



a fresh on squeeze data

www.freshsqueezekids.com

Eradicating Data Bias



From a business as local as a lemonade stand to a company operating on a global scale, relying on data is a must. But trusting that data is trickier. The expression “numbers don’t lie” embodies a belief in the unemotional truth revealed by data, and the analytical prowess of machines seemingly incapable of bias or error. AI projects, however, and the algorithms that power them, are programmed by the human hand, and the same inconsistencies, oversights and prejudices that plague our human world exist in the digital one — even more so since machines are incapable of reasoning on the fly the same way humans can.

Recognizing and eradicating data bias is a major undertaking, necessary to protect the health and livelihood of all communities. It’s becoming increasingly urgent as AI



technologies grow in prominence. Algorithms are already involved in innumerable decisions with real-life consequences: where people work and live, how much money they make, and what healthcare treatment they do or do not receive.



How then does bias infect data, and what is being done to confront the powerful—but often invisible—hand of imperfect algorithms?

Bias in Action

Fixing a problem starts with understanding it, and the past few years provide numerous examples of biased data in action.

Research has shown that algorithmic hiring models can contain bias in numerous ways and can occur in processes ranging from natural language analysis used in resume scanning to speech recognition and commercial AI facial analysis. The broad result of this ever-growing reliance on algorithmic models in the job market is the biased employment outcomes for different cross-sections of the population.

San Francisco became the first major American city to **ban the use of facial recognition technology** in the midst of growing concerns over both its tendency to produce **much higher false positive results** for African American and Asian faces, and unease over privacy issues.

In 2019, **a study revealed** that an algorithm widely used in US hospitals for allocating treatment to patients systematically discriminated against African Americans. This bias led to less care and fewer total dollars spent in comparison to white patients with similar or in many cases better overall health. The algorithm assigned risk scores to patients based on accrued healthcare costs from the previous year. It did not account for systemic racism within the healthcare system nor a historical



legacy of mistrust within Black communities that can lead to patients delaying or refusing care entirely.



While examples vary the impact is clear: fairness is not a given when computers make decisions. Bias can creep into a machine learning model in a variety of ways, some of which reflect larger social prejudices and others stem from specific shortcomings of a particular algorithm. Every stage of model design can be susceptible, be it data collection, curation, or analysis.



Types of Bias

Implicit or unconscious bias can easily transfer from the designer of a model to the model itself, as their own life experience and background can dictate the terms of the program they create in ways that are invisible to them. This is a constant challenge for scientists and engineers attempting to build bias-free algorithms.



Sampling bias is when random data selected from the population does not accurately represent the distribution of the population, leading to incorrect statistical outcomes.



Temporal bias occurs when a model does not account for changes that occur over time, making the original algorithm susceptible to false or misleading conclusions.



Overfitting to training data, edge cases and outliers are other ways bias can influence data and lead to damaging outcomes.

Moving in the Right Direction

Thankfully, the significance of unjust bias in data has created energy around finding solutions.



It starts with prioritizing more diversity and inclusivity within the data science teams that build the models in the first place. Accounting for a range of perspectives across race, gender, class, and geography is less apt to fall short if the design team reflects that same diversity.



Cloudera is among an emerging cohort of companies **investing in individuals** intent on investigating the larger ramifications of data on our society. Cloudera views this as a starting point for data-centric organizations looking to drive the future of more equitable computer models.



Awareness of the dangers and prevalence of algorithmic bias has caused industries, organizations and governments to reexamine the processes behind them. The discriminatory healthcare algorithm mentioned above put a glaring spotlight on determining patient outcomes, forcing both providers and data organizations to do better or face mounting financial, legal, and cultural penalties.



In April of 2021, the **Federal Trade Commission** declared that businesses using biased algorithms to make automated decisions are violating federal law, with enforcement steps to follow. Besides legal ramifications, companies that design models shown to contain troubling bias can suffer severe financial and reputational damage as enforcement agencies can dig into structural issues of the company while customers look for alternative services.



In 2018, **France pledged** to make all government algorithms open to the public. Similarly, the European Union's **General Data Protection Regulation** from 2018 sets in stone the rights of individuals to data protection and privacy.



In the U.S., **there are** several proposed bills aimed at curtailing algorithmic bias, including the Algorithmic Accountability Act of 2019 and the Consumer Online Privacy Rights Act. Both the Justice in Policing Act and Facial Recognition and Biometric Technology Moratorium Acts are designed to restrict facial recognition technology in law enforcement.

Such directives help, and a long-term commitment from all the invested parties can continue building momentum towards design, analysis, and accountability steps that ensure algorithmic bias is at least mitigated and at best eliminated.

Improving data literacy in today's youth will help eradicate data bias in the future. Share this book with the children in your life to help teach them the importance of data and to further understand how bias impacts their everyday lives.

PRESENTED BY

CLOUDERA

© 2021 All rights reserved

