Explaining Unfairness in GNN-based Recommendation

Ludovico Boratto
University of Cagliari
ludovico.boratto@acm.org

Francesco Fabbri
Spotify
francescof@spotify.com

Gianni Fenu
University of Cagliari
fenu@unica.it

Mirko Marras
University of Cagliari
mirko.marras@acm.org

Giacomo Medda
University of Cagliari
giacomo.medda@unica.it

Abstract

Nowadays, research into personalization has been focusing on explainability and fairness. Several approaches are able to explain individual recommendations, but unfairness explainability has been limited to finding user/item features mostly related to biased recommendations. In this paper, we devised a novel algorithm that leverages counterfactuality methods to discover user-item interactions as user unfairness explanations in GNN-based recommendation. Our approach perturbs the graph topological structure to find an altered version (counterfactual explanation) that minimizes the disparity in utility between the protected and unprotected demographic groups. Experiments on four real-world graphs showed that our method can systematically explain user unfairness on three state-of-the-art GNN-based recommendation models. Moreover, the generated unfairness explanations are connected to properties related to the networks topological aspects. Source code and datasets: https://github.com/jackmedda/RS-BGExplainer.

1 Introduction

Context and Motivation. Even though novel recommender systems improve the users' satisfaction, researchers have gradually increased their focus towards the importance of preserving additional properties, such as fairness [1, 2] and explainability [3]. Usually, these properties are independently investigated, one at a time. Considering the main case study in our work, i.e. algorithmic fairness, it is important for service providers to understand *why* the model behind their platform is unfair.

State of the Art. Explanation techniques are receiving increasing attention not only in the field of recommender systems, but also in the more general domain of Graph Neural Networks (GNNs) [4].

Explanation methods in recommendation mainly rely on post-hoc [5, 6] and path-based [7–9] techniques. Conversely, counterfactual techniques have been proved to be effective for deriving explanations of GNNs outcomes in several downstream tasks [10–12], including recommendation [13].

In GNNs, counterfactual reasoning implies manipulating the topological structure of the graph, to find the minimal set of perturbations that changes the model prediction. The same approach has been used to mitigate unfairness in several downstream tasks [14–16], but only [17] focused on identifying the patterns that globally explain unfairness issues in GNNs for node classification.

The few approaches explaining unfairness in recommendation exploit surrogate models [18] and counterfactual reasoning [19]. Both of them identify user and item features causing recommendation disparities. Conversely, [20] adopts interaction dependency graphs to uncover model instability. The latter, which has global implications as the unfairness phenomenon, affects the system as a whole.

Open Issues. The existing unfairness explanation strategies face significant limitations, and as far as we know, no analogous method was ever proposed in GNN-based recommender systems.

L. Boratto et al., Explaining Unfairness in GNN-based Recommendation (Extended Abstract). Presented at the Second Learning on Graphs Conference (LoG 2023), Virtual Event, November 27–30, 2023.

First, algorithms designed to explain recommendations or GNNs predictions depend on local aspects, such as knowledge-aware paths [7, 8] or node neighborhoods [10–13], preventing their adoption to explain unfairness, which affects the system globally. Second, current solutions for unfairness explanation are inapplicable on datasets missing specific features required for the targeted task [18, 19], their functioning constraints the explanation to specific issues (e.g., model instability [20]), they adopt specialized fairness notions at the instance level [21] or node dropping strategies [17][1].

Our Contribution. In this work, we propose a framework, named `GNNUERS` (**GNN**-based **U**nfairness **E**xplainer in **R**ecommender **S**ystems), which perturbs the original bipartite graph in order to identify a set of user-item interactions leading to consumer unfairness in GNN-based recommender systems. Driven by counterfactual reasoning, consumer unfairness can be uncovered by the selected edges, i.e. the user-item interactions, because their removal from the original graph can result in fairer recommendations for the end users during the inference phase, thus globally explaining under which conditions a model generates disparities. The original bipartite graph is modified under the assumption that certain users did not interact with the items associated with the selected edges (in the counterfactual world), while, in reality, the graph includes such edges (in the actual world). We validate our proposal on three state-of-the-art models under four datasets, measuring key utility metrics and performing further analyses on the unfairness patterns discovered by our method.

