

Algorithm Bias in Artificial Intelligence: Implications and Impact within our Society

Oluchi Emmanuella Obinna w21575872

ITAI-1370: Artificial Intelligence History

Prof. Machuria Johnson

September 30, 2023

Abstract

Artificial intelligence [AI] has become a crucial component of daily life, influencing choices in a variety of areas, from hiring practices to criminal sentencing. AI can connect data, identify patterns, and produce results across disciplines and borders. Humans can be freed from boring or repetitive activities by using AI, which can be more consistent than humans and swiftly adapt to changing inputs (Stahl, 2021). However, there are issues with using AI in these important fields. Algorithm bias, which occurs when machine learning models render biased or discriminatory judgments, is a major issue within the AI community. This refers to the systemic mistakes and prejudices that AI systems may exhibit as a result of skewed data or biased design. When it comes to AI technology, algorithm bias can have a big impact on both individuals and society as a whole. The notion of algorithm bias, its ramifications for AI technology, and its wide-ranging effects on people and society will all be covered in this essay.

Algorithm Bias in Artificial Intelligence: Implications and Impact within our Society

Data bias is frequently where algorithm bias starts. Humans select the data that algorithms utilize and determine how the algorithms' outputs will be used, which leads to AI bias (Marr, 2022). Data is the lifeblood of AI, and it can produce biased results if it is tainted with old prejudices or lacking variety. Another aspect of algorithm bias is model bias. Assumptions form the basis of model construction, and if these assumptions are skewed, the resulting AI system will also be prejudiced. For instance, a facial recognition system that has been primarily trained on one race group may perform poorly when used on other racial groups, perpetuating racial biases. These biases are greatly influenced by the procedures used for data collecting and the way the data is presented. Biased statistics can amplify social gaps, reinforce preexisting preconceptions, and further underrepresent minority groups.

Inequalities within Algorithm Bias

The possibility of societal inequities being reinforced and sustained by algorithm bias is one of the most urgent problems associated with it, and it has an impact on justice and equity. AI programs frequently draw their learning from past data, which can include societal biases. For instance, AI algorithms educated on past recruiting data may unintentionally prejudice against particular demographic groups throughout the employment process. Individuals with speech problems, asymmetry, or different facial features may be subject to discrimination by AI systems. Certain communities can be further marginalized when various gestures, gesticulations, and communication styles are misread. Autistic people and people with physical or cognitive and sensory limitations are particularly susceptible to prejudice by AI algorithms. Misidentification of facial or speech patterns can have serious repercussions and directly put the lives of individuals affected in danger (Randieri, 2023). This worsens already-existing disparities and thwarts efforts to create a workforce that is more inclusive and varied. Because biased AI can exacerbate socioeconomic inequality rather than reduce them, this has alarming ramifications for society.

Diving into Racial and Gender Prejudice

Algorithm prejudice has troubling societal repercussions. It may strengthen bias, particularly racial and gender prejudices. Biased AI can exacerbate social polarization, reinforce negative stereotypes, and widen gaps across populations. As biased AI systems maintain unequal access to opportunities and resources, socioeconomic gaps may get worse. It may keep institutionalized discrimination alive. Equal opportunity is undermined by discrimination, which also intensifies oppression. Biased hiring processes can keep qualified people out of the workforce and prolong poverty and unemployment cycles. Furthermore, arbitrary restrictions on access to healthcare, financial, and educational opportunities can exacerbate socioeconomic inequities (Shin, 2020). In the justice system, where AI is used to support judges in making judgments on bail, sentencing, and parole, such biases can erode public

confidence. These algorithms may produce disproportionate results with some groups being treated unfairly if they show racial or gender bias. The effects of algorithm bias are profoundly felt by individuals.

Privacy and Autonomy

Fundamental worries in artificial intelligence include privacy and autonomy inside algorithm bias. Significant privacy issues have been highlighted by law enforcement organizations' use of facial recognition technology driven by AI (Hartzog & Selinger, 2018). With this, the threat to people's anonymity and privacy comes from the technology's ability to identify people in real time via surveillance footage, social media, and other sources. As biased AI systems may engage in intrusive profiling and monitoring, using skewed data to draw judgments about individuals, algorithm bias might jeopardize privacy. As AI develops, it will be able to make conclusions based on small data patterns that are challenging for people to see. As a result, people might not even be aware that their personal information is being utilized to make judgments that will have an impact on them (Rijmenam, 2023). AI systems may restrict personal choices or support discriminatory activities, this intrusion into private lives may have major ramifications for individual liberty and privacy rights. Additionally, people would feel forced to follow prejudiced AI-generated recommendations, which would undermine their autonomy and personal agency. In order to ensure that AI respects people's rights and promotes their freedom of choice and privacy, there must be strong protections and ethical concerns in AI development, which are made necessary by the erosion of privacy and autonomy.

