

Researching Fairness In Resume Screening AI

Moulik Jain

Background

- In recent years, the number of companies using AI/ML resume checkers to see if someone is a good fit for the job has risen dramatically.
- With online job postings in such high demand a resume being denied by these models may never even be seen by a person.
- Details like someone's gender, race, school or even just the words they use in their resume could possibly affect the score that the model gives them for the job.
- For our project we set out to test a specific example of one of these resume scorers and determine if there is a bias in any way.

Prior Work

- In 2018 it was found that Amazons Hiring tool was biased towards men, they then discreetly dropped the program completely.
- There have been multiple articles about historical bias, showing how some models favored male candidates for technical roles.



Introduction

- For this project I will be testing ResumeScreening.AI a website application that allows the user to input a job description, and resumes.
- The website will then assign similarity scores for each resume.
- On their website they have mentioned that some larger companies have implemented their services the main one being a Crypto website.
- While we don't know if this service is similar to others used in the industry it provides a solid starting point for analysis.

Data

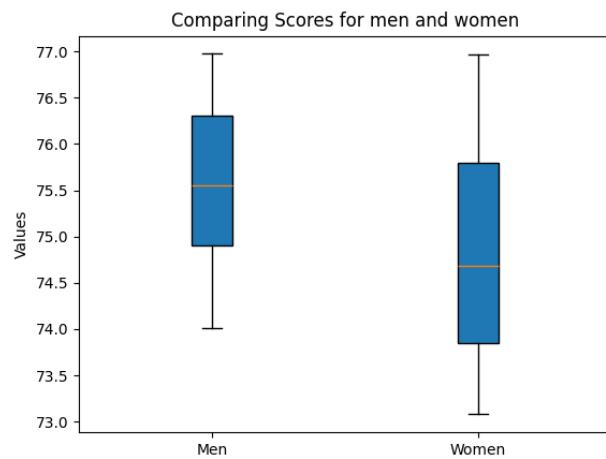
- My first testing started with 20 assumed White names and 20 assumed African American names.
- Some names were made by combing common first and last names and others were simply found online.
- The final test looking at common female names was done on a 50/50 subsection of a two Kaggle datasets. This subsection had 40 from both races. This data was found on an inmate registry and 40 samples were randomly selected.

Methodology

- Because of some initial confusion on the API available, I ended up having to check the resumes semi manually.
- I used a script that took in a sample resume and would change only the sample name (John Doe) to names that are typically assumed to either be white or African American males.
- These names were all found on the internet by either combining popular first names with last names, or just copying and pasting full names.
- The first test consisted of 40 total resumes and the second consisted of 80. Classes were balanced equally for both. This was done for both male and female names.

Results

White Male Avg	African American Avg	White Female Avg	African American Female Avg
75.20%	75.14%	74.740%	74.920%



Results Cont

- The scores only varied around .03 score for every resume.
- Overall, there wasn't a large disparity between scores for races.
- I then tested with 80 total names, even class split the range stayed around .05

Avg White	Avg African American
0.754615385	0.752

- While there was some disparity it wasn't enough to constitute bias

Average for White women	Average for African American Women
0.746666667	0.74975

Other/Future Work

- This presentation showed the partial implementation for testing this website for the final implementation I hope to make some changes.
- Using Kaggle I found some larger datasets where a correlation might make itself more known. The issue is the way that I am currently making and checking the resumes relies on storing them on my computer and manually uploading them.
- While this works for the smaller datasets if I wanted to expand to hundreds or thousands of names I would need to work on a different methodology, maybe stopping the program to clear out and check the resumes and combining the results at the end.
- A key task for future work is going to be finding a matching Mens inmate set to test on and see if gender bias is truly present.

Conclusion

- By working through this basic methodology, I didn't find any good indicators of racial bias. This being said it was slightly confusing to me that some names inherently (regardless of race) scored better than others.
- When testing using the women's dataset from Kaggle it was found that on average women scored around 0-1% lower than men. While this isn't a key indicator it does warrant further research with a larger dataset.
- It should also be noted that the source of the names for men and women was different. Men being from common lists of names and women being from the inmate dataset on kaggle.

Thank You