2 GNNUERS: Methodology

2.1 Problem Formulation

`GNNUERS` aims to explain why a recommender system f is unfair. Typically, f uses the past interactions between users U and items I to learn the users' preference patterns. User-item interactions can be represented by an undirected bipartite graph $\mathcal{G} = (U, I, E)$, where E is the set of edges and $U \cup I$ is the set of nodes. Any GNN, defined as $f(A; W) \rightarrow \hat{R}$, can then be used to model the recommendation task as a link prediction one. A is an $n \times n$ ($n = |U| + |I|$) adjacency matrix representing $\mathcal{G}$, W is the learned weight matrix of f, and $\hat{R}_{u,i}$ is the linking probability between u and i. The top-k items are recommended to u based on their decreasing probability in $\hat{R}_u$.

Following [10, 12], `GNNUERS` leverages counterfactual techniques to generate an adjacency matrix $\tilde{A}$, minimally perturbed w.r.t. A. With a logic similar to [19], $\tilde{A}$ represents a counterfactual explanation of the unfairness if the trained GNN f produces altered and fairer recommendations when f uses $\tilde{A}$ instead of A in the inference phase. Formally, we seek to minimize the following objective function:

$$\mathcal{L}(A, \tilde{A}) = \mathcal{L}_{fair}(A, f(\tilde{A}, W)) + \mathcal{L}_{dist}(A, \tilde{A}) \tag{1}$$

where $\mathcal{L}_{dist}$ controls the distance between $\tilde{A}$ and A, $\mathcal{L}_{fair}$ is the term monitoring fairness.

2.2 Framework

Perturbation Mechanism. Following [10], we generate a perturbed adjacency matrix $\tilde{A}$. Differently from [10], we adopt perturbations at the global level, and leverage a vector $p \in \{0, 1\}^B$ (B is the number of candidate edges, e.g., $B = |E|$) instead of a matrix P to only affect meaningful entries in A, i.e. where $A_{u,i} \neq 0$. p perturbs such entries through a function $h : \mathbb{N}^2 \rightarrow \mathbb{N}$ that maps the 2D indices (u, i) of A to a 1D index for p. Thus, given $t = h(u, i), t < B$, an entry $p_t = 0$ denotes the edge $A(u, i)$ is deleted in $\tilde{A}$, i.e. $\tilde{A}_{u,i} = 0$. Formally, $\tilde{A}$ is populated as follows:

$$\tilde{A}_{u,i} = \begin{cases} p_{h(u,i)} & if\ h(u,i) < B \\ A_{u,i} & otherwise \end{cases} \tag{2}$$

We derive p from an optimization process that uses a real valued vector $\hat{p}$, as [10, 22]. $\hat{p}$ is transformed by a sigmoid function, then its values ≥ 0.5 are rounded to 1, otherwise 0, eventually obtaining p.

Perturbed Graph Generation. `GNNUERS` generates $\tilde{A}$ by iteratively perturbing A with (2), and measuring the unfairness level on the altered matrix of linking probabilities $\tilde{R}$. At first, $\forall t \in [0, B)$,

[1] Node dropping would imply unfairness is caused by specific stakeholders, raising ethical concerns.

we initialize $\hat{p}_t = 0$, such that $p_t = 1$ after the sigmoid transformation and $\tilde{A}_0 = A$ is guaranteed. At the j-th iteration, the perturbation vector $\hat{p}$[2] is updated by (1) based on the altered matrices $\tilde{A}_{j-1}$ and $\tilde{R}_{j-1}$, where the latter was generated by feeding $\tilde{A}_{j-1}$ in input to f. With the updated vectors $\hat{p}$ and p, we obtain $\tilde{A}_j$, and the iterative process will continue until a certain condition is satisfied.

