# The Use of Artificial Intelligence in Hiring and Recruitment

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

**Liam Tolbert**

Spring 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Pedro A. P. Francisco, Department of Engineering and Society

# STS Research Paper

## Introduction

In the current landscape dominated by the proliferation of AI technologies, businesses worldwide are in a race to integrate these advancements into their operations. This trend is especially pronounced in the domain of hiring and recruitment. With the promise of efficiency and objectivity and as they seek to optimize their talent acquisition processes, they increasingly turn to AI and machine learning algorithms to streamline candidate screening and selection. However, beneath the surface lies a sea of ethical challenges that demand careful consideration.

My STS research investigated what potential problems arise in employing artificial intelligence and machine learning hiring algorithms in the recruitment process. It seeked to address the potential problems inherent in these technologies, particularly regarding fairness and bias. Despite their intended purpose of optimizing decision-making processes, AI algorithms are not immune to the biases present in the data they are trained on or the complexities of their design. When data scientists do not catch these biases, the algorithms can have unintended affects or otherwise display unpredictable behavior. As such, when these algorithms are implemented in ways that affect large populations, understanding their issues goes beyond an algorithm's mere technical effectiveness and extends to their broader societal impact.

To provide context, it is necessary to outline the broader landscape of AI and its application in hiring. Artificial intelligence, encompassing machine learning and deep learning, refers to the simulation of human intelligence by an algorithm. These algorithms can analyze data, identify patterns, and make decisions with minimal human intervention. In recent times, entities such as OpenAI, Google, and Microsoft have been pushing the envelope in terms of

innovation on their AI algorithms, developing hugely innovative applications such as ChatGPT, Copilot, and Gemini. These applications, among others, can answer questions, generate images, use logical reasoning, and write code. They are fundamentally changing aspects of our digital culture, generating huge amounts of controversy, and are arguably the face of AI today.

However, most of what we call artificial intelligence is what would be instead classified as machine learning. Machine learning algorithms are software that are designed to "learn" using sophisticated probability models. By feeding the model globs of pre-labeled data, a data scientist can make the model reasonably assume certain features are associated with certain outcomes based on a probability distribution. After training the model with the pre-labelled data, the model is let loose on unlabeled data to validate and test its effectiveness at predicting certain outcomes based on features present in the data that it had learned of before.

For example, if a data scientist wanted to create a machine learning model to classify images as either a dog or a cat, they would feed the model images of dogs and cats labelled as "dog" or "cat". After the model had finished training, they would give the model some additional unlabeled pictures. If the model identifies enough pictures as the correct animal, it is successful. If it doesn't, then some tweaking of the model's math is in order and the process starts again. By using this general example of machine learning, it is quite simple for us to see how it can be applied to hiring. All a machine learning model needs to do is associate certain features with good hires. Then, it will be able to successfully predict, perhaps based on someone's resume or other application details, if they are suitable for the job. While there are many, many features to look for, that is exactly what an automated system is for. Then, presumably, time and money are saved.

Corporations have already observed this connection. According to an article by Andrews & Butcher (2022), as companies scramble to implement machine learning into their recruitment operations, vendors are advertising automated video interview systems, resume filtering algorithms, and even video game tests as applicant screening (p. 148). In their study, the pair claim that marketing for those algorithms suggest they are time effective, cost effective, and accurate in their selection for new hires; they are even touted by their vendors as an impartial way to screen applicants without any form of discrimination (p. 150). Companies are taking the bait. 24% of companies have adopted artificial intelligence in talent recruitment processes and 56% of companies, in the same study, stated in 2020 that they would utilize AI for hiring within a year (Regina, 2023, p. 1084). Additionally, from 2018 to 2022, these algorithms' usage has increased by 270%, with over 98% of companies on the Fortune 500 using them (Tilmes, 2023, p. 1).

While AI holds promise in augmenting human decision-making, its indiscriminate use can in fact exacerbate existing biases and discrimination in general. An infamous case study is Amazon's attempt at creating an artificial intelligence hiring algorithm was simply a glorified way to discriminate against women (Andrews & Butcher, 2022, p. 147). These algorithms can be trained on data that is biased against certain groups of people or using features that are correlated with race, gender, or other protected characteristics, which will allow them to reflect and even reinforce existing biases (Regina, 2023, p. 1091). Additionally, algorithms' opacity and complexity will make them even more difficult to contest or verify (p. 1084). If there is a consistent pattern showing these algorithms are biased and with many more of these algorithms in development, many qualified candidates will be overlooked for jobs, which would have a negative impact on the individuals unfairly rejected and subsequently on the whole of society.

Thus, the ramifications of AI's ubiquity extend beyond individual interactions, influencing broader socio-economic dynamics and power structures.

