PakTech Today | Pakistan's Premier Technology Magazine
Vol. 01, Issue 06 • June 2026

TECHNOLOGY FEATURE Industrial News / Technology Report / Research Communication

Can AI Be Biased? Understanding Algorithmic Bias

Exploring how hidden biases in AI systems affect decision-making and why fairness in algorithms is a critical challenge in today's digital world.

By **Abdul Moez Bin Mansoor** | Information Technology Student | Published: June, 2026
moezmansoor44@gmail.com | Position: Researcher / Student

Abstract

Artificial Intelligence has changed a lot since it was first thought of in the middle of the century. Now it is used in areas like healthcare, finance, education and criminal justice. At first Artificial Intelligence was made to be like intelligence and to make decisions like humans do. Now Artificial Intelligence is used to make things easier and to do tasks. Even with all the good things Artificial Intelligence can do people are worried that it is not always fair.
The thing is, people think Artificial Intelligence is neutral and always right but that is not true. Artificial Intelligence can be. This bias usually comes from the data that is used to train Artificial Intelligence from the people who design it and from the way society has been in the past. This study is looking at whether Artificial Intelligence can be biased and why this happens. It is also trying to figure out how this bias affects the decisions that are made and how we can stop it from happening. To do this the study is looking at what other people have written about Artificial Intelligence and bias. What the study found out is that Artificial Intelligence can be biased and this can make the problems we have in society even worse. So it is very important that we make Artificial Intelligence in a way that's fair and good for everyone. To do this we need to work and make sure that Artificial Intelligence is used in a way that is fair, for all people. We need to make sure that Artificial Intelligence is used to help everyone, not some people.

Introduction

Artificial Intelligence is a part of our lives now and it affects things like healthcare, finance, education and the way crimes are handled. A lot of people think that Artificial Intelligence systems are fair because they use math and data to make decisions.. That is not true. Artificial Intelligence systems are made by people.
The people who make Artificial Intelligence systems and the data they use can both introduce biases, into the system.
Algorithmic bias is when Artificial Intelligence systems treat people unfairly because of their race, gender or where they come from. This can happen by accident.
This article is going to look at whether Artificial Intelligence can be biased how that happens and what we can do to make it less of a problem. What studies have found out to get a better understanding of algorithmic bias and Artificial Intelligence.

"AI is not inherently neutral—it reflects the biases present in the data and the society that shapes it."

Technology Overview

Literature Review

Researchers have found a main reason why this happens:

- **Bad Data:** The data used to train AI models is very important. If the data is not complete or does not represent everyone the AI system will probably have the biases. For example, some face recognition

systems make mistakes with people who have darker skin because they were not included in the training data.

- **Human Mistakes:** People. Program AI systems and they can put their own biases into the system without realizing it. These biases can affect how the model is designed what features are. How the data is labelled.
- **Algorithmic Bias**: For instance, an algorithm that is designed to be efficient might not consider fairness, which can lead to results.
- **Bad Cycles**: AI systems can make existing biases worse by creating cycles. For example, police prediction tools might target neighbourhoods more often which can lead to more arrests in those areas and that just makes the algorithms biases stronger.
- **Society's Bias**: AI systems often reflect the biases that already exist in society. If the data used to train AI systems is biased the systems will probably. Even make those biases worse. AI systems like these can be a problem because they can make things unfair, for certain groups of people.

Methodology

The research follows a qualitative research design that relies on secondary data as a method of studying algorithmic bias in Artificial Intelligence. Academic sources like Google Scholar, IEEE Xplore, and ScienceDirect, as well as reports of organizations like the European Commission and UNESCO were searched to gather relevant literature. The keywords such as algorithmic bias, AI fairness, and ethical AI were utilized in the search process. Relevance and recency were the basis of the selection of only peer-reviewed and credible sources. To examine the collected data using thematic analysis, the data were analyzed with identification of the key patterns which were sources of bias, impact and mitigation strategies. Despite the valuable information contained in the study, it is constrained by the use of existing literature and unavailability of primary data.

