



Don't miss out on diverse talent or customers.

Consider the Risks of Algorithmic Bias

By Noah Cotton, Emily Kelly, and Deonesha Williams

→ When Deep Blue beat Garry Kasparov at chess in 1997, it was considered a major milestone in the development of artificial intelligence. A computer outperformed the human brain at a very complex task. Of course, the task was extremely well-defined. Chess rules are specific and coding them into the software wasn't hard. The computer's true limitation was its own raw processing power. Once Deep Blue could perform enough calculations in a shortened amount of time, no human stood a chance.

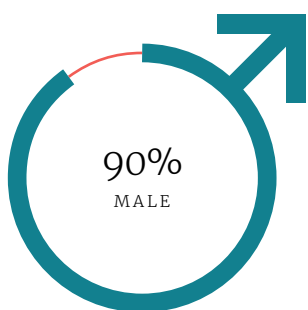
Today, we ask computers to solve much more complicated problems without clearly defined rules and neat grids. In these cases, the humans designing the machines must make choices that will affect the machines' understanding of the rules. But there's one big problem: humans have biases that they unintentionally code into their solutions. This phenomenon is more formally known as Algorithmic Bias, and it is defined as repeatable and systematic errors that create

“unfair” outcomes in favor of one category over another in ways different than initially intended by a system's algorithm.

For decision-making individuals within organizations, Algorithmic Bias can be subtle, but impactful. It can also have a negative influence on the strategic and operational choices that are made if thoughtful consideration is not given to the inputs of each output. While any algorithm-supported decision has the potential to help drive business, there may be unintended consequences to different subsets of employees and market segments. Throughout the article, we'll highlight several scenarios that demonstrate algorithmic biases against specific qualities — such as gender, race, and religion — to show organizational decision-makers the importance of being cautious when attempting to leverage algorithms in their operations.



RESEARCHERS FOUND THAT IN COUNTRIES WITH HIGHER GENDER INEQUALITY, GOOGLE IMAGE SEARCHES FOR GENDER-NEUTRAL TERMS SUCH AS “PERSON” PRODUCED UNEVEN RESULTS WITH RESPECT TO GENDER. IN THE MOST EXTREME CASES, AROUND 90 PERCENT OF IMAGES RETURNED BY THE SEARCH FOR “PERSON” WERE MEN.



Recruiting Challenges

Many artificial intelligence algorithms are trained on large public data sets. For example, a program can use internet search results to learn about any given situation. Unfortunately, those search results may not be neutral and could introduce a bias to the program. In a recent study, researchers at New York University found that in countries with higher gender inequality, Google image searches for gender-neutral terms such as “person” produced uneven results with respect to gender. In the most extreme cases, around 90 percent of images returned by the search for “person” were men.

Let’s consider a hypothetical scenario where, in an attempt to hire an engineer, a company has deployed an algorithm to screen potential candidates. This algorithm might first use a large data set to evaluate what characteristics are common to successful engineers. Perhaps it uses LinkedIn or another professional networking website and makes note of job titles while reviewing people’s profile information, or it may use internal data. In the United States, only about 15 percent of engineers are women. As a result, unless explicitly told otherwise, an algorithm may assume women are inherently less qualified than men to be engineers. And when it comes time to screen applicants, anyone with an “F” under gender may receive a small penalty on their candidate score right from the start. (And if a person lists “Other,” that may prove an even larger hindrance.) An engineering firm presumably

wants to hire the best candidates for the job. By relying on a biased algorithm, however, the firm will likely miss out on hiring top female engineers or other talented individuals who slip through the screening and end up working elsewhere.

Good programmers may be aware of this problem and explicitly code their algorithms to ignore gender in candidate screening. Unfortunately, it’s not necessarily that simple.

Here’s another example: About 90 percent of U.S. armed services veterans are male. And while veterans as a group may be at a disadvantage in a number of scenarios, in this case, we’re focusing only on the correlation with gender.

But what if a particular job, such as electrician, has a history of employing more men than women? Military experience isn’t specifically relevant to the job, but probability suggests veterans may be over-represented in the ranks of electricians. An algorithm selecting candidates for a training program for future electricians, looking for correlations with profiles of current electricians, may give extra weight to military service. In this case, the automated screening may indirectly discount female applicants with other types of career experience. Taking it a step further, members of the LGBTQ community, who were mostly barred from the military until 2011, are even more likely to be excluded by the same implicit criteria.

It’s impossible to predict and correct for all of these nuanced correlations, and harder still to predict the negative externalities they impose on the group relying on the algorithm for decision support.



Talent Development Challenges

From another perspective, algorithmic bias may cause some companies to miss out on diverse talent before they even have the chance to apply. For example, flawed decisions by the marketing department, or in a brand awareness campaign, might impact how the company attracts talent from the outset.

