International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 12 Issue 1, January-2023, Impact Factor: 7.957

The Impact of the Lack of Diversity in Developers and Data on Gender Bias in AI

Anika Reddy Gaddam

Research Scholars Program, Harvard Student Agencies, In collaboration with Learn with Leaders

ABSTRACT

Artificial Intelligence has taken over almost every aspect of our lives. From being used to hire employees to assist in healthcare to being used in transport, it has proven to be applicable in any field (Nadeem et al., 2021). Although its integration into the human world has resulted in advancements and breakthroughs in many fields, AI also poses significant threats to disadvantaged groups, one of the main ones being women. In fact, such high integration of AI into human society has caused it to amplify pre-existing social biases and create new ones (Nadeem et al., 2021). The existing literature reviews the multitude of causes for such bias to be propagated and created by AI algorithms. However, there is limited literature that solely addresses the impact that the lack of diversity of developers and datasets has on gender bias in AI and the relationship between the two in specific. Thus, this investigation adds to the limited literature present on the impact of the lack of gender diversity and provides further detail into the issue by analyzing two case studies. It proves that the lack of diversity in developers and data sets propagates gender bias prevalent in AI. The investigation also highlights the importance of the issue and the caution that must be taken to prevent its negative impacts by providing a thorough understanding of the issue. Through further understanding of one of the main causes of gender bias in AI, proper measures can be put in place to prevent the propagation of such bias.

Keywords: Artificial Intelligence, gender bias, machine learning, algorithms, diversity.

Subject: Technology

INTRODUCTION

AI refers to the simulation of human intelligence in machines, and is designed to ‘think’ like and mimic human beings. The term can also be applied to a machine that displays traits corresponding to human intelligence like learning or problem-solving (Frankenfield, 2022).

One of the subjects of AI that is often confused with AI itself is Machine Learning. Machine Learning (henceforth, ML) refers to the concept that algorithms can learn from and adapt to new data without human assistance (Frankenfield, 2022). Hence, AI and ML are very different from each other. AI is more of an all-encompassing term while Machine Learning is a subset of AI.

Gender bias in AI

As mentioned, Artificial Intelligence has become a crucial system in society. However, this integration has become a double-edged sword, introducing new ethical and research concerns while advancing many fields of study. Research has shown that some areas of concern are accountability, fairness, and human employment. As AI is integrated into our daily lives, any biases that it might create will have negative impacts all over society. Hence, it’s imperative that biases, especially gender bias which could affect large amounts of people, are mitigated (Nadeem et al., 2021).

Several studies have indicated that the primary reason for gender bias in AI is improper training data as well as a lack of diversity of developers (Nadeem et al., 2020) To explore the effect that these two factors have on gender bias in AI, I will be observing two case studies: AI used in talent assessment and AI used in self-driving cars. These case studies highlight the direct impact that the lack of diversity of developers and training data has on gender bias in AI, along with the societal implications created resultantly.

CASE STUDY 1

Gender bias in hiring

Amazon attributes most of its e-commerce dominance to automation. Hence, when the company decided to expand this automation to its hiring process in 2014, it expected success along similar lines. It assembled a team that designed a hiring tool that used AI to provide job candidates with scores from one to five. The group created 500 computer models and taught them to recognize around 50,000 terms that appeared in past candidates' resumes. However, by 2015 the company found that its new hiring tool showed gender bias while rating candidates (Dastin, 2018).

Many researchers attributed this to the lack of diversity in the data the company had used. It used resumes submitted to Amazon over the past 10 years to train the hiring tool. However, most of these resumes came from men, an indication of male dominance in the industry. For years, top tech companies have struggled to close the gender gap in hiring. Currently, there is a pronounced disparity in the men-to-women ratio of software developers, with men drastically outnumbering women. This trend has also been visible in Amazon's hiring patterns. Thus, the tool picked up on this bias, producing results that discriminated against words found in women's resumes such as the very word "women's" in women's chess club (Dastin, 2018).

The dataset used by Amazon required programmers to identify and calculate appropriate features for the tool to look for. To do this well, the team of developers would have to consult with hiring managers (Lauret, 2019). However, through glimpses into Amazon's current workforce and the output of the programming tool, it can be assumed that special emphasis wasn't given to consultation of female hiring managers who would have been able to reduce the bias in the program by suggesting appropriate features to use for women applicants. Hence, the features selected would have largely benefitted male applicants while putting women applicants at a disadvantage.

The bias created by the data could have had drastic implications if it had not been identified by the Amazon team. It would have denied more women jobs, marginalizing women in the STEM field further. Additionally, many qualified women would have been denied an opportunity to work at Amazon and especially in the development of AI at Amazon. This would further propagate gender bias as women will be further excluded from the crucial decision in the creation of AI. Lastly, the bias in the algorithm could also result in the hiring of incompetent candidates which would reduce the productivity and efficiency of the company – putting it at a disadvantage as well (Lauret, 2019).

CASE STUDY 2

Gender bias in object detection

A recent study conducted by the Georgia Institute of Technology investigated "whether the state-of-the-art object detection systems had an equitable predictive performance on all types of pedestrians" (Wilson et al., 2019) In the study, they examine how well several object detection systems, similar to the ones used by self-driving cars, fared in detecting people of different skin tones and genders. The results detected that the systems were 5% less accurate for darker-skinned people (especially women) regardless of the change of any other variables (Samuel, 2019).

