Implications of the Use of AI Algorithms in the Criminal Justice System from a Constitutional Perspective

Mahant P A[a]

[a]M S Ramaiah College of Law, Bengaluru, India

_____

As a country aspires to develop technologically, the integration and the use of AI have increasingly become common. The integration of these technologies has had a great impact on various sectors of society, and the criminal justice system is no exception. The use of the algorithm has helped in crime reduction, and predicting criminal behaviour with its machine learning techniques to increase the effectiveness of the judicial system. The proponents of the use of these AI algorithms have argued that AI's implementation has proven to be more accurate in predicting recidivism than the traditional methods and reduces crime rates and the time taken in the court of law adjudicating the case. It has also led to better efficiency and resource management in the criminal justice system. However, there are concerns regarding its use such as biases, discrimination, transparency and accountability, and privacy concerns among a host of others. This article attempts to analyze these potential negative externalities associated with the use of these algorithms from a constitutional angle and its implications in the criminal justice system. The focus of my research is going to revolve around the use of AI and its legality concerning the Fundamental Rights guaranteed under our Constitution.

Keywords: *ai algorithms, biasness, transparency, privacy, fundamental rights.*

INTRODUCTION

Conceptually, AI is defined as the ability of a machine to perceive and respond to its environment independently and perform tasks that would typically require human intelligence and decision-making processes, but without direct human intervention[1]. AI's integration and use in the criminal justice system have been for some years now around the globe. Recently, in the case of Jaswinder Singh @ Jassi v State of Punjab and Anr[2], the honorable justice used the help of ChatGPT Open AI while granting bail to the accused, where cruelty was a factor for presenting a broader understanding of bail jurisprudence. Post this judgment, the use of Artificial Intelligence algorithms and their implications in the criminal justice system in India comes to the forefront yet again. In recent crimes, the country's judiciary has been at the forefront while adopting advanced technologies from live streaming of court proceedings, e-courts, and AI transcribers. Concurrently, all the possible drawbacks and repercussions shall be weighed against its uses of the system for its effective implementation into the country's judicial criminal system.

AI ALGORITHMS AND FUNDAMENTAL RIGHTS

1. Right to a Free Trial: Free Trial means a trial in which there is no bias or prejudice for or against the accused, witnesses, or parties to a case. This Hon'ble Supreme Court in the case of Zahira Habibullah Sheik and Ors has upheld the importance of fair trial. v The State of Gujarat and Ors[3], where the court held that 'each one has an inbuilt right to be dealt with fairly in a criminal trial. Denial of a fair trial is as much injustice to the accused as it is to the victim and society.' The Hon'ble Supreme Court in the case of Moti Lal Saraf v State of J&K and Ors[4] observed that the concept of fair trial flows directly from Article 21[5] of the Constitution of India. The concept of fair trial is rooted in International Law and Conventions, as stated in Article 10

---

[1] Christopher Rigano, 'Using Artificial Intelligence to Adress Criminal Justice Needs' (*National Institute of Justice*, January 2019) <https://www.ojp.gov/pdffiles1/nij/252038.pdf> accessed 21 November 2023
[2] *Jaswinder Singh Jassi v State of Punjab & Anr* Crm (M) 36727/2022
[3] *Zahira Habibullah Sheikh and Ors v State of Gujarat and Ors* (2004) 5 SCC 353
[4] *Moti Lal Saraf v State of J&K* and *Ors* (2006) 10 SCC 560
[5] Constitution of India 1950, art 21

of the Universal Declaration of Human Rights[6] and Article 14(1) of the International Covenant on Civil and Political Rights[7]. These principles ensure that everyone has the right to a free and fair trial unbiased decision-making, and presumption of innocence, among others. The AI algorithms are developed and trained on historical data like arrest records, court documents, and other public data sources which leads to numerous biases which are implemented in some cases without taking into account the other factors which violate the principle of the right to a free trial. The data that is collected and fed into the system leads to different kinds of biases like Data Bias - where AI algorithms may over-represent certain racial or ethnic groups in arrest and conviction data due to historical biases that perpetuate systematic discrimination; Contextual Bias - AI algorithms which are used for predicting future crime likelihood, may introduce bias due to their context, which may disproportionately affect certain groups; Feedback Loop Bias - AI algorithms can reinforce biases in law enforcement by directing focus on specific areas, leading to increased arrests and perpetuating bias and many more.