Loss Function Optimization. Grounded on the fairness notion of demographic parity, we set $\mathcal{L}_{fair}$ as the absolute disparity in recommendation utility S across demographic groups, as in [23, 24]. As recent works that focused on unfairness explanation or mitigation in recommendation [19, 25–27], we focus on a binary setting, with sensitive attributes comprised of two demographic groups. We use U_- or *unprotected*[3] (U_+ or *protected*) to denote the group with higher (lower) utility on the evaluation set. For the distance loss $\mathcal{L}_{dist}$, any differentiable distance function can be adopted. Formally:

$$\mathcal{L}_{fair}(A, \tilde{R}) = \left\| S(\tilde{R}^{U_-}, A^{U_-}) - S(\tilde{R}^{U_+}, A^{U_+}) \right\|_2^2 \quad \mathcal{L}_{dist}(A, \tilde{A}) = \beta \frac{1}{2} \sigma \left(\sum_{i,j} \left\| \tilde{A}_{i,j} - A_{i,j} \right\|_2^2 \right) \tag{3}$$

σ bounds $\mathcal{L}_{dist}$ to the same range of $\mathcal{L}_{fair}$[4], while β balances them (we tested values in $[0.001, 2.0]$). We use U_+ or U_- as superscripts on $\tilde{R}$ and A to denote the respective sub-matrices w.r.t. U_+ or U_-.

Normalized Discounted Cumulative Gain (NDCG) was selected as the utility metric S. However, due to its non-differentiability, in the optimization process we adopt an approximated version [24, 28], which we refer as *NDCGApproxLoss*, but still using NDCG in the evaluation phase. Given that our goal is to explain the recommendation unfairness measured on the evaluation set, we use the ground truth labels of the latter to measure *NDCGApproxLoss* during the GNNUERS learning process[5].

3 Experimental Evaluation and Discussion

Experimental Setting. We relied on the artifacts of [23], adopting the same splitting strategy and extending the set of datasets, i.e. MovieLens 1M (ML-1M) and Last.FM 1K (LFM-1K), with TaFeng (FENG) and Insurance (INS). Due to space constraints, we will detail and examine only ML-1M, FENG, and only the sensitive attribute *age*. As [23], age was binarized in *Younger* (Y) and *Older* (O). Relying on Recbole [30], we evaluated `GNNUERS` on GCMC [31], NGCF [32], and LigthGCN [33].

To analyze the explanations, we identified three properties: **Degree (DEG)** is a node's degree, **Density (DY)** describes the tendency of a node to be connected to high-degree nodes, **Intra-Group Distance (IGD)** describes how a node z is close to the other nodes $z' \in Z/\{z\}$, $Z \subset U$ or $Z \subset V$. Let $\mathcal{N}_z$ be the neighbors set of z, N the graph diameter, Γ a function that measures the shortest path length:

$$DY_z = \frac{\sum_{i=1}^{|\mathcal{N}_z|} \frac{|\{z' \mid (\bar{z}_i, z') \in E \wedge \bar{z}_i \in \mathcal{N}_z\}|}{|Z|}}{|\mathcal{N}_z|} \qquad IGD_z = \frac{\sum_{n=1}^{N} \frac{|\{z' \mid \Gamma(z, z') = n\}|}{n}}{|Z|} \tag{4}$$

For a user u, DEG represents the interest towards the available items, DY the inclination to engage with popular items, IGD the propensity to interact with items valued by the demographic group of u[6].

The literature does not include baselines that explain unfairness in the form of user-item edges as `GNNUERS`. We then compare our method with (i) its extension `GNNUERS+CN`, which limits the perturbation to the edges connected to the unprotected user nodes, (ii) CASPER [20], an algorithm that identifies the edge in a interaction dependency graph that causes the highest instability in a model, (iii) *RND-P*, which mimics the `GNNUERS` edges selection process, but based on a random choice[7].

