AI BIAS IN DECISION-MAKING SYSTEMS

Nikita Bahaley*, Monika Metkari,

Sakshi More, Esha Patil and Rohini Mutal

Department of Information Technology,

Pillai College of Arts, Commerce & Science (Autonomous), New Panvel, Maharashtra, India 410206

*Corresponding author E-mail: nikitabahaley@mes.ac.in

Abstract:

Artificial Intelligence is mostly used in decision-making across various sectors such as healthcare, finance, education, recruitment, and criminal justice. While AI systems have achieved significant technological success, they often present a one-sided picture to users. AI systems may produce unfair or discriminatory outcomes. Such bias creates inconsistencies that can misrepresent reality and negatively impact individuals and communities. This study investigates the existence of AI-induced bias in decision-making systems and examines public awareness and perception regarding this issue. A survey was done to collect data from students and working professionals to evaluate their level of awareness and to understand the extent to which they have experienced or observed biased behavior in AI systems. The findings aim to highlight the importance of fairness, transparency, and ethical practices in AI development and implementation.

Keywords: *AI, Ethical, Bias, Detection, Data, Algorithmic-Bias, Decision Making.*

Introduction

These days, machines that learn from experience shape choices in fields like hospitals, schools, money management, hiring, even courtrooms. From mountains of information, they pull patterns, offering guesses or advice so groups can move quicker without slowing down. Seen as steady and trustworthy - unlike people's instincts - they gain trust easily. Still, research lately hints these tools might copy unfairness found either in their training records or how their logic was built. Sometimes machines treat people unfairly because they learn from flawed examples. One reason lies in how information is collected - some voices get left out more than others. When past patterns carry old prejudices, software picks them up like echoes. Darker faces were often misread far more often than paler ones, scientists found while testing face scanners. This mismatch revealed hidden gaps in what tech sees as normal. Decisions made by code might lean one way without anyone noticing at first. Mistakes pile up quietly if nobody checks who gets affected most. Patterns repeat unless someone steps in to question where they

began. Blind spots grow larger when assumptions go unchecked too long. Who teaches the teacher matters just as much as the lesson learned. [1].

When artificial intelligence moves into everyday uses, unfair results might lead to deep societal harm. Hiring decisions, loan approvals, or police actions shaped by skewed algorithms often deepen gaps already present in society. Looking at why these biases appear matters - spotting them early, lessening their impact, comes down to clear rules, open methods, and building AI the right way [5][6].

Literature review

Lately, more people are studying how artificial intelligence can carry unfair patterns. Because these systems learn from past information, they sometimes copy mistakes humans made before. One look at facial scanning tools showed some struggled to recognize certain faces accurately. Joy Buolamwini and Timnit Gebru found those errors happened more often for women and folks with darker skin. When data lacks variety, the results tend to favor one group over another - this imbalance leads to uneven performance across different people. Machines aren't neutral just because they're machines; their choices reflect what they've seen during training [1].

Some studies looked at how artificial intelligence can carry bias in law enforcement tools. Backed by work from Julia Angwin and her team, one deep review checked software meant to judge a person's legal risk. It turned up results showing these tools often misjudged people based on race [2]. When old records shaped by social unfairness feed into machine learning, the outcomes tend to repeat past wrongs. Questions about right or wrong, open processes, and who answers for mistakes grow louder because of it.

One surprise turned up when machines picked job seekers - they leaned toward men. That happened at Amazon where software learned from old hiring choices filled mostly with males. When patterns repeat like that, outcomes tilt without anyone noticing right away. Now groups including IBM and Europe's governing body stepped in with rules to balance things out. Their approach highlights openness, responsibility, fairness, while pushing back against skewed results. Making sure automated choices serve everyone equally became central to how these systems grow. What once slipped through cracks now shapes new standards behind smart tools.

Objectives of the research

- The major objectives of this study are as follows:
- To grasp the idea of AI bias in decision-making.
- To recognize the sources of bias in AI.
- To understand the consequences faced by society in AI bias.
- To study public opinion about AI bias.
- To carry out statistical hypothesis tests.
- To propose potential solutions to address AI bias.