When an AI algorithm is created and deployed, many actors influence and are influenced by it. Data scientists and programmers who create the models, the data being collected, and the clients whose needs are fulfilled by these models will of course inject their influence, big and small; noticed and unnoticed. These influences will ultimately affect the entities that are subject to a model's predictions, such as those who are deemed more or less likely to recidivate in predictive recidivism algorithms or those who are rejected from a position by automated hiring models.

In light of these considerations, responsible implementation of AI becomes very necessary. While AI offers potential benefits, its implementation requires careful planning and oversight. Data scientists play a pivotal role in ensuring that fairness considerations are not overlooked but are integrated into the very design and deployment of AI systems.

**Methodology**

To reiterate, my research question is what potential problems arise in employing artificial intelligence and machine learning hiring algorithms in the recruitment process. For my methodology, a literature review will be the primary source of data for my research. I've selected four studies and academic articles from the past one to two years that cover at least most aspects of the consequences of unregulated AI algorithms in hiring and recruitment efforts. These papers provided the theoretical and conceptual framework for my research analysis by providing statistics, examples, trends, public responses to the issue, and more. I've also selected a couple

case studies from this literature based on relevance to my research question, strength of consequences to certain actors, and how well it demonstrates the issue highlighted by my research question.

Usage of case studies is a valuable approach for my research. The case studies I've selected provide rich and detailed case descriptions that provide strong evidence that I will use to support my research question. Case studies are also a reliable source of evidence because they have most likely been circulated around the media, verified, and investigated thoroughly before being used in the literature I read. Lastly, there is potential for comparative analysis across multiple existing case studies. It's important to examine more than one case study to properly prove a relevant argument. Perhaps multiple case studies will highlight different aspects or perspectives of the problem highlighted by my research question.

Additionally, through the lens of Actor Network Theory, we can see how AI hiring algorithms become a powerful actant within the recruitment process. This theory posits that both human and non-human actors (technologies in this case) shape and are shaped by the networks they interact with. Data scientists and programmers, by their background and experiences, may encode their own biases into the algorithms they design. The data used to train these models can also be biased, reflecting societal prejudices or perpetuating existing hiring patterns. Clients, with their specific needs and perspectives, can further influence the algorithms towards favoring certain demographics or skillsets. These combined influences can have a significant downstream impact. The algorithm acts as a powerful force, influencing the decisions of recruiters and ultimately impacting the opportunities of job seekers. Unfortunately, the actors most negatively

affected by biased algorithms are often those from underrepresented groups, further marginalizing them within the job market.

**Literature Review**

In her article, Regina (2023) argues that the key argument of the source is the importance of fairness in machine learning and the challenges that make it difficult to achieve. She provides evidence that biases inherent in datasets can lead to unfair outcomes, such as when the labeling process was performed in an already unfair manner, or if there are under-represented populations in the dataset (p. 1091). She also emphasizes the need for representative datasets and the importance of developing a benchmarking framework to allow a more unified and standard evaluation process for fairness mechanisms (p. 1110). I want to argue that there are a diverse range of causes for bias in these models and individual attention needs to be paid for each model's potential fairness issues.

In their paper, Andrews & Bucher (2022) specifically focus on gender inequality in AI hiring algorithms. Their key argument is that AI hiring practices can perpetuate gender inequality and discrimination (p.148). The evidence offered to support this argument includes studies and examples of how algorithms and recommendation engines can be biased against women, particularly in male-dominated industries. They also argue about the importance of diversity and inclusion both in the workplace and in the implementation of any AI technologies, the potential for AI to be used as a tool for promoting gender equality, and lastly the legal implications of gender inequality in these algorithms. Because this source focuses so much on gender inequality in machine learning hiring models, I am using this information to, just like with a previous source discussing ableism, show that gender inequality is a large facet of algorithmic unfairness.

Zeide (2022) argues in their article that artificial intelligence and automated decision-making systems can perpetuate patterns of inequality and create a barrier to opportunity, which the authors refer to as the "Silicon Ceiling" (p. 403). The Silicon Ceiling is conceptualized as not just a black box of “algorithmic opacity”, but as an imperceptible barrier of entry (p. 403). The source provides evidence from various studies and examples to support this argument, including how algorithmic assessments can reinforce historical biases and how automated platforms, with access to a wider, online candidate market, in their search for unrealistically skilled candidates can exclude marginal candidates from consideration for jobs or educational opportunities. The source offers several secondary arguments to support the main argument, including the need for greater transparency and accountability in AI systems, the importance of promoting diversity and inclusion in the development and deployment of these systems, and the potential for legal and regulatory frameworks to address the harms caused by the Silicon Ceiling. I am using this information to offer more evidence of the extent of problems hiring and recruitment models have beyond simply discriminating against marginalized groups. I want to show that the problem is deeply entrenched in the fabric of today’s society and cannot be fixed with one blanket, “band-aid” solution.

