



ALGORITHMIC BIAS AND THE RIGHT TO FAIRNESS: A LEGAL AND SOCIAL INQUIRY INTO AI IN EDUCATION AND PUBLIC ADMINISTRATION

Sanika Dehury* Dolly Nayak*

There is growing worry about algorithmic bias in Artificial Intelligence (AI) systems, particularly in education and public administration in India. With algorithms providing AI grading, the distribution of welfare, predictive policing, and other applications, bias and opacity in decision-making will only reinforce social inequalities. The paper assesses how algorithmic bias contravenes certain constitutional protections - especially Articles 14, 15, and 21 of the Indian Constitution - without transparency, accountability, and the possibility of legal remedy. Using globally comparative frameworks like the EU AI Act and Canada's Directive on Automated Decision-Making, the article illustrates that there is currently no comprehensive regulation development in India. With real-world evidence, such as exclusion due to Aadhaar and educational algorithm bias, the paper makes a case for the immediate need for algorithm audits and impact assessments, and the inclusion of human-in-the-loop. The article concludes with recommendations to consider fair outcomes, dignity, and democratic oversight when developing legislation and policy related to AI's deployment in contemporary public governance systems in India.

The swift incorporation of Artificial Intelligence (AI) into public administration and education is a seismic shift in the delivery of services and decision-making activities by states and institutions. Forever relegated to theoretical research or limited application, we now see AI systems perform service delivery tasks such as predictive policing, automated welfare distribution, and algorithmic grading in schools and universities. Proponents of these technologies demand that AI replace inefficient, subjective, slow processes that are subject to human errors and bureaucratic delay. With this technological potential, however, there is a growing amount of research that demonstrates the undesirable and often discriminatory consequences of increasing algorithmic decision-making. Algorithmic bias is one of the most

*BA LLB (HONS.), SECOND YEAR, NATIONAL LAW UNIVERSITY, ODISHA.

*BBA LLB (HONS.), SECOND YEAR, NATIONAL LAW UNIVERSITY, ODISHA.

important considerations: AI outputs systematically shift in one direction with biased training data, incorrect design assumptions, and/or structural inequalities baked into the digital infrastructure. Upon fair consideration, these biases are not usually intentional, and while the generative effect can be everything from a little to a lot, to marginalised communities, the effects can be overwhelming as the biases feed into each other in ways that seem zen and are hard to disentangle. Unlike a human decision-maker, algorithms do not offer an explanation about why something is decided, cannot be interrogated like a co-actor, and function typically as a 'black box', raising questions about accountability and transparency.

The risks in India are particularly high. The vast socio-economic diversity present in India, compounded by persistent historical forms of inequality informed by caste, gender, and linguistics/geography, has placed us in a position where there is the potential for AI systems to replicate inequality instead of alleviating it. For instance, an AI system trained on data that embodies years of biased policing or biased assessment and, thus, represents the inequities that also occur with deployment in the real world, could reinforce the inequities to alleviate them. The deployment of uncritically guided innovations by public institutions risks normalising systemic exclusion under a veil of algorithmic neutrality, all presumed to be adopting modern technologies and practices. The Indian Constitution provides a strong normative basis in support of establishing fair and equal treatment of persons. Article 14 (Right to Equality),¹ 15 (Prohibition of Discrimination),² and 21 (Right to Life and Personal Liberty)³ in explicit terms impose on the state an obligation to treat all persons in a non-arbitrary, fair, and transparent manner. Yet, the introduction of algorithmic governance raises fundamental questions: Will automated systems be answerable to constitutional principles? Will citizens have a right to question algorithmic decisions? Will our existing legal frameworks suffice in dealing with the harms associated with algorithmic and digital systems?

