Daniel Kandie

Abstract

This research focuses on multi-stakeholder fairness aware recommenders. That is, recommendation systems that must consider the welfare of multiple stakeholders other than just the end user in the recommendations made. This research focuses on the role that an initial algorithm has on the overall performance of fairness aware re-ranking and how the accuracy-fairness trade-off is affected as accuracy increases. It explores the evaluation of fairness from the perspective of the Black feminist theory of intersectionality. Overall the aim of this research is to inform the design of fairness-aware recommendation systems adopting a definition of fairness that is grounded in the theory of intersectionality.

Introduction

The continued digitization of day-to-day services has led to a rise in the use of recommender systems. These sift through vast troves of data to present users with a manageable list of relevant items. The key role that recommender systems play has therefore found them used in services ranging from web search to e-commerce to the growing "gig economy" (Burke, Sonboli, & Ordonez-Gauger, 2018). The increased use of recommender systems has also found them placed within situations where accuracy of recommendation results is not the only consideration as users are not the only stakeholders involved. This requirement has brought the need for 'multi-stakeholder' recommender systems that do not only consider users as the ultimate stakeholders of the recommendation results (Abdollahpouri, Burke, & Mobasher, 2017). Additionally, recommender systems have the need to, at best, ensure fairness in results and at least, prevent discrimination in their recommendation lists. This has therefore given rise to fairness-aware recommender systems which aim to incorporate fairness criteria within their recommendations (Sonboli, Eskandanian, Burke, Liu, & Mobasher, 2020). Placed together these two form multi-stakeholder fairness-aware recommender systems which is the focus of this paper.

The two main research questions that give structure to this research are:

'To what extent does the performance of fairness aware re-ranking depend on the initial recommendation algorithm?'

'To what extent does fairness aware re-ranking allow for the promotion of items with intersecting protected features?'

These questions therefore allow for a comparison of the interactions between both recommendation algorithms and re-ranking algorithms.

This research intends to make three key contributions. The first is a replication of prior studies on multistakeholder fairness-aware recommendation. The second is an exploration of the effect of various base recommender algorithms on the performance of re-ranking algorithms. Although there is a growing body of research on both fairness-aware recommender algorithms as well as re-ranking algorithms there is little research exploring the influence of various underlying algorithms on the performance of re-ranking algorithms. This research aims to contribute to this part of the growing body of literature on fairness-aware recommender algorithms. Finally, this research adopts an intersectional view of fairness when evaluating the performance of re-ranking algorithms in ensuring fairness. This contrasts the majority of the research which evaluates fairness on various criteria evaluated independently.

Intersectional Fairness

Fairness in a more general sense promotes the equality of people and the absence of discrimination. Within an algorithmic context, it is often interpreted in response to the question: 'are members of different groups who are of equal "merit" treated equally by the algorithm?' (Kasy & Abebe, 2021, p. 1). A development to the theory of fairness calls it to consider historical injustices against certain groups and thus equates fairness more to equity across individuals and groups rather than simply equality (Schelenz, 2021).

A key insight in the conceptualisation of fairness is made when it is understood through the theory of intersectionality. Intersectionality is a Black feminist theory, first proposed by Kimberlé Williams Crenshaw in 1989 that challenged the view of discrimination along a variety of dimensions as mutually exclusive (Crenshaw, Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics, 1989). Crenshaw, focusing on the lived experiences of Black women in the United States of America who experienced both racism as well as sexism, highlighted that the effect of the two "is greater than the sum of racism and sexism," and therefore any analysis into the discrimination faced by black women would be inadequate if it fails to account for the intersectionality of the two systems of oppression (Crenshaw, 1989, p. 140).

Crenshaw further elaborated on the role of intersectionality not as a 'new form of identity' but rather as a framework to "account for multiple grounds of identity when considering how the social world is constructed" (Crenshaw, 1991, p. 1245). Additionally, she made the distinction between 'Structural Intersectionality' referring to "race, gender and class" based discrimination, of which any social interventions only focused on one dimension of discrimination would be insufficient for someone implicated by all three. Additionally, 'Political Intersectionality' highlighted that "women of color are situated within at least two subordinated groups that frequently pursue conflicting political agendas" (Crenshaw, 1991, p. 1252) (Cooper, 2015).

As the core of this theory focuses on the effects of multiple systems of oppression on individuals it has gained widespread use beyond its initial context. For example, Intersectionality has been applied to population health research to understand health inequalities and design intervention strategies (Bauer, 2014). This theory has also been applied within an African context incorporating relevant identities such as social status, and class as well as historical contexts (Meer & Müller, 2017).

Within an algorithmic context the theory of intersectionality has been applied by Buolamwini and Gebru (2018) to evaluate the accuracy of computer vision algorithms. They found that in two widely used computer vision training datasets, 'IJB-A' a US government benchmark released by the National Institute of Standards and Technology and 'Adience' a gender classification benchmark, Black females were under-represented. Furthermore, when evaluating facial recognition classifiers, they found that all classifiers: performed better on male than female faces, difference in error rate of 8.1% - 20.6%; performed better on lighter faces than darker ones, difference in error rate of 11.8% - 19.2%; and worst on darker female faces, error rate 20% - 34.7%. This provided strong evidence showing that when left unchecked machine learning algorithms perpetuate forms of discrimination.

Within fairness-aware recommender systems, the application of an intersectional lens is still quite novel with most research applying a binary viewpoint of 'protected versus unprotected' categories when addressing matters of fairness. Additionally, Schelenz (2021) points out that, particularly for recommender systems that consider user characteristics within their recommendation, there is a need to contextualize the various user characteristics to understand them at a structural level rather than an individual level. Schelenz (2021) gives the hypothetical example of a recommender system for documentaries that considers the educational level of users; in the absence of contextualising the role of race and gender disparities in access to education, the recommender system may privilege certain groups of individuals in recommending educational documentaries.

