# Two-Sided Fairness in Non-Personalised Recommendations (Student Abstract)

**Aadi Swadipto Mondal**[*]**, Rakesh Bal**[*]**, Sayan Sinha**[*]**, Gourab K Patro**

Indian Institute of Technology Kharagpur, India - 721302.

{aadismondal, rakesh.bal, sayan.sinha, patrogourab}@iitkgp.ac.in

## Abstract

Recommender systems are one of the most widely used services on several online platforms to suggest potential items to the end-users. These services often use different machine learning techniques for which fairness is a concerning factor, especially when the downstream services have the ability to cause social ramifications. Thus, focusing on the non-personalised (global) recommendations in news media platforms (e.g., top-$k$ trending topics on Twitter, top-$k$ news on a news platform, etc.), we discuss on two specific fairness concerns together (traditionally studied separately)—user fairness (Chakraborty et al. 2019) and organisational fairness (Burke et al. 2020). While user fairness captures the idea of representing the choices of all the individual users in the case of global recommendations, organisational fairness tries to ensure politically/ideologically balanced recommendation sets. This makes user fairness a user-side requirement and organisational fairness a platform-side requirement. For user fairness, we test with methods from social choice theory, i.e., various voting rules known to better represent user choices in their results. Even in our application of voting rules to the recommendation setup, we observe high user satisfaction scores (table 2). Now for organisational fairness, we propose a bias metric (eq-1) which measures the aggregate ideological bias of a recommended set of items (articles). Analysing the results obtained from voting rule-based recommendation, we find that while the well-known voting rules are better from the user side, they show high bias values and clearly not suitable for organisational requirements of the platforms. Thus, there is a need to build an encompassing mechanism by cohesively bridging ideas of user fairness and organisational fairness. In this abstract paper, we intend to frame the elementary ideas along with the clear motivation behind the requirement of such a mechanism.

## Two-Sided Fairness

**Dataset** The dataset used for this research is a publicly available news dataset that includes news articles (in Norwegian) in connection with anonymised users (Gulla et al. 2017). From this dataset, we use the 865 articles with publicly available news contents and the news consumption history (how much time a user spends reading different articles) of 63260 anonymised users.

---

[*]The first three authors have contributed equally.

**Non-personalised Recommendation** Let $\mathcal{U}$ and $\mathcal{A}$ be the set of users and articles (items) respectively. Most of the recommender systems rely on some form of scoring for user-item pairs (e.g., ratings). Here we derive such scores from the "Active Time" of a user on an article (the total amount of time in seconds the user spends on the article). We derive the score for user $u \in \mathcal{U}$ and article $a \in \mathcal{A}$ pair as: $Score(u, a) = \dfrac{ActTime(u, a)}{\sum_{u' \in \mathcal{U}} ActTime(u', a)} \cdot N_a$. Here $ActTime(u, a)$ is the total active time of the user $u$ on article $a$ and $N_a$ is the number of users who have viewed the article at least once. Next, we linearly scale all the scores of each user separately to floating-point numbers in the range (1 to 10), in order to ensure similar scales for fast and slow readers. This choice of scores is purely dataset-specific and can be replaced with other suitable score metric as per the requirement. Here, a higher score signifies more affinity of a user towards the article. However, there are many user-article pairs for which a score does not exist, owing to the fact that the user might not have viewed that particular article. We find these missing values using Non-Negative Matrix Factorisation (NMF) (Koren, Bell, and Volinsky 2009).

**Non-negative Matrix Factorisation (NMF):** We obtain a sparse matrix $V$ from the normalised scaled scores obtained earlier. The rows of $V$ represent the users, and the columns represent the articles. We then perform NMF on $V$ and obtain $\hat{V}$ which provides us with the predicted scores for the missing values in $V$. $V^*$ denotes the final score matrix, which comprises the original scores from $V$, as well as the predicted scores from $\hat{V}$ (for the missing entries in $V$). NMF was set on training with hyperparameters $\alpha = 0.0002$, and $\beta = 0.02$. Convergence was assumed when $\Delta cost < 0.001$ (around $2,500$ epochs).

**The Recommendation:** As we talk about non-personalised recommendation, the goal is to find $\kappa$-sized recommendation which would be liked by most of the users. For this purpose, it is wiser to use $V^*$ as it—being a complete score matrix—can provide insights on all the user-article pairs. Next, we discuss two kinds of fairness desirable in such recommendations and also the obstacles in achieving them.