Methodology

1. Research design

It has analysis and is an understandable design. The Primary data collected through Google form. Secondary Data: Research papers, case studies, and ethical AI guidelines.
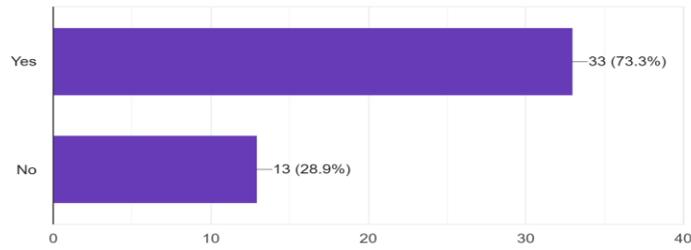
Are you aware of Artificial Intelligence?
45 responses

Figure 1: This response shows how many are aware of Artificial Intelligence

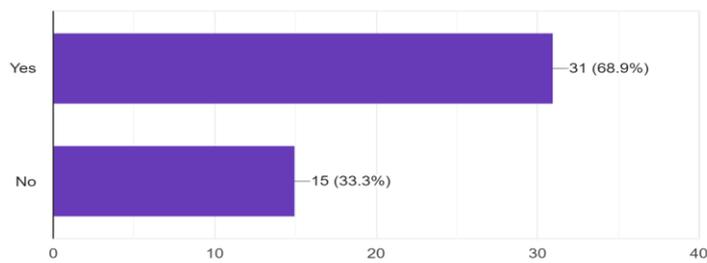Have you heard about AI bias before?
45 responses

Figure 2: This response shows that they have heard about AI Bias before or not

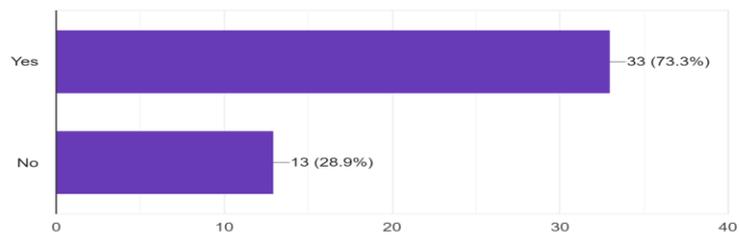Do you think biased AI decisions can affect people's lives?
45 responses

Figure 3: This response shows how many people think that AI decisions can affect people's lives
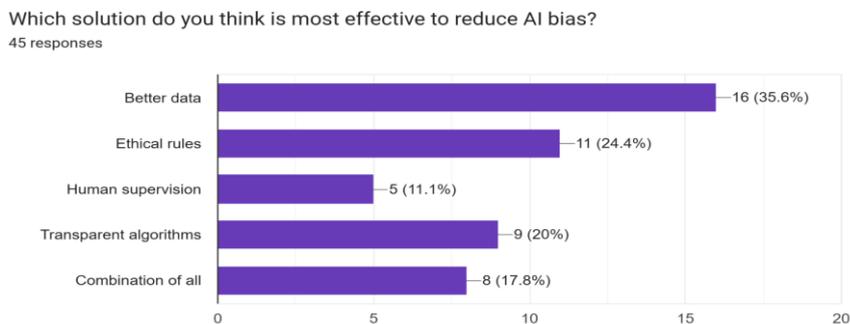
Which solution do you think is most effective to reduce AI bias?
45 responses

Better data	16 (35.6%)
Ethical rules	11 (24.4%)
Human supervision	5 (11.1%)
Transparent algorithms	9 (20%)
Combination of all	8 (17.8%)

Figure 4: This response shows how can we reduce AI bias decisions, what are the solutions for that

Where should AI NOT be allowed to make final decisions?
45 responses

Hiring & firing	18 (40%)
Medical treatment	10 (22.2%)
Education evaluation	14 (31.1%)
All of the above	6 (13.3%)
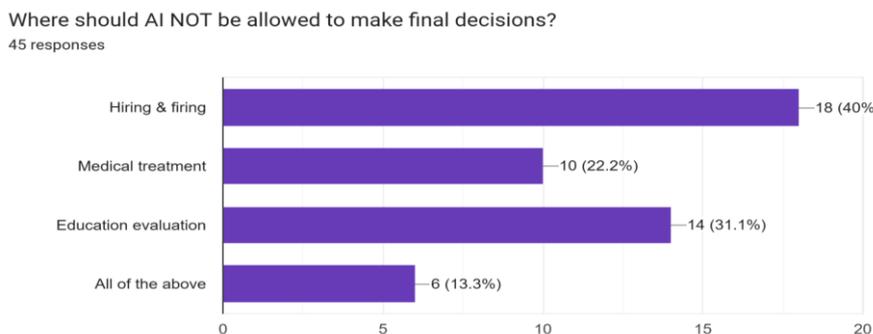
Figure 5: This response shows where AI decisions should be restricted

