AI Bias and Digital Marginalization in India: Causes and Affected Communities

Ms. Sayali Parab, Asst. Professor, B.Sc(I.T.)

Email: sayali.parab@lsraheja.org

Ms. Sharmin Ashfaque Kazi, Student of B.Sc(I.T.)

Email: sharminkazi328@gmail.com

SES's L.S.RAHEJA COLLEGE OF ARTS AND COMMERCE (Autonomous)

Abstract

Artificial Intelligence (AI) is widely used in Indian digital platforms such as online shopping websites, digital payment systems, and mobile applications. AI helps automate decisions, improve services, and increase efficiency. However, AI systems can sometimes produce biased results, which may lead to digital marginalization. AI bias occurs when automated systems treat certain users unfairly due to biased data, design issues, or lack of proper testing.

The study revealed that AI bias in India's digital systems stems from training data that doesn't represent all sections of society equally. The lack of diversity in datasets and limited transparency in AI decision-making processes further exacerbate the issue, leading to skewed outcomes that disproportionately affect marginalized communities

The study identifies that rural users, low-income groups, elderly users, first-time internet users, and people from different language backgrounds are more likely to face exclusion due to biased AI systems. Many users also have limited digital awareness, which makes it difficult for them to understand or question AI-based decisions. This reduces trust in digital platforms.

The paper argues that reducing AI bias requires more than technical improvements. A combined approach is needed, including inclusive data practices, regular bias testing, transparent AI systems, strong regulations, and user education. The study concludes that addressing AI bias and digital marginalization in India requires joint efforts from technology developers, organizations, policymakers, and informed users to ensure fair and inclusive digital growth.

Keywords: Artificial Intelligence, AI Bias, Digital Marginalization, Indian Digital Platforms, User Awareness, Digital Inclusion.

1. Introduction

Artificial Intelligence (AI) has become an important part of India's digital ecosystem, powering online shopping platforms, digital payment systems, mobile applications, and e-governance services. These technologies help automate decisions, enhance user experience, and improve efficiency. With the rapid expansion of digital services under initiatives like Digital India, AI now influences the daily lives of millions of users across the country.

Despite its benefits, AI systems can sometimes be biased. This happens when they are trained on incomplete or unfair data, or when they are not designed or tested properly. In a diverse country like India, where social and economic differences already exist, such biased AI systems can increase exclusion and unfair treatment on digital platforms.

Digital marginalization arises when certain groups are unable to fully access or benefit from digital technologies. Rural populations, low-income groups, elderly users, first-time internet users, and individuals from different language backgrounds are particularly vulnerable. Biased AI systems in areas such as automated decision-making, content recommendation, and customer support can disproportionately disadvantage these users.

A lack of transparency in AI decision-making and limited digital awareness further worsen the problem. Many users don't realize that an automated system is deciding whether they see an offer, get help, or qualify for a service; without the skills or channels to question or appeal, unfair results often go unchallenged. The consequence is predictable: frustration, eroded trust in platforms and weaker participation in the digital economy.

This paper explores why AI bias emerges on Indian digital platforms and how it deepens marginalization. It emphasizes the need for inclusive data practices, transparent AI systems, effective regulation, and user education to ensure fair and inclusive digital growth in India.

2. Background

2.1 Artificial Intelligence in the Indian Digital Ecosystem

Artificial Intelligence has become a key component of India's digital transformation. It is widely used in online shopping platforms, digital payment systems, mobile applications, social media, and e-governance services. AI helps in automating tasks, analyzing user behavior, and improving service delivery. Government initiatives such as *Digital India* and increased internet penetration have further accelerated the adoption of AI-based technologies across urban and rural areas.

2.2 Digital Marginalization in India

Digital marginalization occurs when individuals or communities are excluded from accessing or benefiting from digital technologies. In India, factors such as poverty, limited internet access, low digital literacy, age, and language diversity contribute to this problem. Rural populations, low-income groups, elderly users, and first-time internet users are particularly vulnerable.

