ARTIFICIAL INTELLIGENCE AND SOCIAL BIAS: GENDER, RACE, AND POWER DYNAMICS

Mr. Saket Paras Bisandre
Assistant Professor,
Nabira Mahavidyalaya, Katol
Email: bisandresaket@gmail.com
Mob. No.: 8698889019

Crossref DOI - https://doi.org/10.63665/rh.v7i2.27

Abstract :

AI is increasingly deployed to make decisions in many areas of modern life, including education, healthcare, law enforcement, hiring for jobs, and social media, among others. Despite the fact that AI is often thought to be fair and impartial, new research illustrates that these systems can actually echo and make worse existing unfairness based on gender and race. The current paper focuses on the inner side of power in AI and discusses how discriminatory ideas embedded into data, programming, and ownership cause unfair effects against already vulnerable groups.

This paper, viewed through the lens of gender, race, and power, demonstrates the need for the establishment of ethical rules, the use of data that represents all, and the inclusion of diverse groups within AI development. Correcting bias in AI is not only a technical but also a social and political obligation. The paper concludes that fairness and justice in AI building require one to reflect on power, to advocate for equity, and to ensure that technology plays a supportive role in creating a fairer world.

Keywords : Artificial Intelligence, gender bias, racial inequality, discrimination, ethics in AI.

Introduction :

Artificial Intelligence (AI) has become a powerful force shaping everyday life. It influences decisions related to education admissions, medical diagnoses, job recruitment, criminal justice, and digital communication. AI systems are often promoted as objective tools capable of reducing human error and bias. However, recent research challenges this belief by revealing that AI systems frequently reproduce existing social inequalities. These inequalities are particularly visible in relation to gender and race.

This paper argues that AI should not be understood merely as a technical system but as a structure deeply connected to social power. Bias in AI is not accidental; rather, it emerges from the data used, the choices made during programming, and the institutions that control technological development. By examining AI through the lens of gender, race, and power, this study highlights how technology can reinforce discrimination instead of eliminating it.

AI, Power, and Social Structures :

Power plays a central role in shaping AI systems. Decisions about what data is collected, which problems AI should solve, and who benefits from these technologies are often made by powerful institutions such as governments, corporations, and elite technological groups. These actors influence AI development in ways that often reflect their own interests and worldviews.

AI systems are therefore not neutral. They are embedded within social and political structures that privilege certain groups while marginalizing others. When power remains concentrated in the hands of a few, the resulting technologies tend to reflect dominant norms related to gender and race. This imbalance creates systems that unintentionally discriminate against women and racial minorities.

Gender Bias in Artificial Intelligence :

Gender bias in AI has become increasingly visible across multiple platforms. AI tools trained on historical data often reflect patriarchal norms that undervalue women's roles and contributions.

For example, automated hiring systems may prioritize male candidates because past employment data favors men in leadership and technical positions.

Similarly, language-processing systems frequently associate men with professional identities and women with domestic roles. Such patterns reinforce harmful gender stereotypes and limit opportunities for women. These biases are not caused by malicious intent but by the uncritical use of biased data and a lack of gender diversity in AI development teams.

The absence of women and non-binary individuals in decision-making roles within the tech industry further strengthens these biases. When AI systems are designed without diverse perspectives, they fail to account for the lived experiences of marginalized genders.

Racial Inequality and Algorithmic Discrimination :

Racial bias in AI is particularly evident in areas such as law enforcement and surveillance. Predictive policing systems often rely on historical crime data that disproportionately targets racial minorities. As a result, these systems reinforce cycles of surveillance and criminalization within marginalized communities.

Facial recognition technologies have also been criticized for higher error rates when identifying people of color. Such inaccuracies can lead to wrongful arrests, exclusion, and social harm. These outcomes highlight how racial inequality becomes embedded in AI systems through biased data and unequal representation.

Racial discrimination in AI reflects broader systemic racism within society. When AI systems are trained on unequal social realities, they learn and reproduce those inequalities. Addressing racial bias in AI therefore requires confronting structural racism beyond the technological sphere.

Data, Programming, and Ownership :

One of the key sources of bias in AI lies in data. Data used to train AI systems often lacks diversity and fails to represent marginalized groups accurately. When datasets are incomplete or skewed, AI outcomes become unfair.

Programming choices also shape how AI behaves. Developers make decisions about classification, prioritization, and risk assessment. These decisions are influenced by cultural assumptions and institutional pressures. Without ethical oversight, such choices can unintentionally harm vulnerable populations.

Ownership of AI technology further intensifies power imbalances. When AI systems are controlled by a small number of corporations or governments, marginalized communities have little say in how these technologies affect their lives. This lack of accountability raises serious ethical concerns.

Ethical Responsibility and Social Obligation :

Correcting bias in AI is not only a technical challenge but also a social and political responsibility. Ethical AI development requires transparency, accountability, and fairness. Developers and institutions must recognize their role in shaping social outcomes through technology.

The establishment of ethical guidelines can help prevent discrimination in AI systems. These guidelines should emphasize fairness, inclusivity, and respect for human rights. Regular audits and impact assessments can also help identify and address bias before harm occurs.

Importantly, ethical AI must involve diverse voices. Including women, racial minorities, and marginalized communities in AI development ensures that multiple perspectives are represented. Such inclusion is essential for building technology that serves society as a whole.

Toward Fair and Inclusive AI :

Creating fair AI systems requires reflecting on power dynamics embedded within technology. Developers must ask critical questions: Who benefits from AI? Who is excluded? Whose values are prioritized? Addressing these questions helps shift AI development toward equity and justice.

Inclusive datasets, participatory design, and ethical governance are key steps toward reducing bias. AI should support human dignity rather than reinforce inequality. When guided by ethical principles, technology can contribute to a fairer and more just world.

Conclusion :

Artificial Intelligence has immense potential to improve human life, but it also carries the risk of reinforcing social inequality. This paper has shown that bias in AI is deeply connected to gender, race, and power. Discriminatory outcomes emerge from biased data, programming decisions, and unequal ownership of technology.

Ensuring fairness and justice in AI requires more than technical solutions. It demands social awareness, ethical responsibility, and political commitment. By reflecting on power and advocating equity, society can guide AI development toward inclusive and just outcomes. Technology must ultimately serve humanity, not deepen existing divisions.

Works Cited :

- Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press, 2019.
- Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in
- Commercial Gender Classification." *Proceedings of Machine Learning Research*, 2018, pp. 1–15.
- Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021.
- Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, 2018.
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016.