Ensuring Equitable AI: Mitigating Algorithmic Bias in Social EconomyApplications| Authored By : Ms. Saloni Shashank Patil, Research Scholar,Department of Law, Chhatrapati Shivaji Maharaj University, Panvel, Navi Mumbai | Volume VI Issue IV(July-August,2026), Available At Link

CLICK HERE TO READ DISCLAIMER

CLICK HERE TO SEE EDITORIAL BOARD

CLICK HERE TO SEE PUBLISHER DETAILS

Abstract

The artificial intelligence system creates social economic systems that span from financial inclusion to healthcare and from welfare to education while algorithmic bias serves as a threat that will increase social inequalities and create obstacles against sustainable development. The paper examines how algorithmic bias in AI systems shows various patterns which produce different effects on social economic systems that operate in India through its study of marginalized communities. The paper employs critical legal and ethical analysis methods to assess various case studies which include biased credit scoring and discriminatory healthcare AI and exclusionary welfare algorithms and language bias in education and equity gaps in green tech. The study reviews India's Digital Personal Data Protection Act of 2023, RBI digital lending guidelines, UNESCO AI Ethics Recommendations and EU AI Act.

The main findings show that biased training data and developer diversity gaps together with hidden algorithms and caste and gender proxy variables create systematic disadvantages for women, rural people and low-income individuals which results in loan rejections, incorrect medical diagnoses, welfare exclusion, educational and green technology access disparities. The Indian legal system contains essential regulatory deficiencies because it lacks compulsory algorithm assessments and does not establish adequate requirements for system transparency and its accountability system remains ineffective.

The paper introduces an equity-by-design framework which requires conduction of third-party bias audits together with diverse datasets and explainable AI for high-stakes decisions.The framework mandates community co-design and legal

accountability for discrimination. This study establishes the first India-focused algorithmic bias research which combines cyber law and fairness theory with social economy perspectives to study multiple sectors while providing practical solutions for inclusive development and ethical AI governance that supports SDG 1, 4, 5, 8, and 10.

Keywords: algorithmic bias, equitable AI, ethical AI governance, social economy, digital discrimination, inclusive innovation

1. Introduction

The use of artificial intelligence (AI) technology has grown to become essential for various social economic systems which include financial inclusion programs, public welfare initiatives, healthcare systems, educational institutions and green-technology governance operations.[1] The implementation of AI technology improves operational efficiency and expands service availability, but evidence shows that algorithmic bias creates unfair decision-making patterns which harm underrepresented communities and prevent inclusive sustainable progress.[2] AI-driven credit scoring systems and welfare distribution platforms in India, which operate similarly to systems in other countries, will maintain their existing social disparities unless their inherent biases receive proper identification and elimination.[3]

---

[1]The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India) (recognizing the expanding role of automated processing in public and private sectors).

[2] Ziad Obermeyer, David Powers, Christi Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 179 J. AM. MED. ASS'N 1 (2019) (documenting discriminatory outcomes from biased algorithmic decision-making).

[3] Reserve Bank of India, Master Direction – Digital Lending (2023) (highlighting risks of embedded

The research evaluates how algorithmic bias creates challenges for AI-based social economic systems in India while establishing an equity-based framework for artificial intelligence governance.[4] The analysis is driven by four documented issues:Through its Samagra Vedika program, the AI welfare system in Telangana prevented access to more than 1.86 million ration cardholders;[5]The Apple Card's AI credit scoring system resulted in lower credit limits for women who maintained identical financial conditions; [6] The United States mortgage-underwriting models demonstrated that Black and Latino applicants experienced higher rates of application denial; [7] AI diagnostic tools show poorer functional performance when assessing participants having darkened skin tone in comparison to their performance in other groups.[8]These cases show that bias creates significant consequences because it controls who can obtain their essential requirements for work, health care, public assistance programs, and educational access.[9]

The study has three objectives: the first objective aims to develop an understanding of how algorithmic bias functions in social economy AI systems, and the second

---

bias in digital credit and welfare platforms).

