

Algorithm bias: Everything you need to know

Author: Tatenda Emma Matika . April 2021

The world nowadays is pro-inclusion. This includes inclusion for minority groups, for strong opinions that go against society, for underserved societies, and so on. Various campaigns have been carried out in support of such issues. The most famous campaigns are the [Black Lives Matter](#) campaign in the USA and the recent Stop Asian Hate campaign in the USA as well. In light of these movements, it is not only people who should be inclusive but also the technology we use. Some people will recall the [Google algorithm's bias](#) towards black people and the recent gender bias that Facebook's algorithm shows in job adverts. These kinds of biases promote inequality and discrimination and the world does not appreciate this. The technology we use is developed by people and it is up to them to make sure the algorithms that power their technology are not biased.

Algorithms and Algorithm Bias

An algorithm is a set of instructions to perform a specific task. In data science, it is a series of instructions or steps that determine how a program collects, reads, processes, and analyses data to generate output. Algorithm bias refers to systematic and repeatable errors that create unfair outcomes, for example giving advantages to one group of users over others.

The book by Joseph Weizenbaum, called Computer Power and Human Reason, suggests that bias can result from data used in a program or from the way a program is written. As stated in the first paragraph, programs are developed by humans for computers to follow. This means that these programs follow the way their developer thinks and views the world. The developer's views determine how they understand the problem they are trying to solve through the algorithm and therefore determine how they write the algorithm. We can all attest that as humans, we all are biased in how we view the world. The algorithms also incorporate the developer's biases.

Algorithm bias can also result from bias in the data that is used for the algorithm. In machine learning, the aim is to create an algorithm that learns through practice. One way of achieving this is to give an algorithm data from which it learns. Multiple factors can affect the data and result in bias. In a [previous article](#), we explored the issue of data bias with a focus on data bias in [HR analytics](#). We discussed how data is prone to errors during collection, processing, and preparation before being fed into the model. Simple errors such as omitting a variable can

result in huge biases in the algorithm. The data might also be correct, but be skewed towards one group of users. For example, the data might contain more male leaders and this will result in an algorithm that assumes that more leadership roles should be suggested for men than for women.

Examples of Algorithm Bias

1. It is reported that as many as 60 women and ethnic minorities were denied entry to St. George's Hospital Medical School in the US per year from 1982 to 1986, based on the implementation of a new assessment system that denied entry to women and men with foreign-sounding names based on historical trends in admissions.
2. Google Photos' facial recognition algorithm once tagged black people as chimpanzees.
3. Recently, Facebook's job adverts algorithm showed gender bias in job adverts. It is reported that the algorithm would promote more technical jobs to men than women. This was discovered in a study by researchers from the University of Southern California in the USA.
4. A study found that an algorithm that was used to predict which patients would likely need extra medical care favored white patients over black patients. Even though the race was not used as a variable used in this algorithm, healthcare cost history, which is highly correlated to race, was used. For various reasons, black patients had lower healthcare costs than white patients with the same conditions on average.
5. A COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm used in the USA's court systems to predict the likelihood that a defendant would commit another crime predicted twice as many false positives for recidivism for black offenders (45%) than white offenders (23%).
6. In 2015, Amazon realized that their algorithm used for hiring employees was found to be biased against women. The algorithm was biased because it was based on the number of CVs submitted over the past ten years, and since most of the applicants were men, it favored men over women.
7. A study by Joy Buolamwini at the Massachusetts Institute of Technology found that three of the latest gender-recognition AIs, from IBM Microsoft and Chinese company Megvii, could correctly identify a person's gender from a photograph with 99% accuracy, but only for white men. For dark-skinned women, the accuracy was as low as 35%.
8. A 2015 study showed that in a Google images search for "CEO", just 11 percent of the people it displayed were women.

There are various other examples of algorithm bias and the ones mentioned here are some of the famous examples. For HR practitioners, algorithm bias is a big issue when it comes to gender balance. [Recruitment systems](#) and job advertising systems should be as fair as possible and special care should be taken to ensure that the data used in training the systems' algorithms fairly represents all groups.

Tatenda Emma Matika is a Business Analytics Trainee at Industrial Psychology Consultants (Pvt) Ltd a management and human resources consulting firm.

<https://thehumancapitalhub.com/articles/algorithm-bias-everything-you-need-to-know>