Unfairness Explainability Benchmark. We first investigated the capability of `GNNUERS` to select counterfactual explanations that effectively optimize $\mathcal{L}_{fair}$. The evaluation is performed by randomly

[2] We constrain the floating-point oscillations of $\hat{p}$ values, preventing a deleted edge from being restored.

[3] The gradient computation of *NDCGApproxLoss* was not tracked for the unprotected group, so as to explain the unfairness in terms of the interactions that favored the other group.

[4] We adopt $\sigma(x) = |x|/(1 + |x|)$ which grows slower as x increases w.r.t. the logistic function.

[5] For other tasks, e.g., mitigation, having access to the ground truth labels is a less realistic assumption [29].

[6] Due to space constraints, we omit the table with datasets and graph properties statistics.

[7] If the selected edges do not optimize $\mathcal{L}$, the early stopping criterion might be triggered after a few iterations.

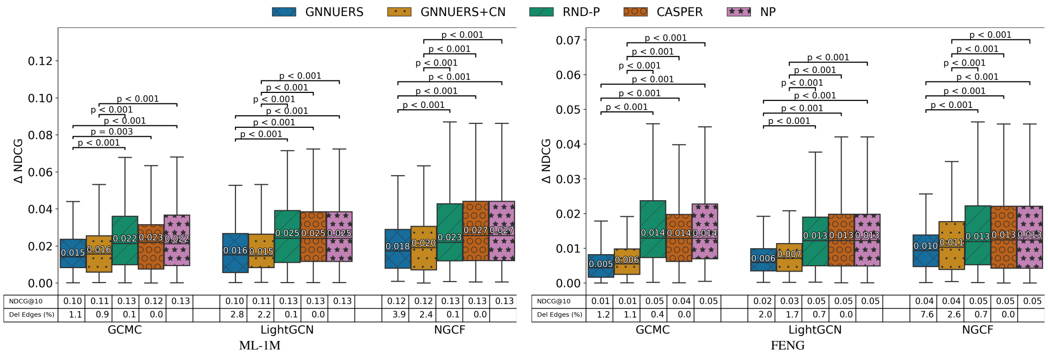

Figure 1: Distribution of |ΔNDCG| measured between age subgroups. A Wilcoxon signed-rank test is performed between each pair of boxes; p-value is shown if it is $< \frac{0.05}{m}$ according to the Bonferroni correction.

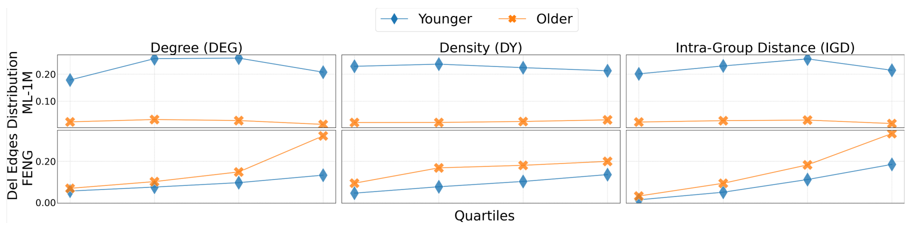

Figure 2: `GNNUERS` on NGCF: deleted edges distribution (*Del Edges Distribution*) over the quartiles (Q1-Q2-Q3-Q4) defined for each *age* group through sorting the nodes by each graph property.

sampling 100 subgroups of users coming from both demographic groups and measuring the utility disparity with |ΔNDCG|, i.e. $\mathcal{L}_{fair}$. This approach mimics the learning process adopted by `GNNUERS`.

Figure 1 shows the |ΔNDCG| distribution across the age subgroups. On the bottom, we include the NDCG@10 after the perturbation and the deleted edges fraction. Our methods significantly narrow the |ΔNDCG| distribution compared with `NP` (No Perturbation) by perturbing just 1% of the edges in some settings. Conversely, `CASPER` and `RND-P` are inefficient in reducing |ΔNDCG|.