[4] INDIA CONST. art. 14, 15, 21.

[5] Al Jazeera, *Telangana's AI-Driven Welfare System Excludes 1.86 Million Ration Cardholders (2024)*, https://www.aljazeera.com.

[6]N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation 1–3 (Mar. 23, 2021), https://www.dfs.ny.gov.

[7] Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022).

[8]OBERMEYER ET AL., supra note 2.

[9]BARTLETT, MORSE &WHITMANE, supra note 7 (linking algorithmic bias to exclusion from essential services).

objective seeks to discover incomplete sections within legal and regulatory systems which include India's Digital Personal Data Protection Act 2023 and its digital-lending regulations, and the third objective intends to create specific rights-based solutions which regulators, developers, and policymakers can implement.[10] The paper begins with a literature review that covers algorithmic bias and AI ethics, and later presents a methodology section which describes the critical-legal and case-study approach and then shows data analysis results through five essential cases, and finally the discussion links research findings to inclusive development and ethical AI governance and AI applications in education.

2. Literature Review

2.1 Conceptualizing Algorithmic Bias

Algorithmic bias is commonly understood as systematic, unfair differentiation which produces different outcomes in artificial intelligence systems that disadvantage specific social groups.[11] Bias can emerge at multiple stages which include data collection as well as feature engineering and model training and deployment. O'Neil (2016) established the foundation for her research by describing most AI systems as "weapons of math destruction" which demonstrate how technical models transform historical bias into neutral assessment scores.[12]The latest studies demonstrate that bias

[10]The Digital Personal Data Protection Act, 2023, § 4–6, 10, No. 22, Acts of Parliament, 2023 (India); RESERVE BANK OF INDIA, supra note 3 (identifying regulatory gaps relevant to automated decision-making).

[11]Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV.671 (2016).

[12]CATHYO'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY(Crown Publishers 2016).

extends beyond technical aspects because it includes social hierarchy relationships and data governance decisions and institutional structure.[13]

Scholars identify multiple types of bias which include **statistical bias** that causes different error rates for various groups and **selection bias** which results from using unrepresentative training data and **aggregation bias** which hides group differences by showing only overall averages. The biases in social economy contexts result in actual damage through four specific outcomes which include excluded credit access, revoked welfare benefits, incorrect medical diagnoses and diminished educational chances.

2.2 Bias in FinancialInclusion and Welfare

AI-based credit scoring needs to develop new methods for assessing creditworthiness because it depends on using location, device type and transaction patterns as proxy variables which show racial, gender, caste and rural residency information.[14] Research shows that algorithmic models used by digital lending platforms in India and other countries continue to create credit barriers which predominantly affect women and rural borrowers who lack official credit records. Automated welfare systems have been shown to reject qualified recipients because their algorithmic rules fail to recognize actual community conditions and they treat spelling errors and

[13] SAFIYA UMOJA NOBLE, ALGORITHMS OF OPPRESSION: HOW SEARCH ENGINES REINFORCE RACISM (N.Y.U. Press 2018).

[14] Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022).

informal economic activities as legitimate criteria.[15]

The automated social-protection systems used in India and other countries operate without transparent processes and human monitoring while lacking proper methods for people to challenge decisions according to Amnesty International (2024) and other organizations.[16] The systems create higher risks for rights violations instead of improving access to services.

2.3 Bias in Healthcare and Education

AI diagnostic systems used in healthcare settings which were developed through training on light-skinned urban datasets demonstrate higher error rates when applied to dark-skinned patients especially from rural areas.[17] Public health initiatives, including *Ayushman Bharat* campaigns, show that these errors create higher health inequities instead of decreasing them. AI-backed tutoring systems and language educational tools used in educational settings show better performance for students who speak English as their primary language, which results in non-English speakers facing penalties while digital access gaps between groups of people become wider.[18]

2.4 Legal and Ethical Frameworks

The EU AI Act (2024) classifies AI used in credit scoring, welfare allocation,

---

[15] VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (St. Martin's Press 2018).

