THE CONSEQUENCES OF ALGORITHMIC DECISION-MAKING

Rohan kumar Jawarkar, University of Mumbai

ABSTRACT

Algorithmic decision-making has gained widespread acceptance as an innovative way to addressing the claimed cognitive and perceptual constraints of human decision-makers by offering "objective" data-driven recommendations. Despite this, numerous incidents of algorithmic prejudice continue to emerge when firms deploy Algorithmic Decision-Making Systems (ADMS). In domains such as health, hiring, criminology, and education, harmful biases have been discovered in algorithmic decision-making systems, generating growing social concern about the influence these systems are having on people's well-being and livelihood. As a result, algorithmic fairness strategies try to figure out how ADMS treat different people and groups, with the goal of detecting and correcting detrimental biases.

Keywords: ADMS, Harmful, Decision-Making.

INTRODUCTION

Data on sensitive traits or protected categories must be available for demographic-based algorithmic fairness strategies to work. Previous study has found that information on demographic segments including race and sexuality is frequently unavailable due to a variety of organisational problems, legal constraints, and practical concerns. Some privacy rules, such as the GDPR in the EU, not only consider data subjects to express meaningful consent before their data is acquired, but also restrict the gathering of sensitive information like race, gender, and sexuality (Diakopoulos, 2016).

Some business privacy rules and standards, such as Privacy protection, require firms to be deliberate in their data collection activities, gathering only the information they need and can designate a use for. Given the ambiguity around whether it is appropriate to ask consumers for sensitive statistical profile, most legal and regulatory teams advise their companies to evoke a sense of caution and gather such data only if legally needed. As a result, privacy concerns frequently take precedence over maintaining product fairness, as the trade-offs between bias mitigation and individualised privacy are uncertain (Goodman & Flaxman, 2017).

When sensitive demographic data is acquired, companies face a number of practical issues during the purchase process. Self-reporting systems are used by numerous organisations to acquire sensitive demographic data. Self-reported data, on the other hand, is frequently fragmentary, unreliable, and unrepresentative, owing in addition to the absence of incentives for people to submit correct and complete information. In certain circumstances, practitioners opt to infer restricted categories of people based on proxy data, which is a highly incorrect strategy. Corporations also have a hard time capturing unobserved qualities like handicap, sexuality, and religion because these classifications are frequently lacking and unmeasurable (Kasy & Abebe, 2021).

Overall, determining how to describe and categorise demographic data is a never-ending task, as demographic categories move and alter over time and between settings. Once demographic data is acquired, antidiscrimination laws and policies prohibit organisations from using it since knowing about sensitive categories exposes them to legal risk if discrimination is discovered without a plan in place to properly mitigate it. Faced with these obstacles, businesses interested in using demographic-based algorithmic fairness strategies have requested instruction on how to obtain and use demographic data properly (Kissell & Malamut, 2005).

Prescripting statistical notions of justice on algorithmic systems without taking into account the social, financial, and political structures, in which they are entrenched, on the other hand, may fail to help marginalised groups and hinder fairness efforts. As a result, establishing guidelines necessitates a greater awareness of the costs and trade-offs associated with using and not using demographic data. Attempts to detect and reduce damages must take into account the larger frameworks and power structures in which algorithmic systems as well as the data they use are embedded. Finally, while my effort is driven by the acknowledged unfairness of ADMS, it's important to remember that discriminatory practices aren't the only potential consequences of the system.

Focusing on constructivist teaching datasets and algorithms is often misguided, as recent papers and reports have argued, because suggested debiasing procedures are only relevant for a subset of the types of bias ADMS introduces or reinforces, and are likely to divert attention away from other, potentially more significant harms. In the first particular instance, harms from tools like recommender systems, content moderating effect systems, and desktop vision systems may be characterised as the consequence of multiple forms of bias, but trying to resolve bias in those systems typically entails adding more contextual factors to understand better differences among groups, rather than simply treating groups more equally (Marjanovic et al., 2018).

While many ADMS are clearly biassed in the second situation, the main cause of harm may be the system's deployment in the first place. One such example is pre-trial detention risk scores. Using statistical relationships to determine whether something should be held in custody, or, in other phrases, potentially severe punishment individuals for factors beyond their control and past decisions unconnected to the charges against them, is a significant departure from legal standards and norms in and of itself, but the majority of the debate has centred on how biassed the predictions are. Trying to collect demographic information in these situations will almost certainly cause more harm than help, as demographic information will divert attention away from the harms that are already there.

CONCLUSION

Algorithms have been criticized as a method for obscuring racial prejudices in decision-making. Because of how certain races and ethnic groups were treated in the past, data can often contain hidden biases. For example, black people are likely to receive longer sentences than white people who committed the same crime. Algorithms are used for calculation, data processing, and automated reasoning. Whether you are aware of it or not, algorithms are becoming a ubiquitous part of our lives.

REFERENCES

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, *59*(2), 56-62.

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a right to explanation. *AI Magazine*, *38*(3), 50-57.

Kasy, M., & Abebe, R. (2021). Fairness, equality, and power in algorithmic decision-making. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 576-586).

Kissell, R., & Malamut, R. (2005). Algorithmic decision-making framework. *The Journal of Trading*, *1*(1), 12-21.

Marjanovic, O., Cecez-Kecmanovic, D., & Vidgen, R. (2018). Algorithmic pollution: Understanding and responding to negative consequences of algorithmic decision-making. In *Working Conference on Information Systems and Organizations* (pp. 31-47). Springer, Cham.