A further analysis on our methods impact on NDCG (not reported here) proved that the latter is decreased more for the unprotected group w.r.t. the protected one, highlighting the edges are specifically selected to provide an explanation of the prior advantage in NDCG for the former group.

Edges Selection Process. We analyze the perturbed edges by profiling the respective user nodes with DEG, DY and IDG. User nodes were split by demographic groups and the nodes in each group were partitioned in quartiles: the data points order was defined by each property value in each quartile, e.g., low-DEG (Q1). For each group and each quartile, the number of perturbed edges was taken and normalized by the total number of edges perturbed in each experiment. We constrain the analysis to NGCF, where `GNNUERS`[8] minimally affects the protected group NDCG to explain the unfairness.

Figure 2 highlights the distribution of deleted edges is significantly higher for the unprotected group. Hence, unfair recommendations are highly dependent on such group interactions. On ML-1M, the over-representation and over-DEG of the Younger is enough to deviate the recommendations in their favor. No relevant pattern is reported across quartiles, but the Younger nodes themselves could be related to the unfairness, given that `GNNUERS` prunes more their edges compared to the Older ones. On FENG, `GNNUERS` deletes more edges connected to high-DEG and high-IGD (Q4) unprotected users (Older), i.e. Older users with the most interactions and the closest ones to other Older nodes. Older users in FENG are less represented, but their higher average DEG and IGD reflects such observations.

Discussion. In this work, we introduced `GNNUERS`, a framework to explain unfairness in GNN-based recommender systems. We proved the effectiveness of our method measuring its performance in unfairness explainability, and we also characterized the edges selected as explanation. Still, a more comprehensive investigation including other features, according, for instance, to some requirements defined by service providers, could provide more informative findings. Nonetheless, `GNNUERS` is flexible to be extended to substantially different GNN-based architectures, e.g., GAT.

[8] `GNNUERS+CN` was not analyzed because it affects only the unprotected group

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.*, jul 2022. ISSN 1046-8188. doi: 10.1145/3547333. URL https://doi.org/10.1145/3547333. Just Accepted. 1

[2] Ludovico Boratto, Gianni Fenu, Mirko Marras, and Giacomo Medda. Practical perspectives of consumer fairness in recommendation. *Inf. Process. Manag.*, 60(2):103208, 2023. doi: 10.1016/j.ipm.2022.103208. URL https://doi.org/10.1016/j.ipm.2022.103208. 1

[3] Yongfeng Zhang and Xu Chen. Explainable recommendation: A survey and new perspectives. *Found. Trends Inf. Retr.*, 14(1):1–101, mar 2020. ISSN 1554-0669. 1

[4] Hao Yuan, Haiyang Yu, Shurui Gui, and Shuiwang Ji. Explainability in graph neural networks: A taxonomic survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(5):5782–5799, 2023. doi: 10.1109/TPAMI.2022.3204236. URL https://doi.org/10.1109/TPAMI.2022.3204236. 1

[5] Léo Brunot, Nicolas Canovas, Alexandre Chanson, Nicolas Labroche, and Willeme Verdeaux. Preference-based and local post-hoc explanations for recommender systems. *Inf. Syst.*, 108: 102021, 2022. doi: 10.1016/j.is.2022.102021. URL https://doi.org/10.1016/j.is.2022.102021. 1

[6] Jinfeng Zhong and Elsa Negre. Shap-enhanced counterfactual explanations for recommendations. In Jiman Hong, Miroslav Bures, Juw Won Park, and Tomás Cerný, editors, *SAC '22: The 37th ACM/SIGAPP Symposium on Applied Computing, Virtual Event, April 25 - 29, 2022*, pages 1365–1372. ACM, 2022. doi: 10.1145/3477314.3507029. URL https://doi.org/10.1145/3477314.3507029. 1

[7] Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, and Mirko Marras. Post processing recommender systems with knowledge graphs for recency, popularity, and diversity of explanations. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, pages 646–656. ACM, 2022. doi: 10.1145/3477495.3532041. URL https://doi.org/10.1145/3477495.3532041. 1, 2