[16] Amnesty Int'l, Entity Resolution in India's Welfare Digitalization (2024), https://www.amnesty.org.

[17] Ziad Obermeyer, David Powers, Christi Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 179 J. AM. MED. ASS'N 1 (2019).

[18] RUHA BENJAMIN, RACE AFTER TECHNOLOGY: ABOLITIONIST TOOLS FOR THE NEW JIM CODE (Polity Press 2019).

healthcare, and education as "high-risk" which requires human monitoring and operational transparency and regulatory evaluations before implementation. [19] UNESCO's AI Ethics Recommendations (2021) establish fairness and accountability and transparency standards for AI systems which should centre around human needs while requiring bias audits and stakeholder participation in their development.[20] The Digital Personal Data Protection Act, 2023 (DPDP Act) of India establishes data consent requirements and data minimization rules however it does not require automated systems to maintain algorithmic fairness or explain their decision-making process.[21] Indian cyber law research identifies existing research gaps which relate to bias and algorithmic discrimination.[22] The existing literature gap provides a reason for the current study to merge legal-regulatory research with technical and social-economic analysis.

3. Methodology

3.1 Research Design

The study uses a qualitative research design which incorporates critical legal analysis together with ethical assessment methods and case study research methods which combine doctrinal research with policy evaluation. The research investigates

[19]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (AI Act), Annex III, pts. 5(a)–(b).

[20] UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021), https://unesdoc.unesco.org.

[21]The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).

[22]CascaGNUL, Algorithmic Discrimination and Indian Cyber Law (2025); JUSCORPUS, Gaps in India's AI Governance Framework (2025).

documented algorithmic bias cases in social economy systems to assess which existing legal and governance frameworks serve as suitable training and testing resources for artificial intelligence models. The research identifies research problems through exploratory research while it develops normative research solutions which identify harmful effects and their underlying causes.

3.2 Data Collection and Case Selection

The paper depends on academic articles and technical reports and policy documents as its secondary data sources. The study selects five key cases as representatives of various social economy fields.

a. **The Samagra Vedika algorithmic welfare exclusion case from India:**demonstrates how artificial intelligence technology results in welfare and ration card cancellations.[23]

b. **The Apple Card gender bias case from the United States:** demonstrates how AI credit scoring systems result in women receiving lower credit limits.[24]

c. **The United States Wells Fargo mortgage racial bias case:**This case study shows that automated underwriting systems often deny Black and Latino

[23] Amnesty Int'l, *Entity Resolution in India's Welfare Digitalization* (2024), https://www.amnesty.org; see also Al Jazeera, How an Algorithm Denied Food to Thousands of Poor in India's Telangana (Jan. 24, 2024), https://www.aljazeera.com.

[24] N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation (Mar. 23, 2021), https://www.dfs.ny.gov.

applicants more than others.[25]

d. **The AI healthcare diagnostic racial bias case (global):**This case study finds that dermatology tools make great level of mistakes when used on patients, with darker skin tones.[26]

e. **The AI education chatbot language and socioeconomic bias case (India/global):**shows that English-dominant urban models create disadvantages for rural students who do not speak English.[27]

These cases were selected because they have extensive documentation and public availability of information and their content covers multiple subthemes which include financial inclusion and welfare and healthcare and education and AI-supported green-tech access. The Indian and global contexts enable researchers to study cross-context comparisons between two different settings which enhance the policy relevance of their findings.

3.3 Analytical Framework

The analysis is structured around three layers:

---

[25] Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022).

[26]Ziad Obermeyer, Brian Powers, Christine Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 SCIENCE 447 (2019).

