Otitonaiye W.♦

Impact of artificial intelligence algorithm
for passing judgement in judicial system

Abstract: This paper explores the use of AI in the US criminal justice system and how it influences fairness and bias in delivering judgment. Artificial Intelligence (AI) is an innovative technology, which has significantly changed the way humans run their day-today activities, it has been implemented for various tasks such as job employment screening, fraud detection, mortgage application, health care diagnosis, etc., most industries are fast relying on AI for decision making. The judicial system in the United States (US) is beginning to adopt the use of AI tools for passing judgment, a critical and complex task that can determine whether an accused person is incarcerated or free to move around in public. It is a duty that must be carefully executed, as a wrong decision can significantly affect the life of a person. People are divided over the use of AI/ML (Machine Learning) tools in court with fears of an unfair judgment emanating from the use of historical data that might be influenced by human bias. This concern can be eliminated if conscious decisions are made by developers to understand the dataset, use an appropriate algorithm design approach, and periodically review codes. AI tools have the potential to eradicate that fear of human bias with time.

Keywords: artificial intelligence; criminal justice; machine learning; smart court; fairness.

I. Introduction

The judiciary system is progressively adopting the use of Artificial Intelligence (AI) for deciding and passing court judgment, as human judges are gradually relying on machine learning models when making bail and parole decisions that affect the freedom of thousands of people yearly [13]. Research has also shown that various industries are increasingly trusting AI for decision making, such that big companies now employ the use of AI software for the first stage of screening job applications and shortlisting successful candidates before a human recruiter is involved. Mortgage applications, university admission, and insurance claims are using insights gained from predictive analytics tools for decision making [17].

A popular symbol of the judiciary system is the figure of the Lady Justice; a blindfolded woman (objectivity and impartiality), holding a set of scales (weighing of evidence) and a sword (punishment). She is seen trampling on a snake (evil and lies) with one of her legs and has the support of a book on the other leg (law and constitution). This is a symbol of fairness and equality within law administration without recourse to corruption, greed, favor, and prejudice.

Criminal sentencing is one of the most complex responsibilities of judgment, it is a task that involves judges facing multiple and conflicting instructions from the legislature and society. The sentence must be in measurable proportion in retribution to the crime committed. It must be of a suitable length and type, to rehabilitate the defendant before returning to society after punishment. It

---

♦ **Otitonaiye Wilson Sunday** – Master of Science of the Department of Computing and Informatics Bournemouth University, Bournemouth, (United Kingdom). E-mail: s5229500@bournemouth.ac.uk

must discourage the defendant from offending again, as well as others from offending in the first place. The sentence must be long enough to protect society from danger [5]. Despite the complexity of deciding a case judgment, courts are selecting the use of AI algorithms to aid proceedings and ease their processes, due to its capability to perform complex analysis on large datasets and make predictions [5]. To ensure that the judiciary act as an agent of the legislature's will when using these tools, historical datasets of past judgments and stipulated punishment that commensurate with the crime committed as specified by the legislature are provided as input to these tools [5]. A wrong decision by an AI tool can be very costly and affect the life of a defendant, hence the focus of this study is to carry out a literature review on the use of AI tools in passing fair judgment without recourse to bias in the United States (US).

II. Assumptions and background

A critical feature of AI is that it must be intelligent in nature, Grewal in his paper [7] explored different definitions of intelligence, addressed their limitations, and provided a more concise definition. For this study, we will adopt the definition of intelligence as a general mental ability for reasoning, problem-solving, and learning that integrates cognitive functions such as perception, attention, memory, language, or planning.

In this paper, AI, and Machine Learning (ML) will be used interchangeably as ML is a subset of AI [9]. AI is the branch of computer science that deals with the simulation of intelligent behavior in computers as regards their capacity to mimic, and ideally improve, human behavior. To achieve this, the simulation of human cognition and functions, including learning and problem solving, is required [15]. An AI system can be identified as a mechanical simulation system for collecting knowledge and information, collating, and interpreting the information obtained, and disseminating it to the eligible in the form of actionable intelligence [7]. ML was defined by Mitchell in his book [11] as "Machine learning addresses the question of how to build computer programs that improve their performance at some task through experience". ML algorithms learn from experience and can produce output when provided with new inputs, thus, an AI or ML system can be described as an intelligent system with the capability to receive information, learn from it, and become better at a task such that it can make future predictions [18].