[8] Shichang Zhang, Jiani Zhang, Xiang Song, Soji Adeshina, Da Zheng, Christos Faloutsos, and Yizhou Sun. Page-link: Path-based graph neural network explanation for heterogeneous link prediction. In Ying Ding, Jie Tang, Juan F. Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben, editors, *Proceedings of the ACM Web Conference 2023, WWW 2023, Austin, TX, USA, 30 April 2023 - 4 May 2023*, pages 3784–3793. ACM, 2023. doi: 10.1145/3543507.3583511. URL https://doi.org/10.1145/3543507.3583511. 2

[9] Azin Ghazimatin, Oana Balalau, Rishiraj Saha Roy, and Gerhard Weikum. PRINCE: provider-side interpretability with counterfactual explanations in recommender systems. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining*, pages 196–204. ACM, 2020. 1

[10] Ana Lucic, Maartje A. ter Hoeve, Gabriele Tolomei, Maarten de Rijke, and Fabrizio Silvestri. Cf-gnnexplainer: Counterfactual explanations for graph neural networks. In *International Conference on Artificial Intelligence and Statistics, AISTATS 2022*, volume 151, pages 4499–4511. PMLR, 2022. 1, 2

[11] Hao Yuan, Jiliang Tang, Xia Hu, and Shuiwang Ji. XGNN: towards model-level explanations of graph neural networks. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pages 430–438. ACM, 2020.

[12] Bo Kang, Jefrey Lijffijt, and Tijl De Bie. Explanations for network embedding-based link predictions. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases - International Workshops of ECML PKDD 2021*, volume 1524, pages 473–488. Springer, 2021. 1, 2

[13] Ziheng Chen, Fabrizio Silvestri, Jia Wang, Yongfeng Zhang, Zhenhua Huang, Hongshik Ahn, and Gabriele Tolomei. GREASE: generate factual and counterfactual explanations for gnn-based recommendations. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases - International Workshops of ECML PKDD 2022, Proceedings*. Springer, 2022. 1, 2

[14] Chirag Agarwal, Himabindu Lakkaraju, and Marinka Zitnik. Towards a unified framework for fair and stable graph representation learning. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence, UAI 2021*, volume 161, pages 2114–2124. AUAI Press, 2021. 1

[15] Jing Ma, Ruocheng Guo, Mengting Wan, Longqi Yang, Aidong Zhang, and Jundong Li. Learning fair node representations with graph counterfactual fairness. In *WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining*, pages 695–703. ACM, 2022.

[16] Nan Wang, Lu Lin, Jundong Li, and Hongning Wang. Unbiased graph embedding with biased graph observations. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pages 1423–1433. ACM, 2022. 1

[17] Yushun Dong, Song Wang, Jing Ma, Ninghao Liu, and Jundong Li. Interpreting unfairness in graph neural networks via training node attribution. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 7441–7449, 2023. URL `https://ojs.aaai.org/index.php/AAAI/article/view/25905`. 1, 2

[18] Yashar Deldjoo, Alejandro Bellogín, and Tommaso Di Noia. Explaining recommender systems fairness and accuracy through the lens of data characteristics. *Inf. Process. Manag.*, 58(5): 102662, 2021. doi: 10.1016/j.ipm.2021.102662. URL `https://doi.org/10.1016/j.ipm.2021.102662`. 1, 2

[19] Yingqiang Ge, Juntao Tan, Yan Zhu, Yinglong Xia, Jiebo Luo, Shuchang Liu, Zuohui Fu, Shijie Geng, Zelong Li, and Yongfeng Zhang. Explainable fairness in recommendation. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 681–691. ACM, 2022. doi: 10.1145/3477495.3531973. URL `https://doi.org/10.1145/3477495.3531973`. 1, 2, 3