These important considerations are examined by examining the overlap of AI with law and social justice in the Indian context. The paper focuses on two vulnerable areas — we will see algorithmic bias present real threats to fair play in the education and public administration sectors. By drawing upon real-world incidents, legal provisions, and international comparators, this article explores the sufficiency of India's current legal protections to address the algorithmic nature of its farming future. The paper concludes with legal and policy reform

¹ Constitution of India, Art 14

² Constitution of India, Art 15

³ Constitution of India, Art 21

proposals that will ensure technological development does not come at the cost of human dignity or constitutional rights. In a period of rapid technology development and application within education and governance, the immediate need is not just for smarter algorithms, but products of fairer systems. Transparent systems, accountable systems, systems where justice is at the heart of the reasoning - systems that reflect the core values of an Indian democracy.

Algorithmic bias exists when artificial intelligence (AI) produces systematically biased predictions because of what it learned from underlying assumptions in the algorithms or flawed practices in their construction. Algorithmic bias arises from multiple sources. Data bias occurs when historical datasets reflect existing prejudices and injustices of society, which the AI speeds up as it continues to predict. The bias that arises because of design occurs when people who design and engineer algorithms have their values and presumptions within the algorithms they test, which can prioritise one outcome above another. Structural bias occurs when models are built on datasets that take the majority or privileged group(s) as their normative population bases, which can have the effect of erasing minority and underrepresented communities. A well-known example in the United States is the criminal justice COMPAS algorithm that was found to systematically label African American defendants at a higher risk than their White counterparts -- behaviours that fed the systemic racial prejudices. Although the use of AI in Indian public administration is still at an early stage of development and use, there is an increasing chance that similar issues will arise as algorithms are often trained on incomplete datasets or datasets that are socio-culturally biased about the diversity of India's population. This is coupled with the increasingly complex issue created by the "black box" nature of many AI systems, where algorithms do the collecting of data as well as decision-making in ways that remain opaque regarding inputs, decision-making, and logic. Thus, individuals affected by these algorithmic decisions have little to no opportunity to know, challenge, or appeal the reasons behind negative outcomes. As is clear, the lack of transparency and accountability raises serious constitutional questions, especially since India's democracy is based on fairness, the right to be heard, and non-arbitrary action by the state. Without transparency and protections, AI can exacerbate existing inequity and infringe on citizens' rights.

The legal system in India, of which the Indian Constitution is the highest legal system and the law of the land, demands equality and fairness as principles. Article 14 of the Indian Constitution guarantees equality before the law and equal protection under the law. Article 15 prohibits discrimination against Indians based on religion, race, caste, sex, or place of birth.

Article 21, which protects an individual's right to life and personal liberty, was broadly defined by the Supreme Court of India to include rights to dignity, privacy, and procedural fairness. In the important case of *Maneka Gandhi v. Union of India* (1978),⁴ the Court expanded the scope of Article 21 to include natural justice and due process of law. These constitutional safeguards are increasingly under attack in the context of the state deploying Artificial Intelligence (AI). AI systems may mediate the relationship between states and individuals in ways that threaten the doctrine of non-arbitrariness, which is central to Indian constitutionalism, when they make subjective decisions on behalf of the state that deprive individuals of welfare benefits, grade students, and profile suspects in ways that are non-transparent and immune to appeal. The repose of the individual about the authority (i.e. the state) is not protected if a citizen is treated inequitably or unreasonably by a non-transparent algorithm and the citizen has no practical avenue of appeal. The state may be estranged from its constitutional responsibility. The Right to Information, 2005⁵ permits citizens to request access to information from public authorities, but its value is diminished when public authorities outsource decision-making responsibility under legislation to private algorithms that may be non-transparent and outside the fields of public accountability. India does not have a specific regulatory regime, despite strong constitutional protections, for AI use in the public sector. These regimes will raise substantial problems of enforcement or incidentally be limited to enforcement when AI-inflected decision-making produces inequality or arbitrariness that can raise immediate concerns about rights, accountability, and digital justice within a democratic state.