**User Fairness** As the recommendation is global, it should fairly represent the interests of the users on the platform. Thus, we can re-purpose and use multi-winner voting rules

| Left-biased | motstandernes, djevlene, motstander, problem, motstand, scorer, møte, sjef, kontret, balanse, maktet, rød, igjen |
|---|---|
| **Right-biased** | bakgrunnen, grunn, etablerte, opprettelsen, døpt, alternativ, offensivt, defensiv, renter, tilværelsen, høyre, erna, siv |

Table 1: Some of the seed words having high PMI values

in our setup as voting rules are known to fairly represent voters' choices while choosing winners. To map our setup to an election we do the following; election $\mathcal{E} = (\mathcal{C}, \mathcal{V})$, where candidate set $\mathcal{C} = \mathcal{A} = (a_1, ..., a_m)$ is the set of articles, voter set $\mathcal{V} = \mathcal{U} = (u_1, ..., u_n)$ is the set of users, and we consider scores $V^*$ as the corresponding votes. The winners of the election are denoted by the set $\mathcal{W} = (w_1, ..., w_\kappa)$ which is a subset of $\mathcal{C}$ and are elected using various voting rules. In this work, we experimented with six famous voting rules, namely SNTV, STV, $k$-Borda, Bloc voting, Chamberlin Courant and Monroe (Elkind et al. 2017). Next, we find the average satisfaction of all users for the results obtained from each election method like: $Satisfaction = \frac{1}{\mathcal{U}} \sum_{u \in \mathcal{U}} \frac{\mathcal{W} \cap Top\text{-}\kappa_u}{\kappa}$. Here $Top\text{-}\kappa_u$ is the top $\kappa$ articles for user $u$ as per the scores in $V^*$. The satisfaction score for each election method has been presented in Table 2 column 2. A high satisfaction score would ensure that the recommendation results (election winners) hence obtained follow the general principle of user fairness. In all our experiments, the value of $\kappa$ has been taken as 10.

**Organisational Fairness** For organisational fairness (Burke et al. 2020), the set of recommended articles needs to be balanced in terms of political/ideological bias. Before finding political bias of the articles in our dataset, first, we analyse articles from politically-biased media houses (https://tinyurl.com/y2t7bq76) – namely, *Klassekampen* (left) and *Aftenposten* (right), as well as articles talking about eminent people involved in politics pertaining to either of these ideologies. Using co-occurrence analysis of words in the articles of biased media houses, we shortlist some of the words with the highest Pointwise Mutual Information (PMI) scores. Some of the words having high PMI values—which we later use as seed words—have been mentioned in Table 1. We consider that their presence in an article indicates a higher likelihood of it being politically biased. Thus, we designate the articles in our dataset as either left-biased or right-biased based on whether they contain the majority of seed words from the left category or right category. Next, we find the reference bias ($\rho$) of the user population on the platform as the ratio of total time spent on the left-biased articles to that on the right-biased articles[1]. Finally, we design a bias metric [2] which measures how far are the global recommendations from reference $\rho$.

$$Bias = \frac{-leftCount + \rho * rightCount}{leftCount + \rho * rightCount} \quad (1)$$

Here $leftCount$ and $rightCount$ are the number of occurrences of left and right biased seed words in a set of articles. Hence, $Bias < 0$ indicates left-bias and $Bias > 0$ indicates right-bias. Closer the value of $Bias$ to zero, better is

the method in terms of satisfying organisation fairness. The range of $Bias$ is [-1, 1]. We calculated the bias values of the global recommendations obtained earlier using voting rules. The results have been shown in Table 2 column 3.

| Election Method | Satisfaction | Bias |
|---|---|---|
| SNTV | 0.878 | -0.115 |
| $k$-Borda | 0.800 | 0.033 |
| Bloc | 0.856 | -0.240 |
| STV | 0.894 | -0.117 |
| CC | 0.883 | -0.169 |
| Monroe | 0.825 | -0.103 |

Table 2: User Satisfaction and Organisational Bias

**Discussion** For better user fairness, we repurposed well-known voting rules from fair voting theory, and implemented the same in our recommendation setup. Our evaluation shows that they achieve good user satisfaction scores (column 2 of table 2). However, they show high bias values (column 3 of table 2). This is happening as the voting rules do not consider the article content while choosing the global recommendations. On the other hand, organisational fairness—relating to balance in the content distribution—requires the bias to be close to zero. Therefore, there is a need for designing a mechanism which simultaneously cares for user fairness and organisational fairness while designing non-personalised recommendations. All of our resources can be found at https://github.com/americast/electoral.

## References

Burke, R.; Voida, A.; Mattei, N.; and Sonboli, N. 2020. Algorithmic Fairness, Institutional Logics, and Social Choice. In *Harvard CRCS Workshop: AI for Social Good*.

Chakraborty, A.; Patro, G. K.; Ganguly, N.; Gummadi, K. P.; and Loiseau, P. 2019. Equality of voice: Towards fair representation in crowdsourced top-k recommendations. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 129–138.

Elkind, E.; Faliszewski, P.; Skowron, P.; and Slinko, A. 2017. Properties of multiwinner voting rules. *Social Choice and Welfare* 48(3): 599–632.

Gulla, J. A.; Zhang, L.; Liu, P.; Özgöbek, Ö.; and Su, X. 2017. The Adressa dataset for news recommendation. In *Proceedings of the international conference on web intelligence*, 1042–1048.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* 42(8): 30–37.

---

[1]In our experiments, $\rho$ was found to be $1.423$.

[2]More sophisticated methods for bias analysis can also be used with the availability of relevant annotated resources in Norwegian. However this is not under the purview of this abstract.