Data Collection

Relevant articles were collected from databases like Google Scholar, IEEE Xplore and ScienceDirect. The search process used keywords "Bias", AI fairness", "machine learning discrimination" and "ethical AI. The focus was on bias in AI fairness and machine learning discrimination to ensure ethical AI. These keywords helped in finding articles, on Google Scholar, IEEE Xplore and ScienceDirect.

Inclusion Criteria

- Peer-reviewed journal articles were looked at.
- Reports from organizations were considered.
- The focus was, on studies published in the 10 to 15 years.

Analysis Approach

The literature collected was analyzed thematically. The main themes that emerged were sources of bias, real-world examples and ways to reduce bias. The analysis was based on these areas to understand the topic better.

Figure 1: Illustration of algorithmic bias in Artificial Intelligence, showing how biased data can lead to unfair outcomes in decision-making systems.
Credit: Research Gate

Key Findings & Impact

Discussion: Can AI Be Truly Unbiased?

The question of whether Artificial Intelligence can be completely unbiased is complicated. It is possible to make Artificial Intelligence a little less biased. It is very hard to make it completely neutral. There is a reason for this:

1. **Dynamics Nature of Society**: What people think is fair. What the rules are can change a lot over time.
2. **Trade-offs Between Accuracy and Fairness:** Making Artificial Intelligence fairer might sometimes make it a little less accurate, at predicting things.

They want to make sure Artificial Intelligence systems are transparent Artificial Intelligence systems.

Impacts of Algorithmic Bias

Loss of Trust

When people think that AI systems are not fair, they will not trust them. This is a deal because people need to be able to trust the systems that are making decisions about their lives.

Legal and Ethical Concerns

If organizations use AI systems that're biased, they may get in trouble with the law.

Economic Consequences

When AI systems are biased, they can make decisions. Bias can lead to decisions reducing productivity and innovation. This means that companies will not be able to make much money as they could if they were using fair AI systems.

Mitigation Strategies

Diverse and Representative Data

To make AI systems better we need to make sure that the data we use to train them includes all kinds of people. Ensuring that training datasets include populations can reduce bias. This is important because AI systems learn from the data they are given.

Transparency

We need to be able to understand how AI systems make their decisions.

Fairness-Aware Algorithms

Some researchers are working on algorithms that're fair. These algorithms are designed to make sure that AI systems are not biased.

Regular Auditing

We need to keep an eye on AI systems to make sure they are not biased. Continuous monitoring and auditing of AI systems can. This means that we can catch problems before they get out of hand.

Ethical Guidelines and Regulations

Governments and organizations are establishing frameworks to promote AI development. This is important because it will help us make sure that AI systems are used in a way that's fair and good, for everyone.

What's Next?

Future Directions

Additionally, public awareness and education about AI bias are crucial for promoting responsible use.

- Developing diverse and representative datasets
- Designing fairness-aware algorithms
- Increasing transparency in AI decision-making
- Conducting regular audits of AI systems
- Establishing ethical guidelines and regulations

Conclusion

They can be biased. This is because they are made using data and decisions that may not be fair. The bias in algorithms is a problem that can cause unfair treatment of people. If we know where this bias comes from and do something about it we can make AI systems that are fairer and better for everyone. Fixing this problem is not about the technology it is also about doing what is right and working together with people from different fields to make sure AI systems are good for society and AI systems themselves need to be looked at closely to avoid bias, in AI systems.