Artificial Intelligence is rapidly growing in importance in education, to the point where it is now being used for grading, adaptive learning, online exams and assessments, and admissions. The recent COVID-19 pandemic accelerated the use of machine-based algorithmic tools worldwide as many educational institutions searched for alternatives to in-person exam environments. A tragic example of the potential risks of using an AI algorithm is the United Kingdom's 2020 A-level grading debacle, where an algorithm not only downgraded students from disadvantaged or failing schools at a disproportionate rate but at the same time also inflated the scores of private school students, which led to an uproar from the public and litigation, forcing the English authorities to abandon the algorithm. Although India has not yet been confronted with a grading fiasco similar to the United Kingdom, there are some disturbing issues arising. For example, some technologies for online proctoring in universities and law

⁴ *Maneka Gandhi v Union of India* AIR 1978 SC 597

⁵ Right to Information Act 2005 (India)

schools have flagged students with darker skin and facial coverings, i.e., hijabs, but even more troubling is the potential coding of racial and religious bias in the technology. Furthermore, while EdTech platforms based on AI emerge as new assessment methods to help track performance and offer students individualised learning experiences, they often privilege students studying in urban contexts who speak English, rather than those from vernacular or regional contexts. The use of opaque algorithms in educational assessment also raises serious constitutional legal issues because often the judicial agents making decisions are not consistent in how they apply the law. Unequal treatment by flawed or biased algorithms violates the principle of equality, which Article 14 guarantees. In addition, where procedural fairness is lacking in automated decisions, educational measures can violate Article 21, which provides for "the right to life, or personal liberty" including "mental well-being". Further, biased surveillance of education practice may lead to inappropriate academic conduct adjudications or discipline measures against students that interfere with students' rights to express themselves freely and privacy (Article 19(1)(a)). The state must place regulations around the use, features, and practices of educational AI platforms to ensure the principles of constitutional propriety are respected and student rights are protected, thereby requiring algorithmic accountability, explainability, audibility, and human decision-makers where a public authority or publicly funded institution (Hernández, 2023).

The increased incorporation of Artificial Intelligence (AI) into public administration represents a shift in governance, thereby presenting opportunities for greater efficiency and data-driven decision-making, particularly in welfare provision and law enforcement. However, this presents substantial risks, particularly when social context and human rights are disregarded. One of the more grievous examples of algorithmic exclusion in India is Aadhaar, and the use of biometric authentication to distribute welfare. These systems are intended to facilitate efficient disbursement of welfare, but often do not recognise the fingerprints or iris scans of vulnerable people, who can be seniors, people with disabilities, or manual labourers, effectively blocking their access to essential services. The challenges are compounded with AI-driven automation in detecting fraud and adjudicating eligibility for interventions, which can be difficult to gain insight into and have limited recourse for appeal. In 2019, an investigation by *The Hindu* showed in states such as Jharkhand, how automated data-matching processes excluded legitimate beneficiaries from the Public Distribution System (PDS), with little evidence that potential beneficiaries were communicated with to name a meaningful opportunity to appeal, violating their right to food and dignity at a time of need. Indian police

agencies are also weighing the use of predictive policing instruments that rely on historical crime data to predict future crimes. However, predictive policing instruments are trained on biased data - data that reflects caste, class, or communal biases - and can exacerbate systemic discrimination by categorising marginalised persons as "repeat offenders" in far higher numbers than other complainants. In the historic ruling in *Justice K.S. Puttaswamy v. Union of India* (2017),⁶ the Supreme Court recognized privacy as a fundamental human right and incorporated the right to privacy to impose constraints on state surveillance through three principles - legality, proportionality, and necessity, which may not be respected at all for automated surveillance systems. The implications of profiling on an arbitrary and discriminatory basis without protections and safeguards, and automating enhanced AI proliferation with software that puts supremacy over equity are serious. Without transparent, accountable, and human oversight, our society will increasingly privilege AI for governing equity over equality for the unprecedented, marginalised communities that become voiceless and have the least ability to access effective legal remedies.

