

AI can deepen social inequality. Here are 5 ways to help prevent this

By [Tiberio Caetano](#) and [Bill Simpson-Young](#)

22 Jan 2021

From Google searches and dating sites to detecting credit card fraud, artificial intelligence (AI) keeps finding new ways to creep into our lives. But can we trust the algorithms that drive it?



© Mchal Bednarek – [123RF.com](#)

As humans, we make errors. We can have attention lapses and misinterpret information. Yet when we reassess, we can pick out our errors and correct them.

But when an AI system makes an error, it will be repeated again and again no matter how many times it looks at the same data under the same circumstances.

AI systems are trained using data that inevitably reflect the past. If a training data set contains inherent biases from past human decisions, these biases are codified and amplified by the system.

Or if it contains less data about a particular minority group, predictions for that group will tend to be worse. This is called

“algorithmic bias”.

Gradient Institute has co-authored a [paper](#) demonstrating how businesses can identify algorithmic bias in AI systems, and how they can mitigate it.

The work was produced in collaboration with the Australian Human Rights Commission, Consumer Policy Research Centre, CSIRO’s Data61 and the CHOICE advocacy group.

How does algorithmic bias arise?

Algorithmic bias may arise through a lack of suitable training data, or as a result of inappropriate system design or configuration.

For example, a system that helps a bank decide whether or not to grant loans would typically be trained using a large data set of the bank’s previous loan decisions (and other relevant data to which the bank has access).

The system can compare a new loan applicant’s financial history, employment history and demographic information with corresponding information from previous applicants. From this, it tries to predict whether the new applicant will be able to repay the loan.

But this approach can be problematic. One way in which algorithmic bias could arise in this situation is through unconscious biases from loan managers who made past decisions about mortgage applications.

If customers from minority groups were denied loans *unfairly* in the past, the AI will consider these groups’ general repayment ability to be lower than it is.

Young people, people of colour, single women, people with disabilities and blue-collar workers are just some examples of groups that may be disadvantaged.



#BizTrends2021: AI will have a real-world impact in 2021

Dries Cronje 13 Jan 2021



Bias harms both individuals and companies

The biased AI system described above poses two key risks for the bank.

First, the bank could miss out on potential clients, by sending victims of bias to its competitors. It could also be held liable under anti-discrimination laws.

If an AI system continually applies inherent bias in its decisions, it becomes easier for government or consumer groups to identify this systematic pattern. This can lead to hefty fines and penalties.

Mitigating algorithmic bias

Our [paper](#) explores several ways in which algorithmic bias can arise.

It also provides technical guidance on how this bias can be removed, so AI systems produce ethical outcomes which don’t discriminate based on characteristics such as race, age, sex or disability.

For our paper, we ran a simulation of a hypothetical electricity retailer using an AI-powered tool to decide how to offer products to customers and on what terms. The simulation was trained on fictional historical data made up of fictional individuals.

Based on our results, we identify five approaches to correcting algorithmic bias. This toolkit can be applied to businesses across a range of sectors to help ensure AI systems are fair and accurate.

1. Get better data

The risk of algorithmic bias can be reduced by obtaining additional data points or new types of information on individuals, especially those who are underrepresented (minorities) or those who may appear inaccurately in existing data.

2. Pre-process the data

This consists of editing a dataset to mask or remove information about attributes associated with protections under anti-discrimination law, such as race or gender.

3. Increase model complexity

A simpler AI model can be easier to test, monitor and interrogate. But it can also be less accurate and lead to generalisations which favour the majority over minorities.

4. Modify the system

The logic and parameters of an AI system can be proactively adjusted to directly counteract algorithmic bias. For example, this can be done by setting a different decision threshold for a disadvantaged group.

5. Change the prediction target

The specific measure chosen to guide an AI system directly influences how it makes decisions across different groups. Finding a fairer measure to use as the prediction target will help reduce algorithmic bias.



Stumbling blocks to avoid on the AI journey

Harkrishan Singh 21 Oct 2020



Consider legality and morality

In our recommendations to government and businesses wanting to employ AI decision-making, we foremost stress the importance of considering general principles of fairness and human rights when using such technology. And this must be done *before* a system is in-use.

We also recommend systems are rigorously designed and tested to ensure outputs aren't tainted by algorithmic bias. Once operational, they should be closely monitored.

Finally, we advise that to use AI systems responsibly and ethically extends beyond compliance with the narrow letter of the law. It also requires the system to be aligned with broadly-accepted social norms — and considerate of impact on individuals, communities and the environment.

With AI decision-making tools becoming commonplace, we now have an opportunity to not only increase productivity, but

create a more equitable and just society – that is, if we use them carefully.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

ABOUT THE AUTHOR

Tiberio Caetano, chief scientist, Gradient Institute, University of Sydney. Bill Simpson-Young, chief executive, Gradient Institute, University of Sydney.

For more, visit: <https://www.bizcommunity.com>