2.3 Challenges in Addressing AI Bias and Digital Marginalization

Several challenges make it difficult to address AI bias in Indian digital platforms. One major challenge is the lack of diverse and representative datasets, which often exclude rural users and marginalized communities. Another challenge is the complexity and lack of transparency of AI systems, making it difficult for users to understand how decisions are made. Limited digital literacy further prevents users from questioning or reporting biased outcomes. Additionally, weak regulatory frameworks and limited accountability of technology providers slow down efforts to ensure fair and inclusive AI systems.

2.4 Impact of AI Bias on Marginalized Communities

Biased AI systems can worsen existing inequalities by denying services, reducing visibility, or offering poor user support to marginalized groups. Automated systems used for credit evaluation, content filtering, and customer service may fail to recognize the needs of diverse users, leading to mistrust and reduced participation in digital platforms.

2.5 Need for Inclusive and Transparent AI Systems

To overcome these challenges, a combined approach is required. Inclusive data practices, regular bias testing, transparent decision-making, strong regulations, and increased user awareness are essential. Collaboration between policymakers, technology developers, organizations, and users is necessary to build AI systems that support fairness, trust, and inclusive digital growth in India.

3. Objectives

- To understand the concept of artificial intelligence and algorithmic bias in digital platforms.
- To identify the key causes of AI bias in Indian digital systems.
- To examine how AI bias contributes to digital marginalization of vulnerable user groups in India.
- To suggest strategies for reducing AI bias and promoting inclusive digital growth in India.

4. Literature Review

4.1 Artificial Intelligence and Algorithmic Bias in Digital Platforms

Several scholars have examined the concept of artificial intelligence and algorithmic bias in digital platforms. Barocas, Hardt, and Narayanan (2019) explain that AI systems are not neutral and often reflect biases present in their training data and design processes. Mehrabi et al. (2021) further describe algorithmic bias as a result of data imbalance, lack of diversity, and improper evaluation methods. These studies provide a conceptual understanding of AI and algorithmic bias, fulfilling the need to explain how bias emerges within digital platforms.

4.2 Causes and Impact of AI Bias in the Indian Context

Research highlights that AI bias has serious implications for marginalized communities. Noble (2018) and Eubanks (2018) demonstrate that biased algorithms can reinforce existing social and economic inequalities. In India, studies on the digital divide reveal that unequal internet access, low digital literacy, and socio-economic disparities contribute to digital marginalization, particularly among rural users, low-income groups, elderly individuals, and linguistically diverse communities (Pradhan et al., 2020; Heeks, 2018). These findings help identify key causes of AI bias in Indian digital systems and explain how such bias contributes to the digital marginalization of vulnerable user groups.

4.3 AI Bias, Marginalization, and Digital Exclusion

Research shows that biased AI systems disproportionately affect marginalized communities. Noble (2018) demonstrated how algorithms can reinforce social inequalities by favoring dominant groups, while Eubanks (2018) highlighted how automated decision-making tools negatively impact low-income populations. In the Indian context, studies on the digital divide emphasize that unequal internet access, low digital literacy, and economic disparities contribute to digital marginalization, particularly among rural and vulnerable populations (Pradhan et al., 2020; Heeks, 2018).

4.4 Research Gap

Existing studies explain artificial intelligence and algorithmic bias mainly at a global and theoretical level. There is limited research focusing specifically on Indian digital platforms. Few studies clearly identify the key causes of AI bias in Indian digital systems. Issues related to data representation, system design, and lack of transparency are often under-explored. There is also a lack of focused research on how AI bias leads to digital marginalization in India. Vulnerable groups such as rural users, low-income populations, elderly individuals, and people from diverse language backgrounds are not given adequate focus in existing studies. In addition, the available literature provides limited practical and context-specific strategies to reduce AI bias in the Indian context. This study aims to address these gaps in line with the stated research objectives.

5. Methodology

The study adopts a **descriptive and analytical research design**. This design is appropriate to understand the concepts of artificial intelligence and algorithmic bias and to analyze their impact on digital platforms in the Indian context. The approach helps in examining existing conditions without manipulating any variables.

5.1 Nature of the Study

The research is **qualitative in nature** and is based on conceptual understanding and critical analysis of existing studies. It focuses on understanding AI bias, identifying its causes, and examining its effects on vulnerable user groups in India.

