates the alignment between the effect and design of MNAM. Compared with other ablation versions, MNAM-CL achieves improved monotonic fairness, particularly in alleviating *irrelevant* cases. The further reduction of unfairness from MNAM to MNAM-CL proves the effectiveness of data augmentation and contrastive learning. The remaining unfair cases are mostly caused by diminishing marginal effects, which is more acceptable from an ethical perspective. Furthermore, we evaluate monotonic fairness both on single and multiple attributes, where MNAM-CL consistently outperforms with stable performance.

6. Conclusion and Future Work

In this paper, we study individual monotonic fairness in recommender systems, and propose MNAM-CL, a novel framework for reducing such unfairness. It has the capability to eliminate all *reverse* cases. Furthermore, it enhances fairness by data augmentation and contrastive learning according to specific scenarios and attributes. Extensive evaluations verify its effectiveness in modeling monotonic fairness while maintaining recommendation accuracy. Fairness in process is the premise of fairness in result. In the future work, more complex pairwise monotonic fairness would be further explored.

References

[1] H. Yang, Z. Liu, Z. Zhang, C. Zhuang, X. Chen, Towards robust fairness-aware recommendation, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 211–222.

[2] J. Tang, S. Shen, Z. Wang, Z. Gong, J. Zhang, X. Chen, When fairness meets bias: a debiased framework for fairness aware top-n recommendation, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 200–210.

[3] G. K. Patro, A. Biswas, N. Ganguly, K. P. Gummadi, A. Chakraborty, Fairrec: Two-sided fairness for personalized recommendations in two-sided platforms, in: Proceedings of the web conference 2020, 2020, pp. 1194–1204.

[4] L. Chen, L. Wu, K. Zhang, R. Hong, D. Lian, Z. Zhang, J. Zhou, M. Wang, Improving recommendation fairness via data augmentation, in: Proceedings of the ACM Web Conference 2023, WWW '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 1012–1020.

[5] R. Mehrotra, J. McInerney, H. Bouchard, M. Lalmas, F. Diaz, Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems, in: Proceedings of the 27th acm international conference on information and knowledge management, 2018, pp. 2243–2251.

[6] Y. Wu, J. Cao, G. Xu, Y. Tan, Tfrom: A two-sided fairness-aware recommendation model for both customers and providers, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021, pp. 1013–1022.

[7] X. Yi, J. Yang, L. Hong, D. Z. Cheng, L. Heldt, A. Kumthekar, Z. Zhao, L. Wei, E. Chi, Sampling-bias-corrected neural modeling for large corpus item recommendations, in: Proceedings of the 13th ACM Conference on Recommender Systems, 2019, pp. 269–277.

[8] L. Chen, L. Wu, K. Zhang, R. Hong, D. Lian, Z. Zhang, J. Zhou, M. Wang, Improving recommendation fairness via data augmentation, arXiv preprint arXiv:2302.06333 (2023).

[9] E. Gómez, L. Boratto, M. Salamó, Provider fairness across continents in collaborative recommender systems, Information Processing & Management 59 (2022) 102719.

[10] T. Qi, F. Wu, C. Wu, P. Sun, L. Wu, X. Wang, Y. Huang, X. Xie, Profairrec: Provider fairness-aware news recommendation, in: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022, pp. 1164–1173.

[11] T. J. Hastie, Generalized additive models, in: Statistical models in S, Routledge, 2017, pp. 249–307.

[12] R. Agarwal, L. Melnick, N. Frosst, X. Zhang, B. Lengerich, R. Caruana, G. E. Hinton, Neural additive models: Interpretable machine learning with neural nets, Advances in Neural Information Processing Systems 34 (2021) 4699–4711.

[13] H. Zhuang, X. Wang, M. Bendersky, A. Grushetsky, Y. Wu, P. Mitrichev, E. Sterling, N. Bell, W. Ravina, H. Qian, Interpretable ranking with generalized additive models, in: Proceedings of the 14th ACM International Conference on Web Search and Data Mining, 2021, pp. 499–507.

[14] W. Zhang, B. Barr, J. Paisley, Gaussian process neural additive models, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, 2024, pp. 16865–16872.

[15] R. Luo, Y. Wang, Y. Wang, Rethinking the effect of data augmentation in adversarial contrastive learning, in: The Eleventh International Conference on Learning Representations, 2023. URL: https://openreview.net/forum?id=0qmwFNJyxCL.

[16] J. D. Robinson, C.-Y. Chuang, S. Sra, S. Jegelka, Contrastive learning with hard negative samples, in: International Conference on Learning Representations, 2021. URL: https://openreview.net/forum?id=CR1XOQ0UTh-.

[17] S. Ge, S. Mishra, C.-L. Li, H. Wang, D. Jacobs, Robust contrastive learning using negative samples with diminished semantics, Advances in Neural Information Processing Systems 34 (2021) 27356–27368.

[18] P. Zhou, Y.-L. Huang, Y. Xie, J. Gao, S. Wang, J. B. Kim, S. Kim, Is contrastive learning necessary? a study of data augmentation vs contrastive learning in sequential recommendation, in: Proceedings of the ACM on Web Conference 2024, 2024, pp. 3854–3863.

[19] C. Yang, J. Zou, J. Wu, H. Xu, S. Fan, Supervised contrastive learning for recommendation,

Knowledge-Based Systems 258 (2022) 109973.
[20] J. Zhang, K. Ma, Rethinking the augmentation module in contrastive learning: Learning hierarchical augmentation invariance with expanded views, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 16650–16659.
[21] W. Xia, C. Zhang, C. Weng, M. Yu, D. Yu, Self-supervised text-independent speaker verification using prototypical momentum contrastive learning, arXiv preprint arXiv:2012.07178 (2020).
[22] Y. Wu, J. Cao, G. Xu, Fairness in recommender systems: evaluation approaches and assurance strategies, ACM Transactions on Knowledge Discovery from Data 18 (2023) 1–37.
[23] A. Beutel, J. Chen, T. Doshi, H. Qian, L. Wei, Y. Wu, L. Heldt, Z. Zhao, L. Hong, E. H. Chi, et al., Fairness in recommendation ranking through pairwise comparisons, in: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019, pp. 2212–2220.
[24] A. Fabris, G. Silvello, G. A. Susto, A. J. Biega, Pairwise fairness in ranking as a dissatisfaction measure, in: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, 2023, pp. 931–939.
[25] X. Wang, N. Thain, A. Sinha, F. Prost, E. H. Chi, J. Chen, A. Beutel, Practical compositional fairness: Understanding fairness in multi-component recommender systems, in: Proceedings of the 14th ACM International Conference on Web Search and Data Mining, 2021, pp. 436–444.
[26] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, Bpr: Bayesian personalized ranking from implicit feedback, in: UAI 2009, Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, June 18-21, 2009, 2009.
[27] H.-T. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, et al., Wide & deep learning for recommender systems, in: Proceedings of the 1st workshop on deep learning for recommender systems, 2016, pp. 7–10.
[28] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, T.-S. Chua, Neural collaborative filtering, in: WWW, 2017, pp. 173–182.