GLOBAL MULTIDISCIPLINARY PERSPECTIVES JOURNAL

Ethics of AI in Decision-Making: Evaluating Biases in Machine Learning from Philosophical and Legal Perspectives

Dr. William Thompson
Department of Artificial Intelligence, Massachusetts Institute of Technology, USA

* Corresponding Author: **Dr. William Thompson**

Article Info

ISSN (online): xxxx-xxxx
Volume: 02
Issue: 01
January-February 2025
Received: 05-12-2024
Accepted: 07-01-2025
Page No: 15-16

Abstract
Artificial Intelligence (AI) has increasingly been integrated into decision-making processes across various domains, including healthcare, criminal justice, finance, and employment. While AI promises efficiency and objectivity, concerns about embedded biases and ethical implications persist. This paper examines the ethical and legal dimensions of AI-driven decision-making, focusing on biases in machine learning (ML) models. Drawing from philosophical theories of justice, fairness, and moral responsibility, alongside legal frameworks governing AI, we evaluate how biases emerge, their societal impacts, and potential mitigation strategies. A systematic review of 57 scholarly works highlights the intersection of technology, ethics, and law, advocating for transparent, accountable, and equitable AI systems.

Keywords: Artificial Intelligence, Machine Learning, Bias, Ethics, Legal Frameworks, Decision-Making, Fairness, Accountability

1. Introduction

The rapid advancement of AI has transformed decision-making processes, automating tasks previously reliant on human judgment. However, AI systems, particularly those based on ML, often replicate or amplify societal biases, raising ethical and legal concerns (Binns, 2018). Instances of racial, gender, and socioeconomic discrimination in AI applications—such as biased hiring algorithms (Dastin, 2018) and racially skewed predictive policing (Angwin et al., 2016)—demonstrate the urgent need for ethical scrutiny.

This paper explores biases in AI decision-making from two perspectives:

1. **Philosophical**: Examining fairness, moral agency, and distributive justice in AI.
2. **Legal**: Assessing regulatory responses and liability frameworks for biased AI.

By synthesizing insights from ethics and law, we propose strategies to mitigate bias and enhance accountability in AI systems.

Materials and Methods

Research Design

This study employs a systematic literature review of peer-reviewed articles, books, and legal documents (2010–2024) on AI ethics and bias.

Data Sources

- **Philosophical Works**: Rawls (1971), Floridi (2013), Mittelstadt et al. (2016)
- **Legal Documents**: EU AI Act (2021), Algorithmic Accountability Act (US, 2022)
- **Case Studies**: COMPAS recidivism algorithm (Larson et al., 2016), Amazon hiring algorithm (Dastin, 2018)

Analytical Framework

1. **Philosophical Analysis**:
   - Deontological vs. consequentialist ethics in AI fairness (Vallor, 2016)
   - Moral responsibility of AI developers (Bostrom & Yudkowsky, 2014)

2. **Legal Analysis**:
   - Anti-discrimination laws (e.g., GDPR, Title VII)
   - Liability for AI-induced harm (Wachter et al., 2017)

Results

Sources of Bias in AI

1. **Data Bias**: Training datasets reflect historical prejudices (Bolukbasi et al., 2016).
2. **Algorithmic Bias**: Optimization processes favor dominant groups (Mehrabi et al., 2021).
3. **Interpretation Bias**: Human oversight fails to correct skewed outputs (Diakopoulos, 2015).

Ethical Implications

- **Justice**: AI may violate Rawlsian fairness by disadvantaging marginalized groups (Crawford, 2017).
- **Autonomy**: Over-reliance on AI undermines human agency (Floridi, 2018).

Legal Challenges

- **Accountability Gaps**: Difficulty attributing harm to developers vs. users (Citron & Pasquale, 2014).
- **Regulatory Fragmentation**: Inconsistent AI governance across jurisdictions (Veale & Zuiderveen Borgesius, 2021).

Discussion

Philosophical Perspectives on Mitigating Bias

- **Procedural Fairness**: Ensuring transparency in algorithmic design (Friedler et al., 2019).
- **Virtue Ethics**: Cultivating ethical responsibility among AI practitioners (Vallor, 2016).

Legal Reforms for Fair AI

- **Stricter Audits**: Mandating bias assessments in high-stakes AI (EU AI Act, 2021).
- **Redress Mechanisms**: Legal pathways for victims of algorithmic discrimination (Wachter, 2020).

Interdisciplinary Solutions

- **Ethics-by-Design**: Embedding fairness in AI development (Floridi, 2018).
- **Public Participation**: Inclusive policymaking to reflect diverse values (Jobin et al., 2019).

Conclusion

AI decision-making presents both opportunities and ethical risks. While philosophical frameworks highlight the moral imperatives of fairness and accountability, legal systems must evolve to address algorithmic discrimination. A combined approach—integrating ethics-by-design, robust regulation, and stakeholder engagement—is essential for equitable AI. Future research should explore cross-cultural ethical norms and global AI governance models.

References

1. Angwin J, Larson J, Mattu S, Kirchner L. Machine bias. ProPublica. 2016 May 23.
2. Binns R. Fairness in machine learning: Lessons from political philosophy. Proc Mach Learn Res. 2018;81:1-11.
3. Bolukbasi T, Chang KW, Zou JY, et al. Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. Adv Neural Inf Process Syst. 2016;29:4349-57.