Reflections on Coded Bias and AI Ethics and Accountability

Based on "Discuss the risks of deploying and using opaque and unaccountable algorithms in public domains based on Mittelstadt et. al. (2016) and Boyd & Crawford's (2012) readings and issues discussed in the movie Coded Bias."

Siddhant Bali, Roll No. 2022496

The Coded Bias movie is where the Algorithmic Justice League got its start.In addition to showcasing the testimonies of individuals affected by detrimental technology the movie depicts trailblazing women raising awareness of the dangers artificial intelligence poses to civil rights. In January 2020,the documentary made its debut at the Sundance Film Festival(Kantayya, 2020).

Real-World Impact of Algorithmic Bias

In the U.S. House Committee press conference , Joy Buolamwini stated explicitly that the biased algorithm favored "White Males"(Buolamwini & Gebru, 2018).

At the case where a qualified teacher ,loads of certificate holder Teacher in preschool is declared bad teacher fo kids by ai algorithm.it effects the teacher adversely.In China's Civil Scoring System,just based upon the AI Classifications, Many People are restricted to basic amenities such as Metro Transit Vendor Machines etc(Mittelstadt et al., 2016).

The Common major people can be declared and forced on AI controlled by Powerful Authorities but these common people don't have their own AI to counter question it. (Mittelstadt et al., 2016).The main problem is that even if any action is owned and labeled by the government doesn't mean that it's by default trustworthy and ethical.(Boyd & Crawford, 2012).The data collection by Governments and corporations leads to 24/7 Virtual Imprisonment by surveillance

of People ,conflicting their right to privacy,based on Atlantic Plaza Towers interview.(Creemers, 2018).

Surveillance Capitalism and Public Spaces

Silkie Carlo, director of the UK civil-liberties NGO Big Brother Watch, appears in the film monitoring trial deployments of facial-recognition technology by British police forces. (Big Brother Watch, 2020).

The algorithms are so problematic that they literally declare any normal citizen to terrorists . and then things took a horrible turn leaving common citizens who were wrongly accused as harassed.And when in many cases asked about how they were blamed on what grounds,authorities clearly say that AI Predicted.(Garvie et al., 2016).

At Atlantic Plaza Towers in Brownsville,Brooklyn,the building management announces plans to replace traditional keys with a facial-recognition system for tenant entry,The film follows residents forming the local tenants association to oppose the biometric locks.100+ residents including long-time tenant Icemae Downes file a formal complaint with New York State, calling the system a "tagging" of people like animals and an invasion of privacy. (Eubanks, 2018).Experts in the documentary explain how housing surveillance disproportionately harms Black and low-income communities, exacerbating gentrification pressures (Heilweil, 2019).

Roots of Bias, How Algorithms Learn

The core essence of these biases of these algorithms is the stages where it is built. It depends on the Developers ,their biases,thoughts ,their mindset and classification and training the model.and then model learns from them(Mittelstadt et al., 2016).

In 2016, Microsoft launched its AI Model Avatar named Tay,who thinks like a 19 Year old American girl.The aim is to study human behaviors ,thought and mindsets and train like a human. Tay was feeded by internet trolls and all kinds of content and opinion on extreme thought

too.. Soon Tay modeled itself as tweeting racist, sexist, and extremist content (e.g., supporting Hitler) and many more extremes.lead to offline shutdown after 16 Hours of its launch(Vincent, 2016).

AI and its development is first exposed to common major people and from the trained data ,big rich powerful companies use it in unethical ways, Even the department who codes it ,doesn't know how it works.and make AI a black box .which can lead to uncertain future consequences which may or mayn't be controlled(Burrell, 2016).

Philosophical Approaches to Ethical AI

In AI-driven decision-making, utilitarianism focuses on maximizing overall benefits, guiding AI to choose actions that produce the greatest good for the majority. Conversely, deontology emphasizes adherence to moral duties and principles, directing AI to follow predefined ethical rules regardless of outcomes(Binns, 2018). Balancing these approaches is crucial for developing AI systems that make ethically sound decision

Toward Ethical Algorithmic Governance

For algorithmic governance to actually be accountable, it should be required of the policy makers that they require human rights and bias audits before deployment, demand explanations based transparent AIs that can take complex logic and make it meaningful to the human decision, put independent multidisciplinary auditors in charge with full access to proprietary code and data place affected communities in charge of data selection and threshold setting,and create a clear, enforceable path for each person to contest machine decisions,together these steps would convert ethics as a brake on innovation into guardrails to keep the progress of technology aligned with human dignity(Mittelstadt et al., 2016).

References

1. Big Brother Watch. (2020). *Big Brother Watch briefing on facial recognition surveillance*. https://bigbrotherwatch.org.uk/wp-content/uploads/2020/06/Big-Brother-Watch-briefing-on-Facial-recognition-surveillance-June-2020.pdf

2. Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability and Transparency*, 149–159. https://doi.org/10.1145/3287560.3287598

3. Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

4. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

5. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). https://doi.org/10.1177/2053951715622512

6. Creemers, R. (2018). China's social credit system: An evolving practice of control. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3175792

7. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

8. Garvie, C., Bedoya, A., & Frankle, J. (2016). *The perpetual line-up: Unregulated police face recognition in America*. Georgetown Law Center on Privacy & Technology. https://www.perpetuallineup.org/

9. Heilweil, R. (2019, October 25). A Brooklyn landlord tried to install facial recognition. Tenants fought back — and won. *Vox*. https://www.vox.com/recode/2019/10/25/20930816/facial-recognition-apartment-privacy-brooklyn-icemae-downes

10. Kantayya, S. (Director). (2020). *Coded Bias* [Film]. 7th Empire Media.

11. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). https://doi.org/10.1177/2053951716679679

12. Vincent, J. (2016, March 24). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge*. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist