DECISION MAKING

Want Less-Biased Decisions? Use Algorithms.

by Alex P. Miller

JULY 26, 2018

ORLAGH MURPHY/GETTY IMAGES

A quiet revolution is taking place. In contrast to much of the press coverage of artificial intelligence, this revolution is not about the ascendance of a sentient android army. Rather, it is characterized by a steady increase in the automation of traditionally human-based decision processes throughout organizations all over the country. While advancements like AlphaGo Zero make for catchy headlines, it is fairly conventional

machine learning and statistical techniques – ordinary least squares, logistic regression, decision trees – that are adding real value to the bottom line of many organizations. Real-world applications range from medical diagnoses and judicial sentencing to professional recruiting and resource allocation in public agencies.

Is this revolution a good thing? There seems to be a growing cadre of authors, academics, and journalists that would answer in the negative. Book titles in this genre include *Weapons of Math Destruction*, *Automating Inequality*, and *The Black Box Society*. There has also been a spate of exposé-style longform articles such as "Machine Bias," "Austerity Is an Algorithm," and "Are Algorithms Building the New Infrastructure of Racism?" At the heart of this work is the concern that algorithms are often opaque, biased, and unaccountable tools being wielded in the interests of institutional power. So how worried should we be about the modern ascendance of algorithms?

INSIGHT CENTER

Adopting AI

How companies are using artificial intelligence in their business operations.

These critiques and investigations are often insightful and illuminating, and they have done a good job in disabusing us of the notion that algorithms are purely objective. But there is a pattern among these critics, which is that they rarely ask how well the systems they analyze would operate *without algorithms*. And that is the most relevant question for practitioners and policy makers: How do the bias and performance of algorithms compare with the status quo? Rather than simply asking whether algorithms are flawed, we should be asking how these flaws compare with those of human beings.

What Does the Research Say?

There is a large body of research on algorithmic decision making that dates back several decades. And the existing studies on this topic all have a remarkably similar conclusion: Algorithms are less biased and more accurate than the humans they are replacing. Below is a sample of the research about what happens when algorithms are given control of tasks traditionally carried out by humans (all emphasis mine):

- In 2002 a team of economists studied the impact of automated underwriting algorithms in the mortgage lending industry. Their primary findings were "that [automated underwriting] systems more accurately predict default than manual underwriters do" and **"that this increased accuracy results in higher borrower approval rates, especially for underserved applicants."** Rather than marginalizing traditionally underserved home buyers, the algorithmic system actually benefited this segment of consumers the most.

- A similar conclusion was reached by Bo Cowgill at Columbia Business School when he studied the performance of a job-screening algorithm at a software company (forthcoming research). When the company rolled out the algorithm to decide which applicants should get interviews, **the algorithm actually favored "nontraditional" candidates** much more than human screeners did. Compared with the humans, the algorithm exhibited significantly less bias against candidates that were underrepresented at the firm (such as those without personal referrals or degrees from prestigious universities).

- In the context of New York City pre-trial bail hearings, a team of prominent computer scientists and economists determined that algorithms have the potential to achieve significantly more-equitable decisions than the judges who currently make bail decisions, with "jailing rate reductions [of] up to 41.9% with no increase in crime rates." They also found that in their model "all categories of crime, including violent crimes, show reductions [in jailing rates]; and **these gains can be achieved while simultaneously reducing racial disparities."**

- The New York Times Magazine recently reported a longform story to answer the question, "Can an algorithm tell when kids are in danger?" It turns out the answer is "yes," and that algorithms can perform this task much more accurately than humans. Rather than exacerbating the pernicious racial biases associated with some government services, "the Allegheny experience suggests that **its screening tool is less bad at weighing biases than human screeners have been."**

- Lastly, by looking at historical data on publicly traded companies, a team of finance professors set out to build an algorithm to choose the best board members for a given company. Not only did the researchers find that companies would perform better with algorithmically selected board members, but **compared with their proposed algorithm, they "found that firms [without algorithms] tend to choose directors who are much more likely to be male**, have a large network, have a lot of board experience, currently

serve on more boards, and have a finance background."

In each of these case studies, the data scientists did what sounds like an alarming thing: They trained their algorithms on past data that is surely biased by historical prejudices. So what's going on here? How is it that in so many different areas — credit applications, job screenings, criminal justice, public resource allocations, and corporate governance — algorithms can be reducing bias, when we have been told by many commentators that algorithms should be doing the opposite?

Human Beings Are Remarkably Bad Decision Makers

A not-so-hidden secret behind the algorithms mentioned above is that they actually *are* biased. But the humans they are replacing are *significantly more biased*. After all, where do institutional biases come from if not the humans who have traditionally been in charge?

But humans can't be all that bad, right? Yes, we may be biased, but surely there's some measure of performance on which we are good decision makers. Unfortunately, decades of psychological research in judgment and decision making has demonstrated time and time again that humans are remarkably bad judges of quality in a wide range of contexts. Thanks to the pioneering work of Paul Meehl (and follow-up work by Robyn Dawes), we have known since at least the 1950s that very simple mathematical models outperform supposed experts at predicting important outcomes in clinical settings.

In all the examples mentioned above, the humans who used to make decisions were so remarkably bad that replacing them with algorithms both increased accuracy *and* reduced institutional biases. This is what economists call a *Pareto improvement*, where one policy beats out the alternative on every outcome we care about. While many critics like to imply that modern organizations pursue the operational efficiency and greater productivity at the expense of equity and fairness, all available evidence in these contexts suggests that there is no such trade-off: Algorithms deliver more-efficient *and*

more-equitable outcomes. If anything should alarm you, it should be the fact that so many important decisions are being made by human beings who *we know* are inconsistent, biased, and phenomenally bad decision makers.

Improving on the Status Quo

Of course, we should be doing all we can to eradicate institutional bias and its pernicious influence on decision-making algorithms. Critiques of algorithmic decision making have spawned a rich new wave of research in machine learning that takes more seriously the social and political consequences of algorithms. There are novel techniques emerging in statistics and machine learning that are designed specifically to address the concerns around algorithmic discrimination. There is even an academic conference every year at which researchers not only discuss the ethical and social challenges of machine learning but also present new models and methods for ensuring algorithms have a positive impact on society. This work will likely become even more important as less-transparent algorithms like deep learning become more common.

But even if technology can't fully solve the social ills of institutional bias and prejudicial discrimination, the evidence reviewed here suggests that, in practice, it can play a small but measurable part in improving the status quo. This is not an argument for algorithmic absolutism or blind faith in the power of statistics. If we find in some instances that algorithms have an unacceptably high degree of bias in comparison with current decision-making processes, then there is no harm done by following the evidence and maintaining the existing paradigm. But a commitment to following the evidence cuts both ways, and we should to be willing to accept that — in some instances — algorithms will be part of the solution for reducing institutional biases. So the next time you read a headline about the perils of algorithmic bias, remember to look in the mirror and recall that the perils of human bias are likely even worse.

Alex P. Miller is a Doctoral Candidate in Information Systems & Technology at the University of Pennsylvania's Wharton School.

This article is about DECISION MAKING

⊕ Follow This Topic

Related Topics: Data | Technology | Policy