[27]UNESCO, State of the Education Report for India 2022: Artificial Intelligence in Education (2022), https://unesdoc.unesco.org; see also Bias Detection Framework for AI Chatbots Used in Education, RJPN J. (2024).

a. The technical layer studies how bias emerges through its presence in datasets and its effect on model design and feature selection.
b. The legal layer evaluates whether present regulations regarding non-discrimination, data protection and administrative law are being followed or not.
c. The ethical and policy layer examines how operations will affect fairness and accountability and environmental sustainability particularly through their impact on Sustainable Development Goalswhich includes SDGs 1, 4, 5, 8 and 10.

The paper uses a standardized case study template to analyse each case which includes five elements: (1) the AI system's background information;(2) the mechanisms of bias; (3) the group of people who were impacted; (4) the documented damages, and (5) the current or suggested regulatory solutions.

3.4 Limitations

The study uses publicly accessible secondary sources which prevent it from performing primary algorithmic audits and accessing proprietary model code. The research suffers from survivorship bias because it overemphasizes high-profile cases which receive media coverage. The chosen examples provide sufficient diversity to enable researcher to draw generalizable conclusions about equitable AI governance in social economic systems.

4. Data Analysis and Results

4.1 Samagra Vedika Welfare Exclusion Case

The Samagra Vedika platform of the Telangana government combined data from welfare programs to prevent fraudulent activities but used a hidden system that

matched citizen records between different government departments.[28] The system used automatic processes to detect duplicate records which it used to cancel or reject applications without needing thorough human examination. The platform incorrectly terminated 1.86 million ration cards while it turned down more than 1,40,000 applications which mainly affected low-income rural households and elderly individuals and city-based migrants.[29]

The bias mechanism used rigid identity matching (which treated minor name differences as fraudulent) and it restricted users from making appeals and had people verify information only through minimal local methods. The algorithm failed to recognize all common naming patterns which included joint family structures and informal documentation methods. The system created algorithmic exclusion which prevented families from accessing food security for several months and it resulted in one verified instance of death by malnutrition.[30]

This legal case demonstrates that India's digital-governance system lacks both explicit algorithmic fairness requirements and right-to-appeal protections. The DPDP Act 2023 establishes data management standards but it does not mandate organizations to provide complete justification for their automated welfare decision processes or conduct bias testing.[31] The case establishes a regulatory gap between data-protection standards and social-rights protection requirements.

---

[28]Amnesty Int'l, *Entity Resolution in India's Welfare Digitalization* (2024), https://www.amnesty.org.

[29]Al Jazeera, *How an Algorithm Denied Food to Thousands of Poor in India's Telangana* (Jan. 24, 2024), https://www.aljazeera.com.

[30]CascaGNUL, Victims of Algorithmic Harm in India's Welfare (2025).

4.2 Apple Card Gender Bias Case

The credit-scoring platform Apple Card which started operations in 2019 works with Goldman Sachs to create credit limits through its AI-powered assessment system.[32] Public complaints showed that women, including highly qualified professionals, received significantly lower limits than male counterparts with similar or better financial profiles.[33] The Financial Services Commission of New York conducted an investigation which revealed that the system design that claimed to maintain gender neutrality actually produced lower scores for female users.[34]

Analysis shows that the model used proxy variables because it needed to collect data about historical spending patterns and credit-history length which helped it identify the relationships between gender roles and caregiving responsibilities.[35] The case established itself as a major precedent which demonstrated how gender discrimination in algorithms created biased results within the financial services industry.[36] The United States government responded to this case by implementing stricter regulations

---

[31] The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).

[32] N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation 1–3 (Mar. 23, 2021), https://www.dfs.ny.gov.

[33] Incident Database, Incident 92: Apple Card's Credit Assessment Algorithm (2020), https://incidentdatabase.ai.

[34] N.Y. DEP'T OF FIN. SERVS., supra note 1, at 10–14.

[35] David Heinemeier Hansson, Twitter Thread (Nov. 7, 2019) (alleging 20× higher limit for husband than wife despite superior credit), cited in N.Y. Dep't of Fin. Servs., supra note 1, at 5.

[36] N.Y.DEP'T OF FIN. SERVS., supra note 1, at 20–22 (noting systemic transparency deficits and fairness concerns in algorithmic underwriting).

which controlled the use of automated lending systems.[37]

The Apple Card case demonstrates that Indian research should focus on how AI-based lending platforms will create gender discrimination in credit access because these systems lack proper bias audits despite existing legal prohibitions against such discrimination.Thus, the requirement of explainability standards and the need for third-party assessments which test fairness has been underscored.

4.3 Wells Fargo Mortgage Racial Bias Case

The research conducted by Bartlett and Morse together with Whitmane in 2022 discovered that Wells Fargo's mortgage underwriting system showed different rejection rates for Black and Latino applicants compared to white applicants who shared identical financial profiles.[38] The model used neighbourhood and historical lending data as risk indicators, which created a race-based proxy through geographic and credit history data patterns.[39]

The case shows that AI systems can reproduce historical housing discrimination which existed in the past.Even when there isn't a specific reference to race within an algorithm, it still discriminated in the way which denied people equal chances to acquire homes and the ability to accumulate wealth through capital.[40]

---

[37] Consumer Fin. Prot. Bureau, *Order Against Goldman Sachs Bank USA* (Oct. 23, 2024), https://www.consumerfinance.gov.

[38] Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022).

[39] See id. (demonstrating how zip-code and historical-credit proxies embed racialized risk assessments).

[40] See id. (arguing that algorithmic underwriting perpetuates structural wealth gaps).

The European Union (EU) AI Act classified high-risk financial artificial intelligence systems as regulated entities which needed to undergo human evaluation and compliance testing before they could be implemented.[41]

This case indicates to Indian policymakers that they must monitor proxy-based discrimination which occurs in AI-driven financial inclusion programs to prevent AI-based financial inclusion from increasing racial and caste-based discrimination.

4.4 AI Healthcare Diagnostic Bias Case

The research conducted on AI-based dermatology tools indicates that model performance decreases when testing on patients with darker skin who were not included in their primary training dataset which focused on light-skinned individuals.[42] AI health technologies face similar problems because their AI systems trained on urban high-income datasets show poor performance for rural and marginalized population groups.[43] The public health camps and telemedicine platforms use these tools which result in wrong diagnoses and treatment delays and create healthcare equality issues.[44]

Technical fairness metrics fall short of solving social equity problems according to

---

[41] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (AI Act), Annex III, pt. 5(b).

[42] Ziad Obermeyer, Brian Powers, Christine Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 SCIENCE 447 (2019); see also Daneshjou et al., *Disparities in Dermatology AI Performance on a Diverse, Curated Clinical Image Set*, 8 SCI. ADVANCESeabq6147 (2022).

[43] Hindustan Times, *India Risks Widening Healthcare Divide Without Ethical AI Safeguards*, Experts Warn (Feb. 13, 2026), https://www.hindustantimes.com.

[44] Int'l J. Cmty. Med. & Pub. *Health, Artificial Intelligence in Rural Healthcare in India* (2026),

this case study. The healthcare access disparities present in the datasets create structural healthcare access barriers. The system needs three protective mechanisms which include transparent model design and domain-specific bias testing and human-in-the-loop review.[45]

4.5 AI Education Chatbot Bias Case

The educational chatbots and tutoring platforms that use AI technology exhibit design flaws because they only serve English-speaking users who live in urban areas, which results in excluding students from rural India and those who speak Hindi, Tamil or other regional languages.[46] The reports about India's EdTech programs indicate that AI tutors provide inaccurate and less informative responses when they operate in non-English languages and that users with low-bandwidth connections and lower-spec devices experience reduced service quality.[47]

The existing bias manifests through two primary factors which include linguistic differences and socioeconomic backgrounds, thus perpetuating educational inequality. The case connects AI to green technology through its demonstration that AI-based environmental literacy tools and climate-education platforms maintain their focus on

---

https://www.ijcmph.com.