Accountability and Transparency

Any ethical and reliable AI system must have accountability and transparency, but these features are frequently undermined by algorithm bias. It may be difficult to assign blame or hold someone accountable for the results when AI systems make biased judgements. This is

made worse by the fact that AI algorithms, also known as "black boxes," are frequently complicated and opaque (Marr, 2022). When a biased judgment hurts an individual or a group, it can be challenging to track the decision-making process, comprehend why a particular decision was made, or identify who should be held accountable. Furthermore, algorithm bias might result in AI systems that are opaque. Biased underlying data used to train AI models might provide skewed conclusions that are difficult to explain or defend. Because consumers and other stakeholders are kept in the dark about the reasoning behind particular decisions, this lack of transparency has the potential to reduce trust in AI systems. Without transparency, it becomes difficult to spot and address prejudice in AI systems, impeding attempts to develop more equitable and responsible AI technology.

More to Eliminate AI Bias

A multifaceted strategy is needed to address algorithm bias. Transparency and accountability are crucial in the first place. Developers and companies must be open and honest about the data and algorithms they use. For the purpose of identifying and reducing prejudice, they should perform routine audits and assessments. Initiatives like the European Union's AI Act, which aims to “strengthen Europe's position as a global hub of excellence in AI from the lab to the market, ensure that AI in Europe respects our values and rules, and harness the potential of AI for industrial use” (Feingold, 2023), serve as examples of how regulatory bodies can play a significant role in creating norms and standards for algorithm fairness and transparency. AI development procedures should also incorporate ethical frameworks and principles. The key to minimizing prejudice is using diverse and inclusive datasets. More representative data must be gathered, and biases must be removed from the data before it is used to train AI models. Re-weighting and adversarial training are two additional methods that can be used to lessen bias in AI systems. This entails applying diversity and inclusion guidelines, carrying out rigorous testing and assessment, and putting in place ongoing

monitoring and feedback systems. It is crucial to use better methods for data collecting and preprocessing that address historical biases, guarantee representativeness, and regularly audit data for bias detection and reduction. The solution must also include vital elements of awareness and education. The ethical use of AI and the effects of bias should be taught to developers and practitioners in the field. Making people aware of algorithm bias and its effects can promote ethical AI use and hold companies and developers responsible.

Conclusion

Algorithm bias in artificial intelligence is a widespread problem with substantial ramifications for people and society, to sum up. It may provide unfair results, amplify social injustices, and undermine confidence in AI systems. Collective efforts are needed to address algorithm bias, including the development of ethical frameworks, better data practices, and enhanced knowledge and awareness. We can only create an ecosystem for AI that is more diverse, fair, and trustworthy by taking these steps, which also prevent biases and injustices from being reinforced.

References

Stahl, B. C. (2021). Ethical Issues of AI. *Artificial Intelligence for a Better Future*, 35–53.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7968615/>

Randieri, C. (n.d.). *Council Post: Unveiling The Role Of AI Algorithms: Unmasking Societal Inequities And Cultural Prejudices*. Forbes. Retrieved September 28, 2023, from

<https://www.forbes.com/sites/forbestechcouncil/2023/07/19/unveiling-the-role-of-ai-algorithms-unmasking-societal-inequities-and-cultural-prejudices/?sh=2780dba82f60>

Shin, T. (2020, June 4). *Real-life Examples of Discriminating Artificial Intelligence*. Medium.

<https://towardsdatascience.com/real-life-examples-of-discriminating-artificial-intelligence-cae395a90070>

Fintelics. (2023, May 5). *AI and its Impacts on Privacy and Security*. Medium.

<https://fintelics.medium.com/ai-and-its-impacts-on-privacy-and-security-57243e9d0b6f#:~:t>

van rijmenam, mark. (2023, February 16). *Privacy in the Age of AI: Risks, Challenges and Solutions*. Dr Mark van Rijmenam, CSP - the Digital Speaker | Strategic Futurist.

<https://www.thedigitalspeaker.com/privacy-age-ai-risks-challenges-solutions/>

Marr, B. (n.d.). *The Problem With Biased AIs (and How To Make AI Better)*. Forbes.

Retrieved September 28, 2023, from

<https://www.forbes.com/sites/bernardmarr/2022/09/30/the-problem-with-biased-ais-and-how-to-make-ai-better/?sh=7a90e6e34770>

Fintelics. (2023, May 5). *AI and its Impacts on Privacy and Security*.

Medium.[ext=The%20technology's%20ability%20to%20identify,from%20seemingly%20non%20sensitive%20data.](https://fintelics.medium.com/ai-and-its-impacts-on-privacy-and-security-57243e9d0b6f#:~:text=The%20technology's%20ability%20to%20identify,from%20seemingly%20non%20sensitive%20data.)

Feingold, S. (2023, June 30). “On artificial intelligence, trust is a must, not a nice to have,” one lawmaker said. #AI. World Economic

Forum.<https://www.weforum.org/agenda/2023/06/european-union-ai-act-explained/>