ML systems can either learn through a supervised, unsupervised, or reinforcement learning approach. In both the supervised and reinforcement learning approach, an explicit target output is stated with each given input but in the unsupervised learning approach, the target output is not explicitly stated [4]. In both approaches, it learns through identifying patterns by performing complex mathematical equations on the training dataset and then make predictions [11]. An AI system can learn from experience and identify patterns just as human judges, it has the potential to perform the task of passing judgment when provided with all required input datasets and expected target output.

III. Literature review

There has been an increase in the adoption of AI tools in US courts and correctional agencies to support judgment, determine the nature and length of the punishment meted out to a defendant (e.g., sentencing and parole), and even making the decision of releasing an offender before trial (e.g., custody and bail decisions) [17]. It is believed that the use of AI and predictive policing analytics integrated with a computer-aided response and live public safety video enterprises will aid law enforcement and criminal justice professionals to better maintain public safety. It will hasten the response time to incidents, prevent threats, stage interventions, divert resources, and investigate and analyze criminal activity. AI has the potential to be a permanent part of the criminal justice eco-

system, providing investigative assistance, and allowing safety [14]. Other countries like the United Kingdom (UK) are also considering the use of AI system as detailed in a 2016 bill by The Ministry of Justice titled "Transforming our justice system", states its vision to digitize court proceedings as it will play a significant role in ensuring that the legal systems in England and Wales provides a rapid and certain judgment, in a manner that saves people time and money, and reduces the impact of legal proceedings [19].

US courthouses are widely adopting the use of an AI tool called Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) algorithm, developed by a private company Northpointe. It is a case management tool and decision supporting tool for assessing the likelihood of a defendant becoming a recidivist. It attributes a 2-year recidivism-risk score to arrested people. It also evaluates the risk of violent recidivism as a score[12]. Some of the most talked-about cases that COMPAS was used are Loomis v. Wisconsin and Kansas v. Walls, but the public has been divided over the outcome of judgment [20]. ProPublica analyzed the outcome of the COMPAS algorithm, particularly looking at racial bias, and the result showed that a black defendant is twice as likely to be misclassified as a white defendant to have a high risk of violent recidivism, and white recidivists had a 63.2% misclassification rate as low risk than the black defendants [6].

There is also fear of AI algorithmic 'black-box' problem despite their promising analytical nature. The misguided interpretations and inferences resulting from data analytics have engendered enormous debates amongst policymakers, practitioners, and academics in the past [10]. Another concern is that there could be a trade-off between swift judgment and fairness, the use of AI in the judiciary system might minimize the time it will take to pass a judgment, reduce the influence of extraneous factors such as weariness and emotional instability, but it poses a risk of making decisions that reveal different human-made, structural biases that originate from the legal system, potential bias, and discrimination embedded in their data sources or the AI's programming itself [3].

If an important and sensitive responsibility is to be handed over to a machine, it is reasonable to say establishing an easy and fast process is not enough, there must be some level of assurance that it can be fair and unbiased in nature, just like the Lady Justice.

To provide more perspective, it is important to understand what the words; fair and bias means, in the Science, Technology, Engineering and Mathematics (STEM) environment, bias is often used to describe the probability associated with a binary event such as a coin toss [17], but in this paper, both words are referred, as defined in the Oxford Dictionary, fair means "acceptable and appropriate in a particular situation" [12] and bias means "against one group of people, or one side in an argument, often not based on the fair judgment"[12].

Those who worry about the implementation and use of AI tools in criminal justice fear that true fairness cannot be attained as the data and algorithms risk reproducing biases against historically disadvantaged populations. Most ML models are trained using historical data, thus are likely to notice patterns of human bias performed over the years, which can influence prediction [8]. It is critical to address the issue of fairness if this model is to be used and accepted universally. Ao in his paper [1] suggested Three (3) approaches to programming fairness into an algorithm. The first is the willful blindness approach which treats subgroups the same regardless of their distinct difference like race and gender, but it stands the risk of creating an unaware algorithm as it falls short of key human attributes. The second approach is to ensure statistical parity in the outcome, by creating 2 distinct groups (Protected and Non-Protected) and selecting an equal number from both groups, but this method will require constant verification and modification of the thresholds and groups by

someone. It also falls short of the capability to account for subgroups within the group. The final and recommended approach is Predictive equality, which does not force equality in the outcomes but in the performances and accuracy of the algorithm across distinct groups, it should be able to identify specific relevant subgroups, identify a set of metrics for defining fairness and hierarchy within the set, identify attributes of a subgroup that should either be adapted or removed. This constraint of this approach is that it is dependent on a large dataset for training the model, else it could come at a cost of low accuracy.

A computer system or machine does not have actual or subconscious biases, they provide answers which are driven by the code that is given to them, if the code has no inappropriate variables, then appropriate answers will follow, a concept known as Garbage In Garbage Out (GIGO). Thus, flaws of this nature are, at least in theory, readily fixable. The design part of the algorithmic process is key and very essential [2]. If a system is designed with insight and a deep understanding of the historical datasets been used, there is a high probability of eradicating the fear of bias.

The consensus is that AI is changing so many aspects of our lives already. Industries such as transportation, health care, education, and entertainment have seen changes due to its emergence [3]. Rarely did anyone argue that AI will intentionally make bias predictions, even when the arguments are based on unfair decisions, evidence showed that historical data been is the major cause for concern.

IV. Gaps and conclusion

AI has the potential of transforming the justice system, it will not only affect how activities such as (evidence gathering and presentation, jury observation and recommendation, etc.) are done in courts, it has the potential of disrupting the academic training process for human judges. This paper has not looked at the possible impact it will have on future judges and the possibility of skill deficiency due to the reliance on AI.

One can still safely assumed that AI/ML models are not biased in nature and can make a fair prediction if humans make the conscious decision to diligently investigate and understand the datasets used in training algorithms and periodically reviewing them for improvement.

An undeniable advantage of AI decision making over human process is efficiency, AI can handle complex use cases through mathematical computation on time and synthesize thousands of variables instantaneously [2], if incorporated in the justice system appropriately, will not only minimize the time it will take for a case to run in court, but also save cost for everyone involved, it will increase productivity, and in time take away that human bias that we fear so much.

References

1. AO Allsop James. Technology and the future of the courts. 2019. Available at: http://www.fedcourt.gov.au/digital-

2. Bagaric M., Hunter D., and Stobbs N. Erasing the Bias Against Using Artificial Intelligence to Predict Future Criminality: Algorithms are Color Blind and Never Tire University of Cincinnati Law Review, 2020, vol. 4, no. 88, pp. 1037-1081.

3. Buocz T. J. Artificial Intelligence in Court. Legitimacy Problems of AI Assistance in the Judiciary," Copenhagen, 2018. Available at: www.nytimes.com/2011/03/05/science/ 05legal.html

4. Dayan P., Unsupervised Learning - (The MIT Encyclopedia of the Cognitive Sciences.). Paperback Edition. Massachusetts: A Bradford Book, 2001.

5.   Donohue M.E. A replacement for justitia's scales?: machine learning's role in sentencing," 2019. Available at: https://perma.cc/YAB4-79MT].

6.   Flores A.W. and Bechtel K. 38 EEDERAL PROBATION. Vol. 80. Number 2.

7.   Grewal D. A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering," Ver. I, 2014. Available at: www.iosrjournals.org.

8.   Hoffmann A.L. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. Information Communication and Society, 2019, vol. 22, no. 7, pp. 900–915, doi: 10.1080/1369118X.2019.1573912.

