WEDNESDAY, AUGUST 31ST, 2022

THE TEXAS ORATOR
academic-ish

TECH

How Black Box Algorithms Blur the Line Between Ethics and Technological Development

BY SWETHA81200 ON DECEMBER 7, 2018 • (LEAVE A COMMENT)

In a world where Google tells us what to search for, Amazon delivers groceries to our homes, and Tesla creates self-driving cars, the technological capabilities of computers have transcended their original purpose. The eminence of machine learning has rapidly increased in recent years, culminating in a dense web of neural networks that have gotten far too complicated for even their own creators to untangle. To really get to the bottom of the ethical issues of technology, however, it is important to understand the basics of artificial intelligence, neural networks, and black box algorithms.

The concept of programming began with directly giving a machine logical instructions to perform certain tasks. Artificial intelligence, or a machine's ability to make decisions using real-time data rather than pre-programmed instructions, arose next. As artificial intelligence developed further, machine learning evolved. Machine learning utilizes statistical data and analysis to provide computer systems the ability to "learn" from data and self-perfect their performance over time without explicitly being programmed to do so. A good example is the recruiting software, HireVue (https://www.hirevue.com/), which analyzes facial patterns and response times of potential clients, painting a more objective picture of a candidate for hire.

How exactly do researchers get machines to learn? There are various models, from simple statistical models like the Bayesian (https://www.datascience.com/blog/introduction-to-bayesian-inference-learn-data-science-tutorials) inference model to more complicated neural networks. Neural networks, similar to the intricate neural connections in a human brain, are adaptive systems that use a series of interconnected units to adjust outputs based on data received from the outside world. Put more simply, neural networks eliminate (https://towardsdatascience.com/why-deep-learning-is-needed-over-traditional-machine-learning-1b6a99177063) the need for a "middleman" to continuously organize and translate data being fed into machines to actual actions — machines can now act, using outside feedback on their own, to produce some sort of output.

Older machine learning methods are being replaced with this newer, deeper learning through neural networks. Technologies like word completion, autonomous vehicles, and translation apps all utilize neural networks. Besides the consumer sphere, however, this technology is being implemented in spheres such as national defense and government. (https://medium.com/@jayeshbahire/real-world-applications-of-artificial-neural-networks-a6a6bc17ad6a) While this makes operations vastly more efficient, there are many inherent problems with trusting even the most intelligent machine to come up with decisions based on statistical data.

An example that illustrates this on a smaller scale is Google's search results order algorithm. While we may think we have it generally figured out, the actual formula, based on millions and millions of user inputs and research, is so dense and intricate that nobody has actually pinpointed the algorithm. The term "black box (https://www.propublica.org/article/breaking-the-black-box-when-machines-learn-by-experimenting-on-us)" refers to this ambiguity. While a neural network's creators have a general idea of the purpose of the algorithm, and its users are generally well aware of the data inputted and outputted, the path, or "black box" that the input data takes to arrive at an output, is largely hidden.

While this seems harmless when viewed in the context of our everyday interactions with neural networks, such as Google search and translator applications, even these seemingly everyday applications reveal a lot about the potential dangers of letting the machine take control. Take Google Photos' mislabeling (http://www.nydailynews.com/news/national/google-photos-algorithm-mislabeled-people-gorillas-article-1.2278049) of an image of a black couple as gorillas and Amazon bookseller bots (https://www.theatlantic.com/technology/archive/2011/04/want-to-see-how-crazy-a-bot-run-market-can-be/237773/) bidding against each other until a book's price exceeded $23 million.

ProPublica conducted in-depth studies (https://www.propublica.org/article/breaking-the-black-box-when-algorithms-decide-what-you-pay) of both the purely social implications of black box algorithms and their potentially more serious implications in sectors like criminal justice. While it is evident that Google filters search results

and orders them based on an intricate algorithm, what we don't realize is that over time, the search engine can be trained to display partisan data to a user based on millions of prior searches. For example, after analyzing a user's keystrokes over an extended period of time and even his or her usage of certain words, such as "thrice-married" (in reference to President Trump), Google algorithms have "learned" to provide the user with news sources like Huffington Post and the New York Times, which lean more left, ordering them above sources like Breitbart (a right-leaning source) in search results. Of course, the problem is not that these hidden algorithms try to feed information to us out of nowhere — it is that they take our own biases and feed them back to us. In a world where technology is supposed to expand our horizons and bring the world closer together, this is especially dangerous.