[45]OBERMEYER ET AL., supra note 1, at 450–52 (arguing for transparency, bias audits, and clinician oversight).

[46]UNESCO, State of the Education Report for India 2022: Artificial Intelligence in Education 9–11 (2022), https://unesdoc.unesco.org.

[47]Wadhwani Found., *AI in Education: Why Rural and Tier-3 India Risks Being Left Behind* (Dec. 29, 2025), https://wadhwanifoundation.org; see also Parth Mehrotra, *Algorithmic Bias in AI-Based Assessment*, 6 FAST FORWARD JUST. (May–June 2026).

urban areas, which results in rural communities being neglected.[48]

5. Discussion

The case studies examined in this paper show that algorithmic bias occurs in AI-driven social economy systems as a persistent problem which exists beyond technical faults and rare exceptional situations, and it continually harms specific groups of people. The system creates a systematic disadvantage for women, rural people, low-income individuals, marginalized castes, ethnic groups and individuals who have restricted digital access.The AI system uses historical exclusion patterns which it hides behind a facade of neutrality and efficiency to create exclusion in welfare, credit, healthcare and education systems.The Samagra Vedika case demonstrates that an algorithmic welfare system can prevent thousands of qualified families from obtaining food assistance. The Apple Card and Wells Fargo cases show that financial-inclusion platforms treat women and minority applicants unfairly because they deny them credit access despite having similar or better financial qualifications. The educational and healthcare cases demonstrate that AI systems which operate with biased or limited training data deliver inferior performance to users who have darker skin tones or live in rural areas or speak languages other than English.

The main discovery shows that current legal systems lack the necessary tools to handle algorithmic bias problems that exist in these specific fields. The *Digital Personal Data Protection Act* (2023) of India establishes requirements for consent data minimization and purpose limitation but does not mandate organizations to

[48]UNESCO, International Forum on AI and Education: Ensuring AI as a Common Good to Transform

implement algorithmic fairness controls or conduct bias assessments or provide customers with explanation rights for AI-based decisions. The lack of such regulations permits AI technologies to function in welfare and finance and health and education sectors as hidden "black boxes" which prevent individuals from knowing how to seek justice. The *Digital Lending Guidelines* (2022) published by the Reserve Bank of India require lenders to protect their customers through transparent practices but stop short of making fairness assessments mandatory or asking lenders to reveal how their AI systems make decisions.[49] The existing data protection rules create a compliance gap between protection requirements and actual rules for non-discrimination.[50]

The results demonstrate that the issue exists as a worldwide problem.The Apple Card and Wells Fargo cases together with the EU AI Act (European Parliament 2024) demonstrate that jurisdictions which possess strong consumer and data protection laws face difficulties in managing algorithmic bias control.[51]The European Union's

---

Education, Synthesis Report 12–14 (Dec. 7–8, 2021), https://unesdoc.unesco.org.

[49]Reserve Bank of India, Guidelines on Digital Lending (Sept. 2, 2022), https://www.rbi.org.in (requiring Key Fact Statements and data-consent protocols but not mandating algorithmic fairness audits).

[50] The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India) (establishing data-minimization and consent rules without explicit non-discrimination or explainability mandates for automated lending).

[51]N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation (Mar. 23, 2021), https://www.dfs.ny.gov; Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022); Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (AI Act), Annex III, pt. 5(b).

method of designating AI-driven credit scoring and welfare distribution and medical services and educational systems as high-risk artificial intelligence systems requires all systems to undergo human monitoring and complete transparency of their functions and accuracy assessments and compliance testing before they can be used.India can learn from these cases and improve its AI policies because it needs to create new laws that treat algorithmic discrimination as a separate legal violation. The country needs to establish protection measures for AI systems that will be used in social economy applications.

