

Bias in Resume Screening Algorithms

Kevin Pham

Halicioğlu Data Science Institute

DSC 267R: Data Fairness and Ethics

David Danks

June 13, 2023

A. Explain the goal of your data science effort

With the imminent fear of a recession and the mass layoffs happening across the country in different industries, the job market has become increasingly challenging, especially to new graduates who are aiming to enter the workforce. These incoming graduates now have to compete with not only their peers, but also highly seasoned professionals, many of which have still been seeking employment for the last couple of years. With the significant number of candidates applying for jobs, companies need to be able to effectively and efficiently root out candidates during the resume screening process to determine who they want to move on to the next step of the hiring process. As a result of this, many companies opt to develop or use resume screening algorithms in order to whittle down the number of applicants to only those that are deemed competent for the job. The challenge with this is that there may be inherent bias within the algorithm used to screen resumes leading to unfair treatment for different groups of people.

B. Outline a proposed data collection and measurement system for this effort

To begin addressing this challenge, we will first be gathering pre-collected datasets from different open data source websites such as Kaggle and Data World. These datasets would already have information on candidates such as their gender, ethnicity, school major, prior work experience and so on. It is important to have a large data library for the algorithm to refer to. Additionally, depending on the industry and field, we will need more data especially for similar jobs to cross reference [Appendix A]. If the datasets for specific industries are inaccessible, we could default to our own data collection and measurement system. We would use surveys to

collect key topics from a resume such as, but not limited to: gender, ethnicity, school major, prior work experience, class curriculum, etc; essentially looking for all information that may contribute to the person's success in the potential role. Lastly, for each applicant, we would document each of their answers into a field for a csv and add this to our data pool.

C. Identify at least two significant ethical and/or societal concerns about this collection & measurement system

When collecting data to feed to the algorithm, it is important to be wary of the potential harms of using the datasets. In the case that your algorithm is biased, it will cause your predictions to be incorrect and cause harm to the respective affected groups. One significant concern is that using old employment data may lead to racial or gender biases. Given that in the past, there was a higher proportion and presence of specific genders and ethnicities in certain job fields, when screening through resumes, the algorithm may think that other genders or ethnicities are not as desirable or successful in these jobs. As such, the algorithm will become biased towards these parties and omit them from potentially being able to land a job. Another significant concern is transparency. In order for people to use the results of any algorithm, they need to be aware of how the data used was collected. If the data was collected unethically, how can people reliably use this information in order to generalize results? The users and anyone who views the results need to be kept informed and continually aware of any changes to the datasets so that they are not left wondering if there is any bias. The last concern with these datasets is the lack of privacy and security. When applying for jobs, you often divulge a lot of private information such as your home address, email, phone number, etc. In these cases, there is a concern that such

information may be shared or leaked without your consent. Worst case scenario is that your information can be used by someone else for an unintended purpose.

D. Provides responses to the concerns raised in C

In order to circumvent the concerns raised above, it is imperative that the datasets are diversified and representative. The datasets should, for example, consist of information from different industries/fields and a wide range of ethnicities and genders. From this, the algorithm will have a wider scope of data to refer to and the results will be more generalizable and less likely to have bias. In addition, there should be more examination of the datasets, including audits and cross checks to make sure that the data is impartial and there are no causes for concern. To address the issue of transparency, all of the methodology for the collection and use of the data should be publicly available, for example on a social media platform. There should be regular updates to the platform so that the public is constantly kept informed of the latest changes along with documentation to explain the changes. Finally to address the concern of privacy, all of the data collected should be handled with utmost confidentiality as possible. The data collection will only collect essential information and retain anonymity for the sources of data. The data accessibility should be limited to those working with the datasets and the data retention periods should be established from the start, with the individuals having the ability to access and withdraw any of their saved data at any point.