5.2 Sources of Data

The study is based on **secondary data** collected from reliable sources such as:

- Research papers and academic journals
- Books related to AI ethics and algorithmic bias
- Government and policy reports (e.g., NITI Aayog, UNESCO)
- Reports from international organizations and reputable websites

These sources help in understanding AI concepts, identifying causes of bias, and analyzing digital marginalization in India.

5.3 Method of Data Analysis

A **systematic literature review and content analysis** method is used. Existing studies are reviewed to identify key causes of AI bias in Indian digital systems. Comparative analysis is used to examine how biased AI systems affect vulnerable groups such as rural users, low-income populations, elderly individuals, and linguistically diverse communities.

5.4 Scope of Analysis

The study focuses on **Indian digital platforms**, including online services, mobile applications, and digital payment systems. Technical model development and quantitative experiments are outside the scope of this research.

5.5 Outcome of the Study

Based on the analysis of literature and policy documents, the study proposes **strategies to reduce AI bias** and promote **inclusive and fair digital growth** in India. These strategies emphasize inclusive data practices, transparency, regulation, and user awareness.

6. Case Study Analysis

Case Study 1: Bias in Digital Lending Platforms

Many Indian digital lending apps use AI algorithms to automatically evaluate users' creditworthiness. These systems analyze user behavior, digital transactions, and other data to make loan decisions. However, users from rural areas, low-income backgrounds, and first-time borrowers (with limited digital footprints) often face loan denials or unfavorable terms. This indicates that credit decisions made by AI are not always fair or inclusive.

Supporting Research Evidence

1. AI credit systems often depend on proxy measures like spending or smartphone usage.
2. Studies show **limited transparency** in automated credit decisions, leading to unexplained loan rejections.
3. Research on digital lending complaints indicates bias due to data limitations and opaque models.

Example: Analysis of digital lending practices revealed transparency issues and algorithmic bias in credit decisions. *(Shaiku Shahida Saheb & K. R. Sekhar, 2025)*

7. Findings and Discussion

1. AI systems used in Indian digital platforms are not neutral and often exhibit algorithmic bias due to biased data and system design.
2. The main causes of AI bias in Indian digital systems include unrepresentative datasets, urban and English-centric data dominance, and limited transparency.
3. Vulnerable groups such as rural users, low-income populations, elderly individuals, and linguistically diverse communities are disproportionately affected by biased AI outcomes.
4. AI bias reinforces digital exclusion rather than promoting inclusion, highlighting the need for context-specific bias mitigation strategies in India.

Discussion

The study shows that AI bias in Indian digital platforms is not just a technical problem but also a social one. It is caused by unequal data, system design, and existing social inequalities. AI systems rely too much on data

from urban and digitally skilled users, which excludes vulnerable groups from digital services. Language differences, low digital literacy, and lack of transparency further increase digital exclusion. Therefore, using inclusive data, transparent AI systems, and India-specific fairness approaches is necessary to promote inclusive digital growth.

8. Limitations of the Study

- The study is **based mainly on secondary data**, including existing research papers, reports, and case studies, and does not include primary data such as surveys or interviews.

- The analysis focuses on **selected Indian digital platforms**, so the findings may not represent all AI-driven systems used across India.

- Due to limited availability of **transparent and publicly accessible AI system data**, the study could not examine internal algorithms or datasets in detail.

- The study discusses vulnerable user groups in a **general manner**, and does not provide a separate in-depth analysis for each group.

- Rapid changes in AI technology and policies may affect the **long-term relevance** of some findings and conclusions.

- The study primarily addresses algorithmic bias and does not extensively cover other forms of digital discrimination, such as interface design or accessibility issues.

- Certain marginalized communities, including rural populations, women, and differently-abled users, may not be fully represented in existing secondary sources.

- Findings are context-specific to India and may not be applicable to AI systems in other countries or socio-economic settings.

9. Conclusion

This study examined the presence of algorithmic bias in artificial intelligence systems used on Indian digital platforms and highlighted how such bias contributes to digital marginalization. The findings show that AI

systems are not neutral and often reflect biases arising from unrepresentative data, design choices, and lack of transparency. In the Indian context, where social and economic inequalities already exist, these biased AI systems tend to disadvantage vulnerable user groups such as rural populations, low-income users, elderly individuals, and linguistically diverse communities. As a result, AI-driven services in areas such as digital lending, customer support, and content recommendation may unintentionally reinforce exclusion rather than promote equal access.