Secondly, 'All criminal trials are based on the principle that the accused is innocent until proven guilty'. The presumption of innocence is a cardinal principle of our legal system and a basic right of the accused person. Article 14(2) of ICCPR[8] states that 'Everyone charged with a criminal offense shall have the right to be presumed innocent until proven guilty according to law'. Over time, the pronouncements of the Supreme Court have consistently reaffirmed that the presumption of innocence is a human right. This presumption of innocence must condition his/her treatment and the procedure of the trial throughout. In its recent judgment in the case of Suresh Thipmppa Shetty v State of Maharashtra[9], the Hon'ble Apex Court has held that the presumption of innocence is a fundamental as well as a human right and that the cardinal matters of life and liberty are not to be trifled with. Therefore, the use of algorithms in pre-trial detention decisions, bail hearings, and sentencing violates the right of presumption of innocence as the algorithm does not consider the relevant circumstances of the accused. What it does is brand an individual as a convict and predict the possibility of him committing a crime in the

---

[6] Universal Declaration of Human Rights 1948, art 10
[7] International Covenant on Civil and Political Rights 1976, art 14(1)
[8] International Covenant on Civil and Political Rights 1976, art 14(2)
[9] *Suresh Thipmppa Shetty v State of Maharashtra* (2023) SCC Online SC 1038

future based on the data collected which is very skewed and historical without giving any importance to the principles of free trial. There is a flaw in the algorithm when it does not consider the presumption of innocence which leads to biased and irrational decisions. It does not take into account the nature of the offense, circumstantial evidence, or prior criminal records and makes decisions only on the data that it has collected and has been fed which can be disastrous.

2. Biasness and Discrimination: Equality is defined under Article 14[10] of the Constitution. The intent of the Constituent Assembly while drafting Article 14 was to make sure that every Human Being i.e., the citizens, as well as non-citizens, are treated equally without any discrimination before the eyes of law. It is important to note that there is already existing discrimination in our country at different levels based on the Varna system, caste, color, etc. The AI algorithms as noted above trained on the data that is available in the past, thereby historically discriminated communities stand to be disproportionately discriminated. As a result, the use of an AI algorithm will not aid in alleviating the social evil of discrimination but entrench the issue deeper into society. A similar AI algorithm is already in use in the United States of America (USA) named COMPAS. A study by a US media outlet on the COMPAS algorithm showed that though the COMPAS does not use ethnic origin or skin color as an input, the risk assessment tool designed to predict the likelihood of reoffending in Broward County in Florida in 2016 is biased against blacks.[11] The results show that COMPAS was 77% more likely to rate black defendants as 'high-risk' than white defendants, and it was almost twice as likely to mislabel white defendants as lower risk than black defendants[12]. The black defendants were also twice as likely as white defendants to be misclassified as being at a higher risk of violent recidivism. White violent recidivists were 63% more likely to have been misclassified as having a low risk

---

[10] Constitution of India 1950, art 14

[11] Julia Angwin et al., 'Machine Bias: There's software used across the country to predict future criminals. And it's biased against the blacks.' (*Pro Publica,* 23 May 2016) <https://www.benton.org/headlines/machine-bias-theres-software-used-across-country-predict-future-criminals-and-its-biased> accessed 20 November 2023

[12] Jeff Larson et al., 'How We Analyzed the COMPAS Recidivism Algorithm' (*Pro Publica,* 26 May 2016) <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm> accessed 20 November 2023

of violent recidivism, compared with black violent recidivists.[13] Any individual associated with the crime must be equally treated before the courts of the land as propounded in Article 14 of the Constitution. In contrast, using these AI algorithms when sentencing criminals is not based on the facts of the crimes committed and the defendant's criminal history, but also on factors outside his or her control 'or on the possibility' of a future crime that has not taken place[14]. Chief Justice Roberts in the case of Buck v Davis[15] openly acknowledged that 'as an initial matter, this is a disturbing departure from a basic premise of our criminal justice system: Our law punishes people for what they do, not who they are. Dispensing punishment based on an immutable characteristic flatly contravenes this guiding principle'.