What should be the main goal of ethical AI?
45 responses

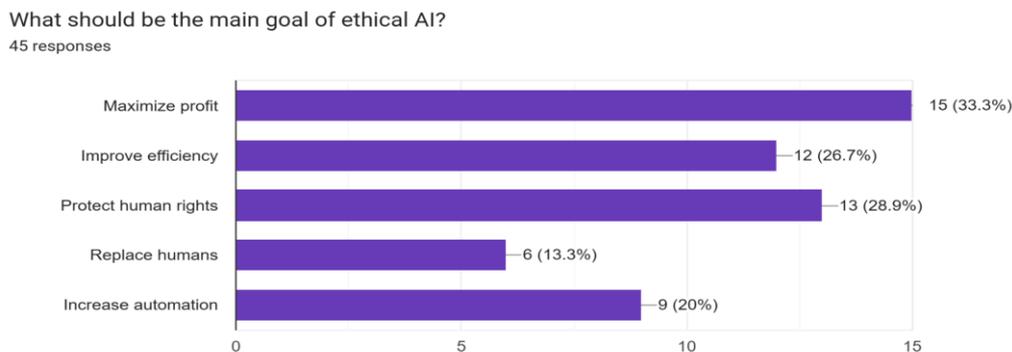|	
Maximize profit	15 (33.3%)
Improve efficiency	12 (26.7%)
Protect human rights	13 (28.9%)
Replace humans	6 (13.3%)
Increase automation	9 (20%)

Figure 6: This Response shows what the main purpose of ethical AI is

2. Data collection table

Table 1: Response collected through survey distribution

Response Category	Number of Responses
Strongly Agree (AI is biased)	18
Neutral	21
Disagree	8
Total	**47**

Table 2: Converted for the Hypothesis Test

Category	Count
Biased (Agree + Neutral)	39
Not Biased	8
Total	**47**

3. Hypothesis testing

Null Hypothesis (H_0): AI decision-making systems are not biased.

Alternative Hypothesis (H_1): AI decision-making systems are biased.

Statistical test used: Chi-Square test

Using R programming, a chi-square test was performed on the data.

If p-value < 0.05, reject null hypothesis.

Results

p-value < 0.05

Decision:

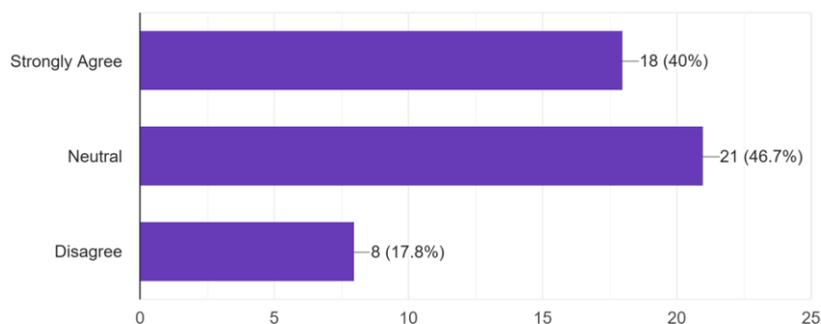Reject Null Hypothesis

Conclusion:

AI systems are biased.

Do you think AI systems can be biased?

45 responses

Strongly Agree	18 (40%)
Neutral	21 (46.7%)
Disagree	8 (17.8%)

```
R RGui
File   Edit   View   Misc   Packages   Windows   Help

R R Console

> observed <- c(39, 8)
> names(observed) <- c("Biased", "Not_Biased")
>
> observed
    Biased Not_Biased
        39          8
> chisq.test(observed)

        Chi-squared test for given probabilities

data:  observed
X-squared = 20.447, df = 1, p-value = 6.131e-06
```

Conclusion

Results from this study indicate there is some evidence showing that AI influenced decisions at various levels. Survey results, along with data generated from the hypotheses tested as part of this research, all showed that most people believe there is an existence of bias within AI systems. Existing AI systems contain bias due to various factors such as previous biases encountered by humans, different ways we developed the algorithms, and not including enough types of data (or diversity) in the sample data used to train AI systems. There are a great deal of documented case studies that demonstrate how the existence of AI has negatively impacted people; there are equivalent examples of how AI has had a negative impact on employment, policing, and finance. Ethical development practices are necessary when creating AI systems. To lessen bias to the use of AI systems through algorithm verification and being more forthcoming (i.e., developed in a transparent way) will help build public trust in how we use AI and should help inspire respect for AI's potential to supplement (rather than replace) human-based decision-making processes. Making ethical decisions while developing AI will result in equality, justice, trust, and social equity. The AI is Biased in decision making systems. It can be controlled through Data Transparency , Ethical rules , Better training data , Human supervision and Government regulations .

Limitations of the study

The study has the following limitations:

- The number of data collected was only 47. The study focused only on the urban academic population.
- The study relied only on self-reporting questions. The study only employed basic Chi-square statistical analysis.

References

1. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.

2. Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. *ProPublica*, 1–17.

3. Dastin, J. (2018). Amazon scraps a secret AI recruiting tool that showed bias against women. *Reuters*, 1–5.

4. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.

5. European Commission. (2019). *Ethics guidelines for trustworthy artificial intelligence*. European Commission, 1–41.

6. IBM. (2020). *Everyday ethics for artificial intelligence*. IBM Corporation, 1–36.