However, some research that applies an intersectional approach to fairness-aware recommendation is proposed by Sonboli et al (2020) who present an approach (focused on item characteristics and user-item interaction in recommendations) which assigns relative weights to various item categories (designating some categories as 'protected' based on a history of underrepresentation in prior recommendation lists) and allows the application of multiple categories. Although this work does not contextualize the categories of each item, its key contribution is presenting an approach that considers the interaction of various categories. Furthermore, although the method allows for items at the intersection of multiple protected categories to be promoted, this is not measured or evidenced within the paper.

Research has also focused on a variety of metrics to measure intersectional fairness. Ghosh et al (2021) propose a worst-case scenario metric that can be used for measuring the performance of fairness-aware recommender algorithms across intersecting protected groups whilst building upon previous measures of recommender fairness. Mary et al. (2019) present an approach that makes use of the Rényi maximum correlation coefficient to consider fairness aware criteria as continuous variables allowing for the intersectionality of various groups. Such research is key as it opens existing recommender systems to further scrutiny on the level of intersectional fairness that they ensure as well as opening the door to further research that seeks to develop fairness-aware algorithms focused on ensuring intersectional fairness.

Recommender Systems

Fairness aware multi-stakeholder recommendation systems generally fall into three categories of fairness: consumer side fairness, producer side fairness or both (Sonboli, Eskandanian, Burke, Liu, & Mobasher, 2020). Consumer side fairness pertains to fairness considerations regarding an individual consumers recommendation list. An illustrative example of this is with job search and recommendation engines which must balance the aims of providing a personalised recommendation to users whilst ensuring that they do not unfairly recommend jobs to different users with the same qualifications based on factors such as gender or race, i.e., they should not consistently recommend high-paying jobs to men whilst recommending lower paying jobs to women.

On the other hand, provider side fairness has to do with fairness concerns pertaining to the overall set of recommendations made by the provider. Fairness in this regard is therefore less concerned with an individual's personal recommendation list but rather that the set of items are recommended in a fair and balanced way. Within the job search example, provider side fairness would be the requirement that the job search engine recommends each category of job in an equitable way and does not over recommend one type of job whilst under-recommending another. This may however be difficult to enforce as various companies may pay to have their job adverts promoted in job search results.

At the intersection of these two is the application of both consumer and producer side fairness. which exists in scenarios where both factors must be taken into consideration when designing a recommender system.

This research intends to compare the effectiveness of a variety of fairness-aware recommendation algorithms against metrics that highlight the intersectional definition of fairness.

References

Abdollahpouri, H., Burke, R., & Mobasher, B. (2017, July). Recommender Systems as Multistakeholder Environments. *UMAP '17: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, 347–348. doi:10.1145/3079628.3079657

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. *Social Science & Medicine*, 10-17. Retrieved from https://doi.org/10.1016/j.socscimed.2014.03.022

Burke, R., Sonboli, N., & Ordonez-Gauger, A. (2018). Balanced Neighborhoods for Multi-sided Fairness in Recommendation. *Proceedings of Machine Learning Research*, 1-13.

Cooper, B. (2015). Intersectionality. In L. Disch, & M. Hawkesworth, *The Oxford Handbook of Feminist Theory.* Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199328581.013.20

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum, 1989*(1). Retrieved from http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics and Violence against Women of Color. *Stanford Law Review, 43*(6), 1241-1299. doi:10.2307/1229039

Ghosh, A., Genuit, L., & Reagan, M. (2021). Characterizing Intersectional Group Fairness with Worst-Case Comparisons. *Proceedings of Machine Learning Research - Workshop on Diversity in Artificial Intelligence: Diversity, Belonging, Equity, and Inclusion*, 1-12. Retrieved from https://arxiv.org/abs/2101.01673

Joy Buolamwini, T. G. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Conference on fairness, accountability and transparency, Proceedings of Machine Learning Research*, 77-91.

Kasy, M., & Abebe, R. (2021). Fairness, Equality, and Power in Algorithmic Decision-Making. *Conference on Fairness, Accountability, and Transparency (FAccT '21)*. doi:10.1145/3442188.3445919

Liu, W., & Burke, R. (2018). Personalizing Fairness-aware Re-ranking. *arXiv preprint arXiv:1809.02921*.

Mary, J., Calauzènes, C., & Karoui, N. E. (2019). Fairness-Aware Learning for Continuous Attributes and Treatments. *Proceedings of the 36th International Conference on Machine Learning PMLR*, 4382-4391.

Meer, T., & Müller, A. (2017). Considering intersectionality in Africa. *Agenda*, 3-4. doi:10.1080/10130950.2017.1363583

Schelenz, L. (2021). Diversity-aware Recommendations for Social Justice? Exploring User Diversity and Fairness in Recommender Systems. *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 404-410. doi:10.1145/3450614.3463293

Sonboli, N., Burke, R., Mattei, N., Eskandanian, F., & Gao, T. (2020). “And the Winner Is...”: Dynamic Lotteries for Multi-group Fairness-Aware Recommendation. *arXiv preprint arXiv:2009.02590*.

Sonboli, N., Eskandanian, F., Burke, R., Liu, W., & Mobasher, B. (2020). Opportunistic Multi-aspect Fairness through Personalized Re-ranking. *ACM Conference on User Modeling, Adaptation and Personalization, 28*, 14–17. doi:10.1145/3340631.3394846