The key argument that Tilmes (2022)’s paper is using is that AI hiring tools can exacerbate biases against marginalized groups, including disabled people, which is its primary focus, and argues that existing fairness measures alone are not enough to be inclusive to disabled people (p. 1). The evidence offered studies that have found that AI hiring tools can discriminate against job candidates based on factors such as their race, gender, and socioeconomic status. The source adds disability onto this list and discusses how disabled people are often excluded from opportunities in the workplace due to ableist attitudes and practices, and how AI hiring tools can

reinforce these biases by relying on narrow definitions of “ability” and “productivity.” The source also argues for the need for more research on the impact of AI hiring tools on marginalized groups, the importance of centering the experiences of disabled people in discussions of algorithmic bias, and the potential for a disability justice approach to inform broader efforts to address algorithmic bias. I am using this information to highlight that although there may be solutions to fairness problems with machine learning models, merely calculating fairness doesn’t do much to completely solve any issue in any specific model, and that ableism is a large chunk of the algorithmic unfairness problem we are facing today.

**Results and Discussion**

Different machine learning models are curated for every different situation. Whether it be purpose, location, data collected, or others, every model will be unique, with their own potential fairness pitfalls. As discussed, my literature also provides some case studies that will work to answer my research question.

Resume screening is a method used by employers to efficiently sift through job applications and identify the most suitable candidates for their organization. AI programs are vital in this process because they analyze huge amounts of resumes and provide recommendations to employers regarding which candidates should be prioritized for contact. For instance, one vendor, Ideal, utilizes AI technology to evaluate resumes and rank applicants based on their alignment with the job requirements (Regina, 2023, p. 1096). In a specific case, an audit of an undisclosed company's AI resume screening tool revealed two factors—being named "Jared" and having played "high school lacrosse"—that were seemingly correlated with job performance for that employer (p. 1096). Despite a potential statistical correlation, the lack of

direct relevance to actual job performance renders these factors discriminatory and devoid of valid defense. It would be truly challenging for an employer to justify these criteria as essential for successful business operations under a business necessity defense.

As mentioned earlier in the introduction, one of the most well-known case studies of hiring discrimination arising from hiring algorithms utilizing artificial intelligence is from Amazon. Around 2018 or 2019, Amazon internally developed a recruiting tool that primarily used resume screening to hire candidates for software engineer openings. Their data scientists fed the model ten years of hiring data to train it in the hopes that it would hire only the most qualified candidates (Regina, 2023, 1097). However, because the model was fed resumes for software engineering, which is a field dominated by men, all the algorithm managed to do was rank any resume that mentioned "woman" or "female" lower. Those included resumes that contained phrases like "Captain, Women's Soccer Team", or "National Women's Chess Champion" and applicants from all-women's colleges (Andrews & Butcher, 2022, p. 147). The program even rejected qualified female candidates before any interviews were made, and accepted candidates that weren't qualified (p. 147). Amazon admitted that the program was not helpful in determining job qualification and scrapped the project before it was used (p. 147).

Video interviewing is a popular option for employers for its virtual and automatable nature. HireVue used AI to analyze video interviews and determine eligibility based on features such as tone of voice, facial expression, and answer content (Regina, 2023, 1094). While automated video interviews are attractive because of cost reduction in hiring processes, many of the nuances necessary to conduct fair interviews are left behind. For example, there is no evidence to show that any expression espoused by an interviewee is any indication of job

qualification (p. 1094). Additionally, those with accents and speech impediments receive negative grades from HireVue's video interviewing AI, and the program struggled with accurately detecting expressions on those with darker skin (p. 1094). Overall, facial recognition algorithms are thirty times more likely to misidentify darker-skinned women than lighter-skinned men (p. 1094). These issues are a race-based, speech-based, and gender-based discrimination issue within these kinds of algorithms. They, although not revealed by HireVue, are caused most likely by an overabundance of lighter skinned, more "vocally proficient" (i.e. no accent nor speech issues) men in the training data fed to the models. Luckily, HireVue, after conducting a third-party audit in 2018 to study their use of AI in job interviews and presumably finding these issues, decided to halt their use of AI (p. 1112).

According to the literature and case studies, there are indeed many implications of employing artificial intelligence/ machine learning hiring algorithms in the recruitment process. From countless forms of discrimination to fantastical hiring criteria, AI hiring algorithms can produce dangerous, misleading, and unhelpful hiring recommendations for companies who are looking for a quick and easy solution to their mass hiring decisions.

AI fairness refers to the principle that AI algorithms should not discriminate against certain groups that are affected by what it pertains to. In many such cases, individuals or groups based on protected characteristics like race, gender, or age are the most affected by the lack of fairness. AI fairness encompasses both the fairness of the outcomes produced by AI systems (e.g., who gets hired) and the fairness of the processes used to build and train those systems (e.g., the data used to train the algorithm). It is essential because AI systems are increasingly used in critical decision-making processes, including hiring. Biases in these systems will cause many

unforeseen consequences, including those introduced in the case studies I have showcased, and others that I have not.