9.   Jakhar D., Kaur I. Artificial intelligence, machine learning and deep learning: definitions and differences. Clinical and Experimental Dermatology, vol. 45, no. 1. Blackwell Publishing Ltd, 2020, pp. 131-132, doi: 10.1111/ced.14029.

10.   Liu H.-W., Lin C.-F., and Chen Y.-J. Beyond State v Loomis: artificial intelligence, government algorithmization and accountability. International Journal of Law and Information Technology, 2019. vol. 27, no. 2, pp. 122-141.

11.   Mitchell T. Machine Learning, ISBN: 0070428077. New York: McGraw Hill, 2017.

12.   Oxford Dictionary of English. Oxford University Press, 2010.

13.   Re R.M. et al. UCLA School of Law; 2017-2019 PULSE Fellow, UCLA School of Law, 2019. Available at: https://perma.cc/9DQ6-MH7E.

14.   Rigano C. Using Artificial Intelligence to Address Criminal Justice Needs (NIJ Journal 280), 2018.

15.   Russell S.J. Artificial intelligence: A Modern Approach. Upper Saddle River, New Jersey: Prentice Hall , 2010.

16.   S. lo Piano Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward," Humanities and Social Sciences Communications, 2020, vol. 7, no. 1, doi: 10.1057/s41599-020-0501-9.

17.   Scantamburlo T., Charlesworth A., and Cristianini N. Machine decisions and human consequences, 2019. Available at: https://belong.co/.

18.   Shaikh R., Sahu T., and Anand V. Predicting Outcomes of Legal Cases based on Legal Factors using Classifiers. Procedia Computer Science. 2020,vol. 167, pp. 2393-2402.

19.   Transform Justice, Transform Justice briefing on the Prisons and Courts Bill. United Kingdom Parliament, 2016.

20.   Završnik A. Criminal justice, artificial intelligence systems, and human rights. ERA Forum, 2020. Vol. 20, pp. 567-583.

Original   source   -   Otitonaiye   Wilson   Sunday   –   Research   gate   - https://www.researchgate.net/publication/347444168_IMPACT_OF_ARTIFICIAL_INTELLIGEN CE_ALGORITHM_FOR_PASSING_JUDGEMENT_IN_JUDICIAL_SYSTEM

Отитонайе У.♦

Влияние алгоритма искусственного интеллекта на вынесение приговора в судебной системе

Аннотация: Исследуется использование искусственного интеллекта (ИИ) в системе уголовного правосудия Соединенных Штатов Америки (США) и то, как он влияет на справедливость и предвзятость при вынесении приговора. Искусственный интеллект - это инновационная технология, которая значительно изменила то, как люди ведут свою повседневную деятельность, она реализует решения для различных задач, таких как проверка занятости, обнаружение мошенничества, подача заявки на ипотеку, диагностика здравоохранения и т. д., большинство отраслей быстро полагаются на ИИ для принятия решений. Судебная система в США начинает использовать инструменты ИИ для вынесения приговора, что является важной и сложной задачей, которая может определить, находится ли обвиняемый под арестом или может свободно передвигаться в общественных местах. Это обязанность, которую необходимо тщательно выполнять, так как неправильное решение может существенно повлиять на жизнь человека. Люди расходятся во мнениях относительно использования инструментов AI/ML (машинного обучения) в суде из-за опасений несправедливого решения, вытекающего из использования исторических данных, на которые может повлиять человеческая предвзятость. Эта проблема может быть устранена, если разработчики будут принимать сознательные решения для понимания набора данных, использования соответствующего подхода к разработке алгоритмов и периодического просмотра кода. Инструменты искусственного интеллекта могут со временем искоренить этот страх перед человеческими предубеждениями.

Ключевые слова: искусственный интеллект; уголовное правосудие; машинное обучение; умный суд; справедливость.

♦ **Отитонайе Уилсон Сандей** – магистр, факультет вычислительной техники и информатики Борнмутского университета, Борнмут, (Великобритания). E-mail: s5229500@bournemouth.ac.uk