[20] Sejoon Oh, Berk Ustun, Julian J. McAuley, and Srijan Kumar. Rank list sensitivity of recommender systems to interaction perturbations. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 1584–1594. ACM, 2022. 1, 2, 3

[21] Yushun Dong, Song Wang, Yu Wang, Tyler Derr, and Jundong Li. On structural explanation of bias in graph neural networks. In Aidong Zhang and Huzefa Rangwala, editors, *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 14 - 18, 2022*, pages 316–326. ACM, 2022. doi: 10.1145/3534678.3539319. URL `https://doi.org/10.1145/3534678.3539319`. 2

[22] Suraj Srinivas, Akshayvarun Subramanya, and R. Venkatesh Babu. Training sparse neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2017*, pages 455–462. IEEE Computer Society, 2017. doi: 10.1109/CVPRW.2017.61. URL `https://doi.org/10.1109/CVPRW.2017.61`. 2

[23] Ludovico Boratto, Gianni Fenu, Mirko Marras, and Giacomo Medda. Consumer fairness in recommender systems: Contextualizing definitions and mitigations. In *Advances in Information Retrieval*, pages 552–566, Cham, 2022. Springer International Publishing. ISBN 978-3-030-99736-6. 3

[24] Haolun Wu, Chen Ma, Bhaskar Mitra, Fernando Diaz, and Xue Liu. A multi-objective optimization framework for multi-stakeholder fairness-aware recommendation. *ACM Trans. Inf. Syst.*, aug 2022. ISSN 1046-8188. doi: 10.1145/3564285. URL `https://doi.org/10.1145/3564285`. Just Accepted. 3

[25] Yunqi Li, Hanxiong Chen, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. User-oriented fairness in recommendation. In *WWW '21: The Web Conference 2021*, pages 624–632. ACM / IW3C2, 2021. doi: 10.1145/3442381.3449866. URL `https://doi.org/10.1145/3442381.3449866`. 3

[26] Ashwathy Ashokan and Christian Haas. Fairness metrics and bias mitigation strategies for rating predictions. *Inf. Process. Manag.*, 58(5):102646, 2021. doi: 10.1016/j.ipm.2021.102646. URL `https://doi.org/10.1016/j.ipm.2021.102646`.

[27] Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. Recommendation independence. In *Conference on Fairness, Accountability and Transparency, FAT 2018*, volume 81, pages 187–201. PMLR, 2018. 3

[28] Tao Qin, Tie-Yan Liu, and Hang Li. A general approximation framework for direct optimization of information retrieval measures. *Inf. Retr.*, 13(4):375–397, 2010. doi: 10.1007/s10791-009-9124-x. URL `https://doi.org/10.1007/s10791-009-9124-x`. 3

[29] Hossein A. Rahmani, Mohammadmehdi Naghiaei, Mahdi Dehghan, and Mohammad Aliannejadi. Experiments on generalizability of user-oriented fairness in recommender systems. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2755–2764. ACM, 2022. doi: 10.1145/3477495.3531718. URL `https://doi.org/10.1145/3477495.3531718`. 3

[30] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Yushuo Chen, Xingyu Pan, Kaiyuan Li, Yujie Lu, Hui Wang, Changxin Tian, Yingqian Min, Zhichao Feng, Xinyan Fan, Xu Chen, Pengfei Wang, Wendi Ji, Yaliang Li, Xiaoling Wang, and Ji-Rong Wen. Recbole: Towards a unified, comprehensive and efficient framework for recommendation algorithms. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management*, pages 4653–4664. ACM, 2021. 3

[31] Rianne van den Berg, Thomas N. Kipf, and Max Welling. Graph convolutional matrix completion. *CoRR*, abs/1706.02263, 2017. 3

[32] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2019, Paris, France, July 21-25, 2019*. ACM, 2019. 3

[33] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yong-Dong Zhang, and Meng Wang. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, pages 639–648. ACM, 2020. doi: 10.1145/3397271.3401063. URL `https://doi.org/10.1145/3397271.3401063`. 3