About the Author
Abdul Moez Bin Mansoor is a Data Automation Specialist and Full Stack Web Engineer based in Faisalabad, Pakistan, currently pursuing a Bachelor's degree in Information Technology. His work focuses on building scalable, data-driven applications, with research interests in data analytics, artificial intelligence, and innovative digital solutions for real-world challenges. They can be reached at moezmansoor44@gmail.com

Keywords: Artificial Intelligence; Algorithmic Bias; AI Fairness; Ethical AI; Machine Learning; Data Bias; Digital Ethics; Cybersecurity

References

1. Binns, R. (2018). Machine learning fairness: Political philosophy lessons. *Machine Learning Research Proceedings, 81*, 149-159. https://proceedings.mlr.press/v81/binns18a.html

2. Buolamwini, J., & Gebru, T. (2018). Gender shades: Commercial gender classification Intersectional disparities in accuracy. *Machine Learning Research, 81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a.html
3. Chouldechova, A. (2017). Fair prediction that is disparate impacted. *Big Data, 5*(2), 153–163. https://doi.org/10.1089/big.2016.0047
4. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *Proceedings of ITCS*. https://doi.org/10.1145/2090236.2090255
5. Nissenbaum, H. and Friedman, B. (1996). Bias in computer systems. *ACM Transactions on information Systems, 14*(3), 330-347. https://doi.org/10.1145/230538.230561
6. Gebru, T., et al. (2018). Datasheets for datasets. *arXiv preprint*. https://arxiv.org/abs/1803.09010
7. Hardt, M., Price, E., & Srebro, N. (2016). Equal opportunity in guided learning. *NeurIPS*. https://papers.nips.cc/paper/2016/hash/9d2682367c3935defcb1f9e247a97c0d-Abstract.html
8. Hutchinson, B., & Mitchell, M. (2019). 50 years of test unfairness. *FAT Conference*. https://doi.org/10.1145/3287560.3287600
9. Kearns, M., & Roth, A. (2019). *The ethical algorithm*. Oxford University Press. https://global.oup.com/academic/product/the-ethical-algorithm-9780190948207
10. Kleinberg, J., Mullainathan, S., & Raghavan, M. (2017). Trade-offs in fairness. *ITCS*. https://doi.org/10.4230/LIPIcs.ITCS.2017.43
11. Mehrabi, N., et al. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35. https://doi.org/10.1145/3457607
12. Mitchell, M., et al. (2019). Model reporting model cards. *FAT Conference.,* https://doi.org/10.1145/3287560.3287596
13. Narayanan, A. (2018). 21 definitions of fairness and their politics. *FAT Conference*. https://doi.org/10.1145/3287560.3287596
14. O'Neil, C. (2016). *Weapons of math destruction*. Crown Publishing. https://weaponsofmathdestructionbook.com
15. Raji, I. D., & Buolamwini, J. (2019). Actionable auditing. *AAAI/ACM AIES.* https://doi.org/10.1145/3306618.3314244
16. Selbst, A. D., et al. (2019). Fairness and abstraction. *FAT Conference*. https://doi.org/10.1145/3287560.3287598
17. Shankar, S., et al. (2017). Representation-free classification. *arXiv preprint*. https://arxiv.org/abs/1711.08536
18. Sweeney, L. (2013). Online discrimination in delivering ads. *Communications of the ACM, 56*(5), 44–54. https://doi.org/10.1145/2460276.2460278
19. Verma, S., & Rubin, J. (2018). Fairness definitions explained. *IEEE Workshop*.
20. https://doi.org/10.1145/3194770.3194776
21. Wachter, S., Mittelstadt, B. and Floridi, L. (2017). Right to explanation. *International Data Privacy Law, 7*(2), 7699. https://doi.org/10.1093/idpl/ipx005
22. Zou, J., & Schiebinger, L. (2018). Artificial intelligence is also sexist and racist. *Nature, 559*, 324–326. https://doi.org/10.1038/d41586-018-05707-8
23. Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Biases in semantics are human-like. *Science*, 356(6334), 183–186. https://doi.org/10.1126/science.aal4230
24. Danks, D., & London, A. J. (2017). Autonomous systems: algorithmic bias. *IJCAI*. https://doi.org/10.24963/ijcai.2017/654