Apart from biasness and discrimination, AI algorithms also violate the principles of Natural Justice. Audi Alteram Partem (hear the other side) forms a facet of natural justice and is a requirement of Article 14 for restoring the principles of natural justice. These AI algorithms are not created or are incapable of hearing any side; it is capable of making decisions based on the data that is trained on limping the criminal justice system. It is a developing norm that judges may highly depend on AI-generated numbers and data that when confronted with a high recidivism prediction, blurs the rational judgment of the judges. Judges, who ought to act as gatekeepers, and forbearers of equality and neutrality, can themselves fall victim to a favorable bias towards the AI's output. Also, they violate the principle of Reasoned Decisions which forms a part of Natural Justice. It is the psychological human tendency to be biased[16]. It is a matter of fact that the inputs or the training methods are fed by developers or any other organization/agency that has been entrusted with its development. This means that the algorithm can be tampered with, by the existence of bias by its developers. Further, it is to be noted that while giving its decision, the AI's capabilities are limited to a binary answer, but it

---

[13] *Ibid*

[14] Danielle Kehl et al., 'Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing' (2017) Berkman Klein Center for Internet & Society, Harvard Law School <https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf> accessed 20 November 2023

[15] *Buck v Davis* [2017] 580 US 1, 21

[16] Paul A. Warren et al., 'A Re-Examination of "Bias" in Human Randomness Perception' (2018) 44(5) Journal of Experimental Psychology: Human Perception and Performance <psycnet.apa.org/fulltext/2017-47517-001.pdf> accessed 20 November 2023

does not give an elaborate reasoning for its decision. Explaining a decision is now universally recognized as one of the basics of effective governance and protection against arbitrariness. This limitation of the AI algorithm violates the rule of reasoned decision-making that is necessary and pivotal in the criminal justice system.

3. Transparency and Accountability: 'Transparency' in general terms is 'the quality or condition of being transparent; state of being transparent; that which is transparent. Transparent means easily seen through, recognized, understood, or detected, manifest, evident, obvious, clear.'[17] Supreme Court emphasized the role of transparency and openness in governance and held that compliance with these principles satisfies the requirements of Articles 14[18], 19[19] and 21[20] of the Constitution[21]. In Reliance Petrochemicals Ltd. v Indian Express Newspapers Bombay (P) Ltd.,[22] the Hon'ble Supreme Court held the right to know as a fundamental right under Article 21 of the Constitution of India. The Hon'ble courts have always recognized the Right to Know and conformity with principles of transparency and openness in the decision-making process as flowing out of the golden triangle of the Constitution. The Right to Information Act 2005[23] was enacted to manifestly uphold the rights that were latently a part of Part III of our Constitution. Given the inherent opaqueness in the algorithm makes it difficult to understand its final output. The algorithm classifies individuals as high or low risk based on a risk assessment score, but the process of predicting outcomes is not transparent[24]. All aspects of the algorithm should be made available for public scrutiny including source code, variable weights, and training data, to identify bias and correct course errors, because transparency is crucial for protecting civil liberties, ensuring public trust, and addressing flaws. So what can be reasonably implied is that the accused is labeled as a high-risk/low-risk individual based on the various parameters which

---

[17] *Oxford English Dictionary* (2nd edn, Oxford University Press 1989) 419

[18] Constitution of India 1950, art 14

[19] Constitution of India 1950, art 19

[20] Constitution of India 1950, art 21

[21] *Global Energy Ltd. v Central Electricity Regulatory Commission* (2009) 15 SCC 570

[22] *Reliance Petrochemicals Ltd. v Indian Express Newspapers Bombay (P) Ltd* (1988) 4 SCC 592

[23] Right to Information Act 2005

[24] Madalina Busuioc, 'Accountable Artificial Intelligence: Holding Algorithms to Account' (2021) 81(5) Public Administration Review <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8518786/pdf/PUAR-81-825.pdf> accessed 20 November 2023

is taken into consideration by the AI while giving its output, where the inputs, training data, and methods are not only biased but also opaque. This means that the accused is not provided the data for his convictions because he/she is unaware or does not know the reasons for the same. Since these inputs are made unavailable to the public, it can be difficult for independent experts to audit its decision-making process and hold these systems accountable.

Since these algorithms are opaque in nature, fixing accountability to them is a strenuous task. Transparency is crucial for understanding AI algorithms and their accountability. A mechanism of checks and balances is necessary to address potential harms and their utility to the public interest. Auditing the algorithm's source code, inputs, and outputs is essential for verifying the decision-making process. An independent review board, including legal, technical, and statistical experts, incarcerated individuals, public defenders, prosecutors, judges, and civil rights organizations, should oversee the audit to ensure transparency and accountability.

