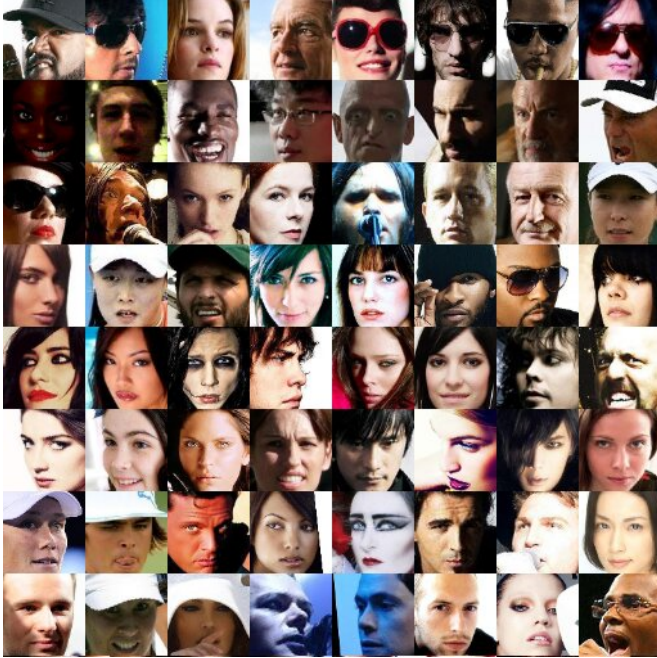


# An AI that 'de-biases' algorithms

29 January 2019



Credit: Massachusetts Institute of Technology

We've learned in recent years that AI systems can be unfair, which is dangerous, as they're increasingly being used to do everything from predicting crime to determining what news we consume. Last year's study showing the racism of face-recognition algorithms demonstrated a fundamental truth about AI: if you train with biased data, you