The policy and managerial perspective show multiple specific interventions which need to be executed.

a. **Mandatory algorithmic bias audits:** It require independent third-party auditors to assess all welfare, lending, healthcare and education AI systems before their deployment to determine their effects on protected groups. The audits need to include disaggregated performance metrics which measure different groups through their gender and rural/urban status and language and caste and all other applicable categories and the audits need to be conducted at regular intervals.

b. **Equity-by-design principles:** Fairness considerations should start at data collection then proceed to feature selection and final model design according to AI developers and public sector agencies. The system needs diverse datasets which include specific bias-mitigation techniques that use re-weighting and adversarial debiasing methods together with a human-in-the-loop approach for reviewing essential decision-making processes.

c. **Explainability and right-to-appeal:** The affected parties should understand the reasons behind the AI system's decision to reject their loan application and welfare assistance request and medical diagnosis assessment whereas they should have the ability to bring their case before a human decision maker. AI educational platforms need to deliver understandable explanations which show their assessment and recommendation procedures.

d. **Community participation:** AI design and evaluation processes need to include participation from marginalized communities. Participatory design enables AI tools to mirror local conditions while honouring cultural and linguistic variation and safeguarding vulnerable populations.

The Digital India program together with the AI Task Force which operates under the National Strategy for AI and the pilot AI initiatives targeting health and welfare services in India create an opportunity to implement algorithmic fairness throughout the country. India currently lacks a specific law which addresses AI fairness yet the DPDP Act, 2023 together with sectoral guidelines can be used to establish methods for addressing algorithmic discrimination. A national framework based on the EU AI Act which has been modified to suit India's federal system and social protection framework would mandate the execution of bias impact evaluations together with public interest algorithm audits and participatory oversight groups. Such a framework would position India as a leader in ethical AI for inclusive development, aligning with

global AI-for-sustainable-development agendas.

The following practical recommendations outline the procedures which essential stakeholders should follow to implement the equity-by-design framework which this paper presents:

a. Policy-makers must establish AI welfare systems and AI lending systems as high-stakes projects which need to undergo bias testing before their implementation and require systems which can clarify their functioning.

b. Governments and NGOs should work together with rural communities and women-led organizations to design AI tools which will help build trust while preventing their exclusion from technological developments.

c. EdTech and green-tech companies need to develop AI systems which support multilingual capabilities and operate effectively in low-bandwidth environments to assist learners from marginalized backgrounds and communities with limited energy resources.

The conference objectives on inclusive social and economic development, ethical AI governance, AI in education and sustainable development receive direct support from these recommendations. AI systems need to incorporate fairness and accountability and transparency as essential elements which regulators and practitioners must

implement to achieve SDGs 1 (No Poverty), 4 (Quality Education), 5 (Gender Equality), 8 (Decent Work and Economic Growth) and 10 (Reduced Inequalities). The paper presents researchers with an extensive future research agenda which includes studying AI bias through empirical research in India's AI backward welfare and education programs, conducting comparative studies of AI governance frameworks across different jurisdictions and creating AI fairness assessment tools designed specifically for Indian cultural contexts.

6. Conclusion

The research demonstrates that algorithmic bias in AI-powered social economy systems creates a major danger which obstructs inclusive sustainable development processes in India because millions of people rely on AI-based welfare systems along with credit and healthcare and education programs. The research demonstrates through five case studies which include Samagra Vedika, Apple Card, Wells Fargo, healthcare diagnostics and AI-based education that technical and social and legal elements work together to create discriminatory results. The findings show that there exists a substantial mismatch between existing legal and regulatory systems and the requirements needed to achieve fair AI governance.

The proposed equity-by-design framework provides detailed guidance which includes mandatory bias audits and the collection of diverse and representative datasets and the use of explainable AI technology in high-risk fields and the active involvement of impacted communities through design processes and the establishment of strong legal responsibility mechanisms. The social economy will benefit from these procedures

which will guarantee that AI systems contribute to inclusive growth and gender equality and poverty reduction and fair educational opportunities instead of worsening current inequalities.

Future research should validate these recommendations by conducting algorithmic audits of Indian AI welfare and lending platforms or by performing cross country studies of AI governance systems. The main point for policymakers to understand is that ethical AI constitutes an essential requirement for achieving sustainable development.

7. References

Books:

cathyo'neil, weapons of math destruction: how big data increases inequality and threatens democracy(Crown Publishers 2016).

safiyaumoja noble, algorithms of oppression: how search engines reinforce racism(N.Y.U. Press 2018).

virginiaeubanks, automating inequality: how high-tech tools profile, police, and punish the poor(St. Martin's Press 2018).

ruhabenjamin, race after technology: abolitionist tools for the new jim code(Polity Press 2019).

Journal Articles:

Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 calif. l. rev. 671 (2016).

Ziad Obermeyer, Brian Powers, Christine Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 science

447 (2019).

Robert Bartlett, Adair Morse & Nancy Whitmane, *Consumer-Lending Algorithms and Racial Disparities in Mortgage Underwriting* (2022).

Daneshjou et al., *Disparities in Dermatology AI Performance on a Diverse, Curated Clinical Image Set*, 8 sci. advanceseabq6147 (2022).

Parth Mehrotra, *Algorithmic Bias in AI-Based Assessment*, 6 fast forward just. (May–June 2026).

Statutes & Regulations:

INDIA CONST. arts. 14, 15, 21.

The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (AI Act), Annex III, pts. 5(a)–(b).

Government / Agency Reports:

Reserve Bank of India, Master Direction – Digital Lending (2023).

N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation 1–3 (Mar. 23, 2021), https://www.dfs.ny.gov.

Consumer Fin. Prot. Bureau, Order Against Goldman Sachs Bank USA (Oct. 23, 2024), https://www.consumerfinance.gov.

News & Online Reports:

Al Jazeera, *How an Algorithm Denied Food to Thousands of Poor in India's Telangana* (Jan. 24, 2024), https://www.aljazeera.com.

Amnesty Int'l, *Entity Resolution in India's Welfare Digitalization* (2024), https://www.amnesty.org.

Hindustan Times, *India Risks Widening Healthcare Divide Without Ethical AI Safeguards*, Experts Warn (Feb. 13, 2026), https://www.hindustantimes.com.

Wadhwani Found., *AI in Education: Why Rural and Tier-3 India Risks Being Left Behind* (Dec. 29, 2025), https://wadhwanifoundation.org.

Incident Database, *Incident 92: Apple Card's Credit Assessment Algorithm* (2020), https://incidentdatabase.ai.

Conference / Research Reports:

UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021), https://unesdoc.unesco.org.

UNESCO, State of the Education Report for India 2022: Artificial Intelligence in Education 9–11 (2022), https://unesdoc.unesco.org.

UNESCO, International Forum on AI and Education: Ensuring AI as a Common Good to Transform Education, Synthesis Report 12–14 (Dec. 7–8, 2021), https://unesdoc.unesco.org.

Int'l J. Cmty. Med. & Pub. Health, Artificial Intelligence in Rural Healthcare in India (2026), https://www.ijcmph.com.

Other (Working Papers, Blogs, Social Media):

CascaGNUL, Algorithmic Discrimination and Indian Cyber Law (2025).

JUSCORPUS, Gaps in India's AI Governance Framework (2025).

CascaGNUL, Victims of Algorithmic Harm in India's Welfare (2025).

David Heinemeier Hansson, Twitter Thread (Nov. 7, 2019) (alleging 20× higher limit

for husband than wife despite superior credit), cited in N.Y. Dep't of Fin. Servs., Report on Apple Card Investigation at 5.

Bias Detection Framework for AI Chatbots Used in Education, RJPN J. (2024).