There is, for certain, a much longer answer to my research question. As I mentioned before, it is, in many cases, individuals/groups of people that are of protected/marginalized groups that are the most affected by the absence of AI fairness. In the case of gender, I have already detailed the Amazon case and how their algorithm failed to accurately predict good hires and kept discriminating against women. It realized that most software engineers hired by Amazon were men, as software engineering is a male-dominated field. Therefore, by using this flawed, biased data, when it came across a candidate who was a woman, it rejected her with no other considerations.

My literature also suggests that while AI hiring tools promise efficiency and less bias, they often exacerbate biases against those with disabilities. The complexity and variability of disabilities, as well as the contextual nature of their manifestation, challenge the effectiveness of algorithmic fairness initiatives. These initiatives tend to oversimplify and flatten variance among disabled individuals, potentially reinforcing the pathologization of disabled people. It suggests that AI fairness alone is insufficient for creating accessible, inclusive AI because of the variability in different disabilities.

Overall, unfair hiring algorithms can disadvantage qualified individuals from underrepresented groups, limiting their access to good jobs and career advancement. This can have a significant impact on their livelihoods and economic well-being. The financial implications can be substantial, with lower wages and fewer opportunities for promotion translating into a wealth gap that widens over time. Furthermore, limited access to good jobs can

restrict an individual's ability to save for a home, invest in education, or build a secure financial future for their families.

If individuals perceive the hiring process as biased, it can lead to a lack of trust in institutions and discourage them from applying for future opportunities, even with qualified profiles. When talented individuals from underrepresented groups feel they are being unfairly excluded because of the constant rejection from AI-driven hiring systems, it can lead to a decrease in self-confidence and motivation. This can create a vicious cycle, where individuals who are discouraged by biased algorithms may be less likely to invest the time and effort required to develop their skills and qualifications, further limiting their future employment prospects. The loss of diverse perspectives from unfairly excluded, potentially qualified candidates can hinder an organization's ability to develop effective solutions and can ultimately limit its success.

These issues have become notable enough that several advocacy groups and U.S. legislators have expressed concern about how AI hiring algorithms could violate Title VII of the Civil Rights Act, including the ACLU, CEOS of private companies such as Somen Mondal of Ideal, and several U.S. Senators (Regina, 2023, p. 1098). This effort led to the creation of the Algorithmic Accountability Act, which has not yet been passed. The Act establishes that businesses that use “high-risk automated decision systems” to make critical decisions to study and report about the impact of those systems on consumers (p. 1099).

**Conclusion**

By examining the evidence and the response to this new AI wave, we can see that responsible implementation of AI is urgently needed. While AI offers astronomic benefits to

society, its implementation requires careful planning and oversight. However, as we've seen in the realm of hiring algorithms, the unabated implementation of AI can have disastrous consequences, particularly in areas that impact fundamental human rights and opportunities. This discussion on AI fairness underscores the urgent need for a responsible approach to AI development and deployment.

Imagine a future where qualified individuals are systematically denied jobs, loans, or even healthcare due to biased AI algorithms. The ramifications extend far beyond the individual. Unfair AI can exacerbate existing societal inequalities, erode trust in institutions, and stifle innovation by limiting the contributions of diverse talent pools. For organizations that deploy these biased systems, the blowback can be severe, damaging their reputation and hindering their ability to attract top talent.

Data scientists play a pivotal role in ensuring that fairness considerations are not overlooked but are integrated into the very design and deployment of AI systems. The responsible implementation of AI is imperative. Just as the benefits of AI are vast, so are the potential pitfalls. We must proactively address these challenges through comprehensive regulations, industry best practices, and a cultural shift towards prioritizing fairness in AI development. We can ensure that AI serves as a force for good, one that empowers individuals, fosters innovation, and propels us towards a more equitable and just future.

# References

Andrews, L., & Bucher, H. (2022, October 1). AUTOMATING DISCRIMINATION: AI HIRING PRACTICES AND GENDER INEQUALITY. Cardozo Law Review, 44(1), 145 - 202.

Regina, G. (2023, June 1). Do You Even Know Me? A.I. and its Discriminatory Effects in the Hiring Process. Hofstra Law Review, 51(4), 1081 - 1116.

Tilmes, N. (2022, June 1). Disability, fairness, and algorithmic bias in AI recruitment. Ethics & Information Technology, 24(2), 1 – 15.

Zeide, E. (2022, December 1). THE SILICON CEILING: HOW ARTIFICIAL INTELLIGENCE CONSTRUCTS AN INVISIBLE BARRIER TO OPPORTUNITY. UMKC Law Review, 91(2), 403 - 436.