The study highlighted that possible reasons for this could be that the datasets used by major companies to train their models have more data pertaining to white men than they do of women, especially colored women. Thus, the algorithm develops a bias toward this section of society (Samuel, 2019). This conclusion is further supported by the fact that many cars use dummies with male physiques during their developmental testing. Hence, standard measurements of women are generally not taken into consideration when designing seatbelts, airbags, or headrests in cars (Niethammer, 2022). Thus, it would be reasonable to assume that when these features are extended to self-driving cars, women's requirements aren't given as much consideration as men. Additionally, even if data from the discriminated classes is used, the algorithm doesn't place enough emphasis on learning from it (Samuel, 2019).

Additionally, it was identified that one possible reason for less emphasis given to data from discriminated classes might have been the lack of diversity of developers working on such projects. As most developers of such systems are white males, the system tends to develop their unconscious biases and forgo important steps in the bias mitigation process. However, if more developers were involved in the creation of such systems, their personal experience would allow them to catch possible avenues of bias entry into the system and mitigate it in ways that their male counterparts cannot (Samuel, 2019).

If these biases are let continue, it could lead to several self-driving cars putting the lives of countless minorities at risk. Improper object detection can lead to many women (especially those of color) being unrecognized by self-driving cars leading to several accidents (Samuel, 2019). Additionally, cars designed to suit a male physique might not be able to cater to women, leading to more women in physical jeopardy than men when faced with similar accidents (Niethammer, 2022). Since the negative implications of such bias have drastic effects, it is imperative that all precautions are taken to mitigate it.

CONCLUSION

From an early age, girls in school are pushed to feel less enthusiastic about a career in IT or AI. Many girls are never encouraged to pursue these fields and hence develop a view that math and science are only meant for boys. However, even when women do enter the IT workforce, they are likely to leave sooner than men do. The main reasons for this include male-dominated work culture, sexual discrimination, gender pay gaps, and a lack of role models in higher-level-senior positions. These factors significantly decrease the much-needed diversity in the AI sector (Women in Tech). This loss of diversity becomes even more destructive considering that diversity amongst employers is one of the primary ways to mitigate bias propagated by data. Data used by many developers during the creation of AI models is biased against women due to historical discrimination of women in society. “Gender ideologies are still present in today’s digital and advanced world and therefore still embedded in the training data by the developers; hence resulting in the algorithm learning stereotypical concepts of gender”. More diversity in the development of such AI systems would lead to these stereotypes being challenged. Additionally, it would ensure that proper precaution is taken when using data that has gender bias embedded in it. This is further illustrated by the case studies explored in this paper. Whether it’s the identification of appropriate hiring features suitable for women, or the utilization of and emphasis on data from discriminated classes, both systems required a more diverse set of developers to combat the bias in their data. Thus, greater efforts must be made to diversify development teams to include more women, and other discriminated classes in order to alleviate the gender bias prevalent in so many AI systems.

REFERENCES

[1]. Dastin, Jeffrey. “Amazon scraps secret AI recruiting tool that showed bias against women.” *Reuters*, Oct. 2018, www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

[2]. Frankenfield, Jake. “Artificial Intelligence: What It Is, and How It Is Used.” *Investopedia*, Jul.2022, www.investopedia.com/terms/a/artificial-intelligence-ai.asp.

[3]. Lauret, Julian. “Amazon’s sexist AI recruiting tool: how did it go so wrong?” *Medium*, Aug.2019,becominghuman.ai/amazons-sexist-ai-recruiting-tool-how-did-it-go-so-wrong-e3d14816d98e

[4]. Nadeem, Ayesha; Babak Abedin; and Olivera Marjanovic, "Gender Bias in AI: A Review of Contributing Factors and Mitigating Strategies". ACIS 2020 Proceedings, 27, https://aisel.aisnet.org/acis2020/27.

[5]. Nadeem, Ayesha; Olivera Marjanovic; Babak Abedin. “Gender Bias in AI: Implications for Managerial Practices.” *20th Conference on e-Business, e-Services and e-Society (I3E)*, Sep 2021, pp.259-270, ⟨10.1007/978-3-030-85447-8_23⟩. ⟨hal-03648154⟩

[6]. Niethammer, Carmen. “AI Bias Could Put Women’s Lives AT Risk – A Challenge For Regulators.” *Forbes*, Mar. 2022, www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=487de4534f27.

[7]. Samuel, Sigal. “A study finds a potential risk with self-driving cars: failure to detect dark-skinned pedestrians.” *Vox*, Mar. 2019, www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin.

[8]. Wilson, Benjamin et al. “Predictive Inequity in Object Detection.” *ArXiv* abs/1902.11097, 2019), pp. 1.

[9]. Women in Technology and IT. “Why Are There So Few Women in Artificial Intelligence?” www.womenintech.co.uk/why-are-there-so-few-women-in-artificial-intelligence-ai#:~:text=Higher%20levels%20of%20attrition%20for,growing%20diversity%20in%20the%20sector.