Indirectly, processes like these, which promote information bias, have the potential to contribute to hyper-partisanship, a problem (https://www.vox.com/the-big-idea/2017/9/5/16227700/hyperpartisanship-identity-american-democracy-problems-solutions-doom-loop) which is already starting to plague our country. The investigation also found (https://www.propublica.org/article/breaking-the-black-box-when-algorithms-decide-what-you-pay) that when parents were searching for college preparation programs for their children, the Princeton Review's price algorithm would change its prices simply based on zip code. The algorithm "learned" to use factors like average household income and percentage Asian population to adjust the pricing for the Princeton Review's services.

Neural networks are utilized for purposes beyond consumer applications, such as in risk assessment modeling (https://www.technologyreview.com/s/609338/new-research-aims-to-solve-the-problem-of-ai-bias-in-black-box-algorithms/), where race technically isn't one of the criteria that algorithms correlate with chances of being granted bail or approved for a loan. However, neural networks, much like the human brain, can quickly learn to correlate certain zip codes with certain income levels. After going through thousands of data inputs, the network can "refine" its algorithm to contain an innate racial bias. What makes the situation so tricky is that nobody knows what exactly goes on under the hood of the black box algorithm.

Even in criminal defense, risk assessment models used in courtrooms have been found (https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing) to demonstrate innate racial bias. The situation is further complicated by the fact that those who implement the algorithms in government are even less (https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html) likely to want to reveal their algorithm to third-party technical specialists because the code is hailed as intellectual property. In 1989, Tim Brennan, a statistician whose main objective was to create an algorithm which measures criminal recidivism, founded the company Northpointe (https://www.technologyreview.com/s/607955/inspecting-algorithms-for-bias/). The name of the algorithm was Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS. It assesses not just risk but also

"criminogenic needs," "criminal personality," "social isolation," "substance abuse," and "residence/stability." Defendants are ranked low, medium, or high risk in each category.

ProPublica conducted an investigation (https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing) into the tool and found that it is decently accurate, predicting recidivism 61 percent of the time. However, there comes a caveat — black offenders are almost twice as likely (https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing) to be labeled as re-offenders but not actually re-offend in comparison to white people. White offenders, meanwhile, are much more likely to not be labeled but go on to perform further crimes. Broward County, in southeastern Florida, decided in 2008 that instead of building another jail, it would start implementing Northpointe's algorithm to decide which criminals were low risk enough to be released on bail. It cited (https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm) COMPAS, at $22,000 a year, as much more cost-effective than employing human capital for the same purpose. Additionally, the sheriff's office indicated that the simplistic charts and data seemed, on paper, much more objective and easy to review than other forms of risk assessment. Brennan explains (https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing) that while his original intent was not for COMPAS to completely replace other forms of risk assessment, it is nearly impossible to eliminate race-related factors from the algorithm, such as income and joblessness, while still maintaining fairly accurate results.

Rich Caruana (https://cacm.acm.org/news/214618-in-black-box-algorithms-we-trust-or-do-we/fulltext), a senior researcher at Microsoft Research in Redmond, Washington, cites his more than 20 years of experience in the field when he explains, "Many people do not realize that the problem is often in the data, as opposed to what machine learning does with the data. It depends on what you are doing with the model whether the data are used in the right or in the wrong way." While the usage of machine learning will rapidly escalate the capabilities of technology as we know it, it is evident that dense black box algorithms are beginning to hold too much power over citizens who cannot see their underlying workings. The companies and governmental agencies which utilize these networks must subject these algorithms to routine testing and research and release findings publicly to improve transparency.

Published by swetha81200

View all posts by swetha81200 (https://thetexasorator.com/author/swetha81200/)