E. Identify at least two different potential ethical concerns about analyzing/modeling these data for that goal

While there are many concerns about data collection, you also need to be mindful during the analysis and data modeling process. One key idea to remember is that most of the data you are working with is historical data. The most intricate and detailed data collection can only explain the past or leading up to the present, and is not entirely accurate when it comes to predicting the future. You can not just use this data and think that the results are going to be completely relevant and precise. Similarly, not all jobs will have the same amount of employment data, some fields will be lacking information. The job market is constantly changing, with new fields consistently rising, and new criteria for employment. In the cases where there is not a lot of information available for an industry or specific job, how can we reliably use this data to determine whether someone will get employed?

F. Explain how you would respond to the concerns in E. through choices in the analysis phase

Given the heavy presence of historical data, there are several steps we can take during the analysis process in order to better represent the data so that it is generalizable for the future. One step is to examine the data and determine the most important features in the data that can be attributed to being successfully employed. From there, we can assign weights to make the key features more impactful for the algorithm. We can also assign heavier weights to datasets that are more recent in order for our algorithm to see that the more recent data is most important. As for new jobs, we can implement code that will identify the similarities in the positions of the new jobs as compared to older jobs and see what features are present in both. While it may not

perfectly categorize the likelihood of getting a job in a newer industry, it will provide similar results based on the key features identified in the analysis phase.

G. Describe appropriate and inappropriate contexts-of-use for the model

Once the model with the algorithm has been developed, there are many appropriate situations where the model can be utilized. For starters, the model can be used first and foremost by recruiters and others who are actively dealing with resumes. Most companies have a platform or email where all of the job applications are sent to. Companies or employees can then apply the model in order to comb through the data and highlight resumes or applicants that stand out and fit their hiring criteria. Another way the model can be used is through a company employee or government employee auditing the model. Before and after the model is put to real-world application, it needs to be audited to look for biases or issues. You don't want to run into issues because you did not check the model and it turns out your model was discriminatory or faulty. As such, companies should focus an internal department to double check their model and consistently make sure the model works as intended. Similarly, the government can have some sort of auditing department that is able to check companies' hiring models such as when someone reports that there was some sort of unfair hiring process. In the case of this, the government employee would be determining whether the model worked as intended or had separate issues. The model can also be used in a teaching or informative manner once released to the public. For example, the model can be used by people who are actively searching for jobs. They can utilize a publicized model and dataset to determine whether their resume or capabilities would be flagged as an interest for a company. In this way, the model would be able to provide reassurance or

foresight on how their applications would go. Another example is that the model can be used as a case study by researchers or students in an academic setting. Teachers can use the model as learning material for students to familiarize themselves with as assessments or projects. Researchers can also use the model to conduct replications or as material to create something better.

In defining many appropriate approaches to using the model, there are also many inappropriate uses of the model. While the model does result in predictions that can be generalizable, it does not mean that this model has a direct causal relationship. One cannot use this model as the single source of validity in affirming that because they were predicted to pass the screening, they will pass the screening. It is pivotal to understand the model and its limitations before utilizing the model. Additionally, you cannot use the model without continually updating the datasets to supply more information on the current job market. As mentioned previously, this model will often use lots of historical data. As such, it is not always accurately reflective of the current market or positions that are in demand. In order for the model to stay impactful and up to date, it needs to be regularly updated for the latest changes so that it will have higher chances of being precise and relevant. As new sources of data and information arise, these sources will need to be examined for bias and then added to the current data pool for integration.

H. Explain how to interpret the model's outputs, perhaps in a context-sensitive manner

Once the model runs, the user needs to be able to properly interpret the results. The model utilizes the inputted information in order to produce predictions on whether a person will

be successful in the applied position based on their resume. This prediction result is not guaranteed in real-world applications, meaning that if it predicts a person's resume passes, it does not mean they will be hired for a job. Or vice versa, if a person's resume does not pass, they will be rejected from every other job. There are a plethora of other factors that can affect an applicant's hiring status and this model's output is merely a guiding tool to narrow down the excessive lists of candidates with the metrics deemed by the company. However, you cannot completely rely on the algorithm and trust it blindly. At some point, there should be human input and judgment in order to double check and reaffirm the results of the predictions. While it is far-fetched for a recruiter to sit and go through all of the applications, the model outputs can serve as a tool to scan through a majority and the recruiter can have the last glance. Doing so, will help to expedite the hiring process and root out inapplicable applicants. At the end of the day, it is the applicant's interactions with the hiring staff and display of knowledge that will have the greatest impact on the hiring result.

