Undergraduate Thesis Prospectus

Algorithmic Bias: A Review of Research

(technical research project in Computer Science)

Hypocrisy or Adaptation: How BP
Defends Its Reputation in the Climate Emergency

(sociotechnical research project)

by

Noah Dela Rosa

November 2, 2020

Technical advisor: Aaron Bloomfield, Department of Computer Science

STS advisor: Peter Norton, Department of Engineering and Society

General Research Problem

How do sociotechnical systems distort perceptions?

Perceptions often depart from reality. Technology can mitigate or exacerbate such discrepancies. For example, algorithmic bias can encode historical social biases, thereby magnifying and perpetuating them (Johnson, 2020). In their public relations efforts, companies more deliberately apply sociotechnical methods to manage public perceptions, particularly in response to unfavorable publicity (Cleeren, van Heerde, & Dekimpe, M.G., 2013).

Algorithmic Bias: A Review of Research

What are researchers' leading findings on algorithmic bias?

Under the CS department and Aaron Bloomfield as acting technical advisor for the fall semester, my goal is to provide analysis-based suggestions for future work to eliminate systematic bias.

In the 21st century, it is unthinkable to not be within reach of a computer. The way these machines operate reflect the design choices made by the programmer. These choices form algorithmic bias, a challenge which regards systematic discrimination derived from design choices or input data. Only recent research has studied fairness in machine learning (Chouldechova and Roth, 2020). Sun, et al., (2020) ascribe the growth of algorithmic bias in algorithms to a shift in the sources of data from being "[based] on reliable labels from" to "more recently… receiving data from the general population [resulting in] bias that is born from ingesting unchecked information, such as biased samples and biased labels". As public infrastructure is reshaped by these algorithms, systemic biases will punctuate society if these algorithms remain unchecked.

Despite an explosion in the volume and rate of published work on algorithmic bias, the field is a new frontier of difficulties and discovery. Questions that would define fairness in machine learning loom large overhead, Chouldechova and Roth (2020) list a few, "What should fairness mean? What are the causes that introduce unfairness in machine learning? How best should we modify our algorithms to avoid unfairness? And what are the corresponding tradeoffs with which we must grapple?". There are currently two truths regarding algorithmic bias. First, any additional work being done is meaningless until a definition for fairness or system of fairness is established. Second, investigating the field through multiple lenses and perspectives is necessary.

To advise on how to eliminate algorithmic bias, I will compare research and findings on the sources of algorithmic bias, methodologies that curb the effect of distorting data, and current working definitions of fairness. I then hope to provide insight on how algorithmic bias is introduced and the best practices for prevention. These findings will hopefully be part of the foundation for a fairer technological future.

Hypocrisy or Adaptation: How BP Defends Its Reputation in the Climate Emergency

How does BP, despite depending on fossil fuels for its revenues, strive to promote a public reputation as a responsible global corporate citizen?

How can BP defend its reputation despite calls to action against climate change? As the globe is faced with continued record-breaking weather (Coumou & Rahmstorf, 2012), companies must address their strategy in preventing climate change. Thus, oil giants like BP, have reimagined their marketing strategies to retain good public standing.

Corporate social responsibility is corporate self-regulation to serve social values (Sheehy, 2015). Over 90% of Fortune 500 firms incorporate social responsibility initiatives within their own brands (Luo and Bhattacharya, 2006). To protect their reputations, many companies make public commitments to social values, though how serious these commitments are is difficult to determine. BP incorporates these strategies to reclaim their image after causing the Deepwater Horizon oil spill due to gross negligence (Robertson and Krauss, 2014).

Obvious participants include BP, which claims to seek net-zero carbon emissions (BP, 2019), Green America, critics of BP who seeks to create a sustainable society (Green America, 2019), American Fuel and Petrochemical Manufacturers, a lobbying group that liaises between the government and the interests of fuel retailers (AFPM, 2019), and the Democratic party which cites climate change as one of our nation's greatest challenges and condemns President Trump for calling it "a hoax invented by the Chinese." (Democratic Party, 2020).

BP claims net-zero carbon emissions is one of their biggest sustainability goals. They've instituted a 'reduce, improve, create' framework into their business model wherein existing emissions are reduced, products are optimized to provide lower emissions, and low carbon business is created through investments and expanding renewable businesses (BP, 2019). Through these measures and others, their claim has been consistent with the findings of Pinko, et al. (2018), whose analysis of eight different fossil energy companies' work toward meeting societal expectations to combat climate change. According to the study, BP was one of only two companies whose score in the "Company-Wide Commitments and Targets to Reduce Greenhouse Gas Emissions" category improved.

References

3

AFPM (2019). American Fuel & Petrochemical Manufacturers, 2019 Annual Report (p. 1, Rep.).
    (2019). Washington, D.C.: American Fuel & Petrochemical Manufacturers.

BP, BP Sustainability Report 2019 (Rep.). (2019).

Chouldechova, A., & Roth, A. (2020). A Snapshot of the Frontiers of Fairness in Machine
    Learning: A group of industry, academic, and government experts convene in Philadelphia
    to explore the roots of algorithmic bias. Communications of the ACM, 63(5), 82–89.
    https://doi.org/10.1145/3376898

Cleeren, K., van Heerde, H.J., & Dekimpe, M.G. (2013). Rising from the Ashes: How Brands
    and Categories Can Overcome Product-Harm Crises. *Journal of Marketing 77*(2), 58-77.

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. Nature Climate Change,
    2(7), 491-496.

Democratic Party, Environment. (n.d.)

Green America, 2019 Annual Report (Rep.). (2019). Retrieved September 16, 2020, from Green
    America website: https://www.greenamerica.org/annual-report-and-financials

Hardner, J., & Rice, R. (2002). Rethinking Green Consumerism. *Scientific American*, *286*(5),
    88–95. https://doi.org/10.1038/scientificamerican0502-88

Johnson, G.M. (2020). Algorithmic bias: on the implicit biases of social technology. *Synthese*
    448.

Luo, X., & Bhattacharya, C. (2006). Corporate Social Responsibility, Customer Satisfaction, and
    Market Value. Journal of Marketing, 70(4), 1-18. Retrieved November 3, 2020, from
    http://www.jstor.org/stable/30162111

Pinko, N., Mulvey, K., Ekwurzel, B., Frumhoff, P., Hurd, N., & Sideris, J. (2018). (Rep.). Union
    of Concerned Scientists. doi:10.2307/resrep24129

Robertson, C., & Krauss, C. (2014, September 4). *BP May Be Fined Up to $18 Billion for Spill
    in Gulf*. The New York Times. https://www.nytimes.com/2014/09/05/business/bp-
    negligent-in-2010-oil-spill-us-judge-rules.html.

Sheehy, B. (2015). Defining CSR: Problems and Solutions. Journal of Business Ethics, 131(3),
    625-648. Retrieved November 3, 2020, from http://www.jstor.org/stable/24703631

Sun, W., Nasraoui, O., & Shafto, P. (2020). Evolution and impact of bias in human and machine
    learning algorithm interaction. PLoS ONE, 15(8), 1–39.
    https://doi.org/10.1371/journal.pone.0235502

Vis, F., Faulkner, S., Noble, S., & Guy, H. (2020). When Twitter Got #woke: Black Lives Matter, DeRay McKesson, Twitter, and the Appropriation of the Aesthetics of Protest. In McGarry A., Erhart I., Eslen-Ziya H., Jenzen O., & Korkut U. (Eds.), The Aesthetics of Global Protest: Visual Culture and Communication (pp. 247-266). Amsterdam: Amsterdam University Press. doi:10.2307/j.ctvswx8bm.18