The issues associated with holding algorithmic systems accountable under Indian law are complex and arise from several legal and institutional shortcomings. First and foremost, the code is closed. Many AI systems developed by private technology firms are protected by commercial trade-secret laws, which practically insulate their logic from the public and even the courts. In the absence of any access, it is typically impossible to discern if a decision was just, lawful, or discriminatory. In addition, there is no jurisprudence dealing with algorithmic bias or automated decision-making. For example, while courts in India have entertained cases that arise from algorithmic systems, they have not engaged the larger implications of algorithmic decision-making, which leave a legal vacuum entirely devoid of standards or precedent as to what can be considered fair, lawful, or due process and accountability in digital contexts. In addition, rather than addressing some of the systemic harms associated with algorithmic discrimination, India's existing data protection framework, such as the Digital Personal Data Protection Act, 2023,⁷ focuses only on accumulating data storage, consent, and related processing. It places no restrictions on algorithmic explainability or fairness and has not conferred a right to appeal or contest automated decisions. Relatedly, we should also worry about institutional inertia. Statutory agencies such as the Unique Identification Authority of India (UIDAI), and the Central Board of Secondary Education (CBSE) have all made use of

⁶ *Justice K S Puttaswamy v Union of India* (2017) 10 SCC 1

⁷ Digital Personal Data Protection Act 2023 (India)

automated systems with different issues; either simply not broaching them, or minimizing their accountability in cases in which systems fail, or in which an individual is harmed as a result of using those systems. In that context and with very few avenues of recourse available for individuals, and most importantly, individuals lower down the socio-economic ladder, algorithmic systems reside in a legal and ethical vacuum; individual citizens bear the brunt of the failures or harms of systems, without recourse or accountability arising because of failures.⁸

Different areas of the world have acknowledged algorithmic bias and are trying to initiate laws and policies aimed at mitigating harm from algorithmic bias. The European Union, through its proposed EU AI Act, has implemented a new risk-based framework that would place extensive obligations on high-risk AI applications, specifically regarding sensitive areas such as education, welfare, and law enforcement.⁹ Canada, through the Directive on Automated Decision-Making, particularly requires federal agencies to undertake algorithmic impact assessments of AI systems before they are deployed, as well as create a framework of transparency and risk evaluation.¹⁰ Although the federal government has not imposed a unified approach in the US, several local jurisdictions have begun to address algorithmic bias. For example, since 2022, New York City has required regular audits of AI-powered hiring tools to identify and mitigate bias.¹¹ Brazil and Kenya, both emerging digital democracies, have started national dialogues related to digital rights and considerations related to the impact of algorithmically powered decision-making in public systems. Global initiatives such as these provide plenty of examples for India to draw from, especially since no specific framework to regulate AI implementation in government processes currently exists in India. India could build upon these models to create a regulatory framework that ensures transparency in AI procurement and use by the public sector, provides for independent algorithmic audits, guarantees human oversight in high-stakes decision-making that affects rights and entitlements, and provides for accessible forms of legal redress for individuals harmed as a result of the use of AI technology. A regulatory framework is necessary not only to fulfil the

⁸ The Hindu, 'Jharkhand's Aadhaar-linked Public Distribution System Leaves Out Many Poor' The Hindu (12 July 2019) <https://www.thehindu.com/news/national/other-states/jharkhands-aadhaar-linked-public-distribution-system-leaves-out-many-poor/article28398081.ece> accessed 18 May 2025

⁹ European Commission, 'Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)' COM (2021) 206 final

¹⁰ Government of Canada, Directive on Automated Decision-Making (2019) <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/automated-decision-making.html> accessed 18 May 2025

¹¹ New York City Council, Local Law No 144 of 2021 (Bias Audit Law for Automated Employment Decision Tools)

constitutional requirements of fairness and equality but also to promote public confidence in AI technologies, enabling modern forms of governance. If left unchecked, the harms of AI and systemic inequalities run the risk of outpacing any benefits.