Another concern is about the lack of due process in the use of AI algorithms in criminal justice decision-making. In the traditional system, an accused is entitled to a range of fundamental and procedural rights as expressed in the Constitution ranging from an oral hearing, to calling witnesses, challenging the evidence, and cross-examining but the same are not available with the introduction of AI algorithms into the system. AI algorithms are not accountable in the same way as judges, as the role of the judge is more complex, requiring a weighing of law and facts, tempered with reasoned discretion, to balance competing interests. On the other hand, AI algorithms do not follow due process, in contrast classifying individuals based on data from a database, may lead to potentially unjust results.

4. Privacy Concerns: Perhaps, one of the biggest drawbacks of these algorithms is privacy. The Supreme Court's Puttaswamy[25] decision established privacy as a fundamental right under Article 21[26] of the Indian Constitution, including informational privacy (including biometrics and other personal data). The court ruled that retaining private data constitutes an interference

---

[25] *KS Puttaswamy v Union of India* (2017) 10 SCC 1

[26] Constitution of India 1950, art 21

with privacy[27]. The Supreme Court in Maneka Gandhi v Union of India[28] ruled that a law denying personal liberty must pass the test of Articles 14, 19, and 21. Sometimes an individual's data is collected without his knowledge or forced compulsion without his consent and indefinitely retained in the database which infringes his fundamental right. The AI algorithms may violate Article 21's right to privacy by collecting, storing, transferring, and sharing data without individual consent, and lack procedural safeguards for data retention and deletion.

When government/government agencies collect individual data under the guise of bringing effectiveness to the justice system, they would not have made any provisions for the storage and retention of personal data. Even after an arrestee is acquitted or released without trial, the Act still allows indefinite retention of data, violating the right to privacy. There is no logical nexus between the collection of such extensive data and storing them in the central database and future or past offending likelihood and the class of individuals included because of the reason that no study suggests that those arrested or detained are more likely to have committed future crimes. Extensive databases do not improve crime prevention, investigation, or prosecution. The Aycaguer v France[29] case emphasized the importance of proportionate data storage periods and the legality of indefinite retention or perpetuity retention in crime investigation and detection. The algorithms do not even mention the dissemination of the personal data. When the data of any person is no longer required or who expects that his/her personal data will be no longer stored or processed then he/she should be able to remove it from the system where the information is no longer necessary, relevant, or is incorrect or is illegitimate. The system does not provide details on deletion procedures, retention periods, and data erasure, which aligns with the principle of right to be forgotten which seriously puts the person's data at risk and in violation of Article 21. Therefore, the indefinite retention of such data and measurements is not necessary for the legitimate aim of aiding future investigations and crime predictions.

Privacy is a fundamental human right protected by international legal frameworks like the Universal Declaration of Human Rights and ICCPR. Courts have upheld this right in cases like

---

[27] *KS Puttaswamy v Union of India* (2017) 10 SCC 1
[28] *Maneka Gandhi v Union of India* (1978) 1 SCC 248
[29] *Aycaguer v France* [2017] ECHR 587

Auto Shankar[30], PUCL[31], and R.M. Mlakani v State of Maharashtra[32] et.al, ensuring that individuals' privacy is protected. Moreover, AI systems are vulnerable to hacking and cyberattacks, exposing sensitive information. This raises concerns about data privacy, particularly if sensitive information is being collected and stored in a way that could be accessed or used by unauthorized individuals without consent, therefore, raising concerns about data privacy and violating Article 21.

ANALYSIS AND SUGGESTIONS

With the advent of the 21st century, there has been a nudge toward the use of AI algorithms and has inevitably been a part of the usual discourse. Around the globe, the use of AI algorithms is used to aid the criminal justice process and not to substitute the entire process. It is a well-known fact that human rationale and judgment always take precedence over AI algorithms. In the Chat GPT case[33], the presiding judge sought an explanation from the AI chat box on the granting of bail. It is worthwhile to note that only suggestions were recorded and the judgment was given by the Hon'ble Judge. The norm should be set, where AI's recommendations can be recorded, but the judge must have the final say using his own reasoning and logic. AI algorithms should not supplant human decision-making and judgment but should supplement them instead. On the human rights front, the application and use of these algorithms should be in consonance with the fundamental rights of the citizens. As stated above, these AI algorithms violate several fundamental rights including, but not restricted to, the right to a fair trial as it lacks the presumption of innocence, violates citizens' rights of equality by labeling accused as high or low risk, making biased decisions, violating natural justice principles, and lacking reasoned decision-making, violates rights by reducing transparency, accountability, and due process of law as the opaqueness of AI algorithms makes it difficult to understand their final output, and interference with privacy and infringement of private data without any procedural safeguards in places.

