Does Fair Ranking Improve Minority Outcomes? Understanding the Interplay of Human and Algorithmic Biases in Online Hiring

Tom Sühr, Sophie Hilgard and Himabindu Lakkaraju





AAAI / ACM conference on ARTIFICIAL INTELLIGENCE ETHICS, AND SOCIETY



Hiring platforms affect people's livelihood

- Hiring platforms connect employers and job candidates
- Prior research demonstrated undesirable algorithmic and behavioral biases
- Fair ranking algorithms were proposed to mitigate those biases
- Evaluation of fair ranking by websearch based clickmodels

But does fair ranking improve the actual outcome of underrepresented groups on hiring platforms?



We created a hiring simulation

- We gathered three datasets from TaskRabbit
- Each dataset consisted of 3 female and 7 male candidates
- We used TaskRabbit's ranking, LinkedIn's Fair Det-Greedy and a random ranking to sort the job candidates
- We tested all algorithms in three different job context

D1	D2	D3
Task Rabbit	FairDet-Greedy	Random
Moving Assistance	Event Staffing	Shopping
Task Rabbit F ← → M	FairDet-Greedy F↔M	Random F↔M

Is fair ranking equally effective in all job contexts and candidate pools? Is the effectiveness of fair ranking dependent on which group is underrepresented?

Controlling for job context, algorithm and data

TaskRabbit	Randomized order of Job
	Contexts
FairDat Graady	1

- Participants were briefed and matched to one algorithm
- Latin-square design for job-context and candidate pool distribution
- Each candidate had to select her four top candidates for each task
- Free text field after the hiring task to describe decision making

Recruiting study participants

- We recruited 1079 participants from Amazon MTurk
- All participants were recruited in the US with at least 5000 approved tasks and 95% approval rate







User Interface Design



1 Selection



Fair ranking improves minority outcomes



Result summary

ranking Fair Fair Random TaskRabbit The effectiveness differs with job context and candidate profiles

