

Analytics

Hiring Algorithms Are Not Neutral

by Gideon Mann and Cathy O'Neil

December 09, 2016



More and more, human resources managers rely on data-driven algorithms to help with hiring decisions and to navigate a vast pool of potential job candidates. These software systems can in some cases be so efficient at screening resumes and evaluating personality tests that 72% of resumes are weeded out before a human ever sees them. But there are drawbacks to this level of efficiency. Man-made algorithms

are fallible and may inadvertently reinforce discrimination in hiring practices. Any HR manager using such a system needs to be aware of its limitations and have a plan for dealing with them.

Algorithms are, in part, our opinions embedded in code. They reflect human biases and prejudices that lead to machine learning mistakes and misinterpretations. This bias shows up in numerous aspects of our lives, including algorithms used for electronic discovery, teacher evaluations, car insurance, credit score rankings, and university admissions.

Algorithms deployed on the front line of HR decision making may cut costs and streamline vetting in busy departments or companies with large hiring needs, but the risk is that they end up excluding applicants based on gender, race, age, disability, or military service — all protected classes under employment law.

INSIGHT CENTER

The Next Analytics Age

Harnessing the power of machine learning and other technologies.

For HBR Subscribers

AI, Algorithms, and Bias

Algorithms can improve our predictions and decisions, but they can also perpetuate our blind spots and biases. Here's what you need to know about the problem—and how organizations can address it.

Show Reading List

At their core, algorithms mimic human decision making. They are typically trained to learn from past successes, which may embed existing bias. For example, in a famous experiment, recruiters reviewed identical resumes and selected more applicants with white-sounding names than with black-sounding ones. If the algorithm

learns what a “good” hire looks like based on that kind of biased data, it will make biased hiring decisions. The result is that automatic resume screening software often evaluates job applicants based on subjective criteria, such as one’s name. By latching on to the wrong features, this approach discounts the candidate’s true potential.

In other words, algorithms are not neutral. When humans build algorithmic screening software, they may unintentionally determine which applicants will be selected or rejected based on outdated information — going back to a time when there were fewer women in the workforce, for example — leading to a legally and morally unacceptable result.

One way to avoid algorithmic bias is to stop making hard screening decisions based solely on an algorithm. Encourage a human review that will ask experienced people who have been through bias training to oversee selection and evaluation. Let decisions be guided by an algorithm-informed individual, rather than by an algorithm alone.

If every HR decision can’t be reviewed by human eyes, your first step should be to accept that algorithms are imperfect. While there are good algorithms that have been properly calibrated to efficiently and accurately measure results, such success doesn’t happen by accident. We need to audit and modify algorithms so that they do not perpetuate inequities in businesses and society. Consider assigning a team or hiring outside professionals to audit key algorithms. The use of multiple algorithms might also help limit blind spots. This way, no single metric, such as SAT score, would exclude a qualified candidate.

Furthermore, you can advocate for a more rigorous observation and control system. Periodically carry out random spot-checks on machine resume decisions and put them through an extensive human review to see which candidates the algorithm has been selecting and why, with a strong emphasis on uncovering potential cases of bias.

Finally, regularly conduct manual reviews of the correlations that the machine learns and selectively eliminate any that may appear to be biased. You shouldn't base your hiring decisions on correlations related to a person's name or extracurricular activities, for example, as these may be indicative of a candidate's age, race, or class — not their qualifications.

HR managers should also make themselves aware of algorithms that can help mitigate their own human biases. For instance, sometimes bias starts with the job listing. Without realizing it, writers sprinkle copy with keywords or phrases that sound hip but contain cues that dissuade certain candidates from applying for the role. In tech recruiting, snippets like “coding ninja wanted” have a connotation that is likely to dampen the interest of older candidates and women. Many businesses are so worried about ads with hidden signifiers that they utilize machine learning platforms like Textio (a Bloomberg Beta portfolio company) to flag potentially problematic words and suggest alternatives that will perform better. Programs like Unitive feature software that helps employers write inclusive job descriptions and holds hiring managers accountable throughout the interview process if they disregard criteria that they said was important.

Can bias and prejudice be completely eliminated from hiring? For both algorithmic and human methods, the answer is “probably not.” But as big data, machine-learning algorithms, and people analytics take on a larger and more influential role in recruiting, HR professionals must consider the consequences of these systems and ensure they always reflect the best human intentions.

GM

Gideon Mann is head of data science at Bloomberg, L.P.

CO

Cathy O’Neil is the founder of **ORCAA**, a risk consulting and algorithmic auditing firm, and the author of *Weapons of Math Destruction*.