---

[30] *R. Rajagopal v State of T.N.* AIR 1995 SC 264
[31] *People's Union for Civil Liberties v Union of India* AIR 1997 SC 568
[32] *R. M. Malkani v State of Maharashtra* AIR 1973 SC 157
[33] *Ibid*

Some suggestions could be as follows:

- We must ensure that free and fair trial is made accessible in accordance with the Indian constitution and other international agreements as it leads to reductions of biases in the justice system.

- We must ensure that the presumption of innocence which is a cardinal principle of our legal system and a basic right of the accused person is adhered where, the presumption must stand and be the guiding principle right from the moment of suspicion, through investigation, throughout the trial process and till the delivery of the verdict and not assume the guilt of the person based on the training data.

- We must aim to reduce the inherent bias and arbitrariness that leads to the discrimination of certain marginal communities in society. The algorithms should be trained in such a manner that the decisions made are not based on factors such as demographics, social media, online activity, etc, to classify individuals accused of a crime as either high-risk or low-risk individuals.

- We must ensure that the observations or decisions made by the AI should give equal importance to the principles of Natural Justice which act as the antithesis of arbitrariness. As in traditional settings, the other party should be given a fair chance and should be heard before arriving at a decision, and made sure that the outputs given by the algorithm should be reasoned.

- We should ensure that the concerned stakeholders and experts make the source code and all the training data available to the public and for auditing purposes to identify bias and correct course errors to bring transparency and accountability in the process which is crucial for protecting civil liberties and ensuring public trust.

- We must ensure that the right to privacy is not breached or interfered with and that provisions are made available concerning the retention and deletion of data, data erasure and the right to be forgotten. An individual's consent should be sought while collecting and storing his/her personal data which may include demographics, biometrics, financial records, and other such personal information.

It is understood that the usage of AI technologies is a double-edged sword. There have been many significant changes and developments in society with the help of Artificial Intelligence. At the same time, it should bear in mind that these technologies should aid the criminal justice system through its working models for bringing efficiency and effectiveness into the justice delivery system and not violate any of the human rights granted, as it is paramount for the well-being of society.

CONCLUSION

The usage of Artificial Intelligence algorithms and systems in the current legal context raises important questions about the balance between technological development and the rights and interests of the state's citizens. On one hand, they ensure and enhance efficiency and accuracy, on the other hand, they raise concerns regarding bias, discrimination, transparency, and privacy. These technologies have proved to be efficient and effective in the criminal justice system in terms of crime prevention and detection and aid both the law enforcement bodies and the justice systems. They have proven to be more accurate in predicting recidivism than traditional methods used by judges. While there has been great progress in AI development, human expertise, context, and ethical considerations should always guide the use of AI[34].

The final decision-making power and human oversight should rest with the judges and the courts to alleviate the inherent bias and overarching outcomes of the AI's algorithms. These algorithms should concertedly address the concerns relating to systematic biases, discriminatory practices, transparency and accountability, privacy concerns, and many more. Multidisciplinary collaborations involving legal professionals, technologists, ethicists, and affected communities are necessary to develop robust frameworks that promote transparency, accountability, and fairness in the use of AI[35]. The inputs should be made available when the need arises for the purpose of auditing by independent experts. The concerned authorities and organizations involved in the development of these AI tools should take steps to minimize the

---

[34] Caleb Jren, 'Implications of AI in the Criminal Justice System' (*INSPIRIT AI*)
<https://www.inspiritai.com/blogs/ai-student-blog/implications-of-ai-in-the-criminal-justice-system> accessed 21 November 2023
[35] *Ibid*

potential for bias and error in the AI algorithm and provide regular updates and improvements. By striking a balance between the benefits and potential risks, we can leverage AI's power to create a more just and efficient criminal justice system[36]. The main objective of this article was to provide AI's potential drawbacks regarding its use in the legal realm. Law and technology are mutually inclusive subjects that go hand in hand. Although Artificial Intelligence systems offer wide-ranging benefits in law enforcement and the justice system, they shall ensure that the fundamental rights and interests of individuals are not violated or infringed upon.

---

[36] Hera Aiman, 'Artificial Intelligence In Criminal Justice System' (*Legal Service India*) <https://www.legalserviceindia.com/ artificial intelligence in criminal justice system.html> accessed 21 November 2023