

Introduction

- I like how you construct the Introduction part, like you're using subtitles to make the structure more clear and readable.
- The first paragraph provides a solid background on gender bias in hiring, and the motivation is really strong.
- It would be helpful to define or clarify some key words earlier in the background part when you first mention it, like gender-coded language, or at least give some examples of it.
- Idk if you cite any previous literature in the background paragraph. You can include some in-text citations if so. Otherwise, including one or two citations directly in the introduction would strengthen the academic framing and show awareness of existing work immediately.
- The research questions are clearly stated.

Data

- It's clear how the data was collected and the source is stated.
- You can talk about the size of the dataset and number of variables in part 1.2.
- I feel like you can specify the date or time range when scraping occurred. I don't know how the website works but I feel like online resume content can change over time.
- For data cleaning, make sure to include a sentence saying that you dropped how many NAs and how many records are left.

Methods

- One quick comment on visualizations: Figure "Distribution of Years of Experience by Gender" has no x and y labels. Make sure you add those.
- For the scoring index, can you do some research to figure out what are the optimal weights for each term?
- How did you evaluate your method's effectiveness and reliability? Maybe try to implement some accuracy checks statistically?

Results

- I like that you compared distributions across gender before moving into the ranking outcomes. One suggestion is to make the connection between each descriptive figure and the research questions slightly clearer, like when presenting differences in masculine-coded language, you can directly note whether and how that feature eventually affects the ranking outputs. Add one or two transition sentences like "This figure supports RQ1 by showing that ..." or "These differences in masculine-coded wording help explain why male-coded résumés get ranked higher."

Discussion

- Nice job on summarizing! It's very clear what the main takeaway is.
- Solid interpretation.
- Maybe separate what affects the accuracy of your findings vs. what affects how far they generalize? Some of these limitations affect the reliability of the patterns we see in this specific dataset, while others mainly limit whether these patterns would hold in other fields or populations.
- I like the reflection on how the dataset structure shapes the model's behavior. One small suggestion is to explicitly connect this to why commercial systems might behave similarly, just to emphasize the broader relevance.
- Nice ethical discussion!

General / Code Stuff

- Good format
- Couple of typos need to be corrected (1 in Model section)
- Empty cell block?
- Comments are very organized → could put comment blocks at the top instead of if (blah blah) # it does this thing
- I could run this on my mac so maybe making it able to be ran on mac? Idk why it was saying something about linux
- Emphasizing how the definition of gender changes the scale is super important – good job
- Being transparent on the ethical concerns of using real applicants is also super crucial in having a good analysis.
- This was super well done and in depth – Tone also seemed super neutral which is great it didn't sound like you were leaning down any avenue and were being very neutral.