I. Identify and explains at least two ways in which proper use of the model could result in increased justice in appropriate contexts

With proper use of the outputs from the model, justice can be administered in a variety of appropriate contexts. This model uses historical and most recent data, along with a variety of objective factors in order to predict whether an applicant's resume passes. In having a model determine the success or failure, there is a reduced risk of intentional bias when reviewing resumes. Because the model is as good as the data it is inputted with, as long as the data is free of bias, the model can reliably generate accurate and fair results. In this way, every applicant is

viewed objectively and identically for key metrics that define success. Additionally, as mentioned earlier, with the significant amount of applications to job postings everyday, it is an extreme task to have recruiters have to manually go through each and every applicant. This takes a considerable amount of time and effort that could be spent on other tasks. By using this model, recruiters are able to significantly decrease the screening workload and complete the hiring process faster, saving time and effort and getting back to the applicants sooner rather than later.

J. Describe and explain at least two potential situations—contexts, uses, etc.—in which the model should be updated or revised (perhaps no longer using it) for ethical reasons

Despite the potential quality of the model, there will always be cases in which the model needs to be updated or revised. First and foremost is when the data being used is outdated or no longer reflective of the present. This can occur if there have been tremendous leaps in a certain field and the skillsets desired have changed or if a certain field is beginning to phase out. In these cases, it is extremely important that the data pool gets updated as soon as possible so that the model's outputs can be viewed reliably. If the data is not updated, then the model may inaccurately move an applicant to the next step in the hiring process when they are unqualified. Another situation in which the model should be revised is if the data being stored for longer than defined in its retention period. In these cases, the protocols set in place for the data need to be revisited and if determined that there is a need for new data, the old data needs to be disposed of as per regulations.

Appendices

A. Data Sources

Data will be obtained through websites such as Kaggle and be used for cross referencing such as the dataset below.

Data Analyst Jobs

Data Card Code (132) Discussion (3)

▲ 529

New Notebook

Abstract

Looking for a job as Data Analyst? Maybe this dataset can help you.

About this dataset

Amidst the pandemic many people lost their jobs, with this dataset it is possible to hone the job search so that more people in need can find employment.

This dataset was created by [picklesueat](#) and contains more than **2000 job listing for data analyst** positions, with features such as:

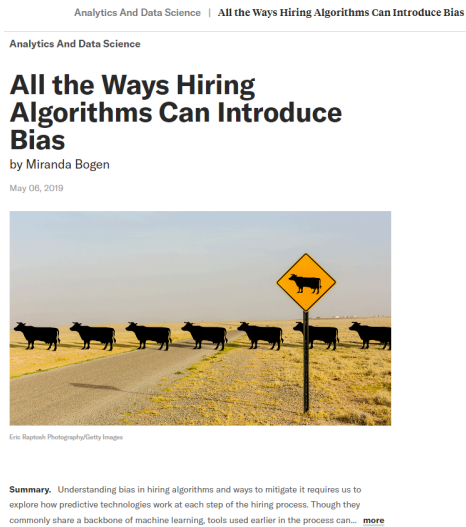
- Salary Estimate
- Location
- Company Rating
- Job Description
- and more.

How to use

- Find the best jobs by salary and company rating
- Explore skills required in job descriptions
- Predict salary based on industry, location, company revenue
- Your kernel can be featured here!
- [Data Engineer Jobs](#)
- [Business Analyst Jobs](#)
- [Data Scientist Jobs](#)
- [More Datasets](#)

References

Bogen, Miranda. “All the Ways Hiring Algorithms Can Introduce Bias.” *Harvard Business Review*, 6 May 2019, hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias.



Raghavan, Manish, and Solon Barocas. “Challenges for Mitigating Bias in Algorithmic Hiring.” Brookings, 6 Dec. 2019, www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/.