India needs to quickly address specific legal and institutional reforms if constitutional values like equality, fairness, and accountability are to be upheld and safeguarded in an age of Artificial Intelligence. First, Indian law must provide an explicitly engaged definition of algorithmic discrimination now in the law, especially where artificial intelligence is employed as part of a public function, such as welfare provision, policing, or education. This will establish an enforceable basis and a basis for remedial action. Second, the law must require algorithmic impact assessments before commencing the deployment of AI systems, especially in governance or education, and before commencing to assess the risk to human rights, fairness, and communities that are already marginalised. Third, individuals adversely affected by automated decisions should have a "right to explanation" that enables them to understand the logic or rationale of automated decisions when they occur and access to appeal or correction. Fourth, and be that as it may, a human-in-the-loop mandate is likewise necessary for all high-stakes decisions to ensure human judgment is preserved in contexts of significance to entitlements, rights, or livelihoods. Fifth, India needs to create a standalone AI regulatory body that will supervise the ethical development of AI, certify systems in terms of fairness, and ensure adherence to constitutional tests. Finally, public participation must be systemic: civil society, technologists, legal experts, and communities impacted should shape, audit, and monitor the deployment of AI systems used in public services. Without inclusive oversight and democratic safeguards to govern AI deployment, AI use could exacerbate structural inequalities and violations of civil liberties. The reforms... all of them, would not just bring the governance of AI into the spirit of India's constitutional framework but also provide a greater degree of transparency, accountability, and public trust in the automated decision-making process.

As India moves forward in realising the potential of Artificial Intelligence in both education and public governance, the country faces a key juncture in addressing the question of whether technological advances will deepen injustices or be appropriated in pursuit of justice and equity. This paper demonstrated that algorithmic systems, which tend to develop in opaque functions, without adequate consideration of socio-legal constraints, threaten key constitutional principles. Regardless of the evidence of problems from AI-regulated grading, to AI-based

surveillance, and its exclusionary effects on beneficiaries of the welfare system or law enforcement powers, what is apparent is that unchecked algorithmic bias can undermine the most fundamental rights contained in Articles 14, 15, and 21 of the Constitution of India. India already faces insufficient legal and institutional frameworks to govern in a digital landscape. The lack of algorithmic accountability, human oversight, and institutional redress mechanisms places all citizens, particularly marginalised communities, at the mercy of automated systems that are arbitrary and unjust. Learning from other countries, India must move quickly to establish a regulatory ecosystem that involves transparency, fairness, and accountability obligations to govern all AI in the public interest. AI must codify a spectrum of rights, including a right to explanation and algorithmic impact reviews before implementation, but also ensure that human judgment replaces machine judgment in decision-making when human beings are likely to be impacted and implement inclusive and transparent approaches to accountability. The path forward must integrate technological progress with constitutional ethics, ensuring that efficiency does not come at the expense of dignity. Ultimately, the introduction of AI in systems of public value should not be framed merely as a technical issue, but as an important public-facing issue in a democratic society. The state will need to take on the responsibility to ensure that automation does not replace accountability and that governments informed by data are still people-centric. Only through principled and participatory regulation will we see India's future use of technology to uplift citizens rather than further marginalise them, and digital development and digital infrastructure that work with us to build an equitable world, just, and that recognises human dignity.

REFERENCES

1. Constitution of India, Art 14
2. Constitution of India, Art 14
3. Constitution of India, Art 21
4. Maneka Gandhi v Union of India AIR 1978 SC 597
5. Right to Information Act 2005 (India)
6. Justice K S Puttaswamy v Union of India (2017) 10 SCC 1
7. Digital Personal Data Protection Act 2023 (India)
8. The Hindu, 'Jharkhand's Aadhaar-linked Public Distribution System Leaves Out Many Poor' The Hindu (12 July 2019) <https://www.thehindu.com/news/national/other->

[states/jharkhands-aadhaar-linked-public-distribution-system-leaves-out-many-poor/article28398081.ece](https://www.jharkhands-aadhaar-linked-public-distribution-system-leaves-out-many-poor/article28398081.ece) accessed 18 May 2025

9. European Commission, 'Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)' COM (2021) 206 final
10. Government of Canada, Directive on Automated Decision-Making (2019) <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/automated-decision-making.html> accessed 18 May 2025
11. New York City Council, Local Law No. 144 of 2021 (Bias Audit Law for Automated Employment Decision Tools)