It may also be happening further upstream at the academic institutional level. Alongside companies and government institutions, educational institutions, both private and public, are using algorithms to improve the quality of education their students receive. While these systems are well-intentioned in theory, the reality may be that the talent pool of diverse applicants is diluted. Biased algorithms, based on guidance

recommendations made by advisors, could impact whether a student is accepted into a graduate program.

Black, Latinx, and other minority students already often face systemic challenges rooted in generational financial inequity, and one of the most effective ways to overcome this obstacle is to pursue higher education. Additionally, these students can pursue challenging courses and degrees in fields like engineering, finance, and computer science to accelerate their careers and earnings.

But what happens if these students are steered away from this education entirely based on their racial identity? A 2021 study conducted by The Markup, a nonprofit that focuses on data-driven journalism, showed that universities using software called Navigate were able to use race as a “high-impact” variable in an algorithm designed to help advisors guide students. The results showed

that in instances where race was used as a variable, Black and Latinx students were many times more likely to be labeled as “high risk.” And if a student is considered a “high risk” of failing a class, it’s unlikely that an advisor would feel comfortable challenging them to take a harder course or enroll in a more competitive program.

Even diverse talent that has completed higher-level education and seeks further development can be algorithmically limited from the talent pool. AI technology used by the University of Texas at Austin’s Department of Computer Science for Ph.D. admissions was discontinued after it was discovered that pre-existing human bias was essentially built into the system by using previous years of admissions numbers as its main data source. It is entirely possible that this machine-learning algorithm unintentionally screened out some of the top candidates simply due to the implicit biases (related to gender, race, or other factors) that humans possess. The downstream impact of that tool being used? A less diverse, more homogenous, and potentially less developed talent pool from which to hire.

Market Development Challenges

Now that we’ve examined how algorithmic bias negatively affects people by gender and race, let’s look at how it can also single out individuals based on religion. As businesses continue to incorporate technology into their daily operations, there are more factors



NIKE DEVELOPED THE SUCCESSFUL **PRO HIJAB** BY USING INCLUSIVE DATA TO TARGET AN AUDIENCE THAT WAS BEING UNINTENTIONALLY EXCLUDED.

PHOTO_NIKE

that need to be taken into account to make conscious business decisions. Consider a branding opportunity: an organization implements a new artificial intelligence system that generates targeted advertising text. This sort of assistance could streamline the production of marketing materials and expand the reach of the team. And while this tool has the potential to be a great addition and resource, if it's relying on natural language, it will have to be fed the proper data to produce unbiased results.

According to a study conducted by Stanford researchers, when an artificial intelligence system was instructed to complete unfinished sentences, the AI formulated sentences that were negative toward one particular religious group: Muslims. The researchers found that when using a particular internet-based data set, the system disproportionately associated Muslims with violence. However, when they swapped the term "Muslims" with "Christians" in the test sentences, the AI went from providing violent associations 66 percent of the time to only 20 percent.

In part, this was due to the concept of "garbage in, garbage out." If you train an AI based on data that humans have put on the internet, the AI will end up replicating whatever human biases are in that data. This aspect of AI programming is easy to overlook when you aren't the affected party.

And while clearly the humans on the marketing team would prevent the publication of actively discriminatory language, it's not hard to imagine a scenario where flawed language associations produce material that unintentionally excludes

a portion of a company's customer base. In these situations, training AI systems with carefully screened inclusive data would make a world of difference for a company's reputation and profit margins.

Opportunity

Some companies are working to address the problems associated with algorithmic biases. Nike, for example, took a head-on approach to religious inclusivity within the fashion industry with its production of the Nike Pro Hijab. Nike looked at its target audience and realized that a portion, whether large or small, was being unintentionally excluded. From there, the company conducted research using an array of inclusive resources and feedback to ensure that its product and marketing campaign appealed to the same audience that was previously excluded.

Since its release in December 2017, the headpiece has attracted steadily growing interest. During the first quarter of 2019, its popularity rose by 125 percent, seeing it enter the top 10 most-wanted items in luxury fashion. Based on those numbers, it seems that Nike may have been on to something by using inclusive data to help set the foundation of its product research and branding.



This article is not meant to suggest that all algorithms are inherently biased. Obviously, the Nike example shows us that data-supported decision-making can, in fact, be more inclusive — and that using creative problem-solving to overcome bias,

rather than trying to build a perfect algorithm, can work wonders. By identifying an underserved group and intelligently targeting it, Nike was able to make a product that was good for its bottom line and good for the community.

Similarly, using algorithms to identify potential talent creates the possibility of screening larger candidate pools, and data-driven approaches can and often do effectively provide assistance to human decision-makers. At the end of the day, we each have to do our part to work with inclusive data for the continued advancement of present and future AI systems. We know that biases can be as complex as the algorithms themselves, but we must commit to reducing how often they occur through deliberate and conscious actions.

Noah Cotton

noah.cotton@jabian.com

Emily Kelly

emily.kelly@jabian.com

Deonesha Williams

deonesha.williams@jabian.com