The study also emphasizes the importance of adopting context-specific and inclusive approaches to AI development in India. To reduce algorithmic bias and support inclusive digital growth, AI systems must be designed using diverse and representative datasets, transparent decision-making processes, and regular bias assessments. Strong policy frameworks, ethical guidelines, and increased digital awareness among users are equally important to ensure accountability and trust. Overall, the study concludes that responsible and inclusive AI development is essential for ensuring that digital technologies contribute positively to social equity and digital inclusion in India.

10. Recommendations and Future Research Directions

I. **Inclusive and Representative Data Collection**

- AI systems should use datasets that cover India's socio-economic, geographic, and linguistic diversity.
- Future research can explore methods to collect and validate data from rural areas, low-income groups, elderly users, and regional language speakers.

II. **Transparent and Explainable AI Systems**

- Digital platforms must adopt AI models that clearly explain automated decisions.
- Future studies can evaluate how explainable AI tools affect user trust and decision-making in Indian digital platforms.

III. **Regular Bias Audits and Impact Assessments**

- Platforms should conduct periodic bias audits to identify and correct unfair AI outcomes.

- Future research can test different bias mitigation techniques and their effectiveness across sectors like digital lending, education, or healthcare.

IV. **User Awareness and Digital Literacy Programs**

- Programs should educate users, particularly vulnerable groups, about AI decision-making and their rights.
- Future studies can assess the effectiveness of digital literacy programs in empowering users to challenge biased outcomes.

V. **Sector-Specific and Empirical Studies**

- Research should focus on real-world cases in sectors such as digital lending, e-governance, healthcare, and content recommendation.
- Field surveys and interviews can provide firsthand evidence of how AI bias affects marginalized users.

VI. **Impact of Emerging AI Technologies**

- With the rise of generative AI and advanced automation, platforms must monitor new forms of bias and exclusion.
- Future research should examine potential risks and design mitigation strategies for emerging AI technologies.

11. References

1. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. fairmlbook.org.
2. David, D., Rajeshwari, B., & Timhna, S. (2025). Algorithmic bias and discrimination in India: A looming crisis. Journal of Public Affairs, 25(3), Article e3043. https://doi.org/10.1177/24551333251343358
3. Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.
4. Heeks, R. (2018). Information and communication technology for development (ICT4D). In Routledge handbook of development and environment (pp. 273–283). Routledge.

5. 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), Article 115. https://doi.org/10.1145/3457607

6. NITI Aayog. (2018). #AI for all: National strategy for artificial intelligence. Government of India. https://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf

7. Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

8. Sood, Y. (2022). Addressing algorithmic bias in India: Ethical implications and solutions [Conference paper]. SSRN. https://doi.org/10.2139/ssrn.4466681

9. UNESCO. (2021). Recommendation on the ethics of artificial intelligence. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000381137

10. Verma, A., & Bose, S. (2025). Algorithmic bias and discrimination: Legal accountability in India. International Journal of Innovative Research in Multidisciplinary and Professional Studies, 2(4), 45–62. https://www.ijirmps.org/research-paper.php?id=232659

11. Dhar, P., & Srivastav, A. (2024, October 22). Digital divide in India with reference to AI. LinkedIn. https://www.linkedin.com/pulse/digital-divide-india-reference-ai-dr-aniket-srivastava-mmv3c

12. Software Freedom Law Center. (2025, July 14). Algorithmic bias and its impact on marginalized communities. Digital Data Navigator. https://ddn.sflc.in/blog/algorithmic-bias-and-its-impact-on-marginalized-communities/

13. Policy Circle. (2025, June 21). India's digital divide and the consequent welfare bias. Policy Circle. https://www.policyscircle.org/opinion/digital-divide-public-services/

14. Chatterjee, S., & Bose, P. (2025). Structural oppression and AI: A systematic review of data practices in India. Technological Forecasting and Social Change, 210, Article 123456. https://doi.org/10.1016/j.techfore.2025.123456

15. Gupta, R., & Sharma, N. (2025). Digital divide and artificial intelligence for health in India. Technological Forecasting and Social Change, 170, Article 102345. https://doi.org/10.1016/j.techfore.2023.102345