The Upshot

HIDDEN BIAS

When Algorithms Discriminate

Claire Cain Miller @clairecm   JULY 9, 2015

The online world is shaped by forces beyond our control, determining the stories we read on Facebook, the people we meet on OkCupid and the search results we see on Google. Big data is used to make decisions about health care, employment, housing, education and policing.

But can computer programs be discriminatory?

There is a widespread belief that software and algorithms that rely on data are objective. But software is not free of human influence. Algorithms are written and maintained by people, and machine learning algorithms adjust what they do based on people's behavior. As a result, say researchers in computer science, ethics and law, algorithms can reinforce human prejudices.

Google's online advertising system, for instance, showed an ad for high-income jobs to men much more often than it showed the ad to women, a new study by Carnegie Mellon University researchers found.

Research from Harvard University found that ads for arrest records were significantly more likely to show up on searches for distinctively black names or a historically black fraternity. The Federal Trade Commission said advertisers are able to target people who live in low-income neighborhoods with high-interest loans.

Research from the University of Washington found that a Google Images search for "C.E.O." produced 11 percent women, even though 27 percent of United States

chief executives are women. (On a recent search, the first picture of a woman to appear, on the second page, was the C.E.O. Barbie doll.) Image search results determined 7 percent of viewers' subsequent opinions about how many men or women worked in a field, it found.

"The amoral status of an algorithm does not negate its effects on society," wrote the authors of the Google advertising study, Amit Datta and Anupam Datta of Carnegie Mellon and Michael Carl Tschantz of the International Computer Science Institute.

Algorithms, which are a series of instructions written by programmers, are often described as a black box; it is hard to know why websites produce certain results. Often, algorithms and online results simply reflect people's attitudes and behavior. Machine learning algorithms learn and evolve based on what people do online. The autocomplete feature on Google and Bing is an example. A recent Google search for "Are transgender," for instance, suggested, "Are transgenders going to hell."

"Even if they are not designed with the intent of discriminating against those groups, if they reproduce social preferences even in a completely rational way, they also reproduce those forms of discrimination," said David Oppenheimer, who teaches discrimination law at the University of California, Berkeley.

But there are laws that prohibit discrimination against certain groups, despite any biases people might have. Take the example of Google ads for high-paying jobs showing up for men and not women. Targeting ads is legal. Discriminating on the basis of gender is not.

The Carnegie Mellon researchers who did that study built a tool to simulate Google users that started with no search history and then visited employment websites. Later, on a third-party news site, Google showed an ad for a career coaching service advertising "$200k+" executive positions 1,852 times to men and 318 times to women.

The reason for the difference is unclear. It could have been that the advertiser requested that the ads be targeted toward men, or that the algorithm determined that men were more likely to click on the ads.

Google declined to say how the ad showed up, but said in a statement, "Advertisers can choose to target the audience they want to reach, and we have policies that guide the type of interest-based ads that are allowed."

Anupam Datta, one of the researchers, said, "Given the big gender pay gap we've had between males and females, this type of targeting helps to perpetuate it."

It would be impossible for humans to oversee every decision an algorithm makes. But companies can regularly run simulations to test the results of their algorithms. Mr. Datta suggested that algorithms "be designed from scratch to be aware of values and not discriminate."

"The question of determining which kinds of biases we don't want to tolerate is a policy one," said Deirdre Mulligan, who studies these issues at the University of California, Berkeley School of Information. "It requires a lot of care and thinking about the ways we compose these technical systems."

Silicon Valley, however, is known for pushing out new products without necessarily considering the societal or ethical implications. "There's a huge rush to innovate," Ms. Mulligan said, "a desire to release early and often — and then do cleanup."

The Upshot provides news, analysis and graphics about politics, policy and everyday life. Follow us on Facebook and Twitter. Sign up for our newsletter.

A version of this article appears in print on July 13, 2015, on page B1 of the New York edition with the headline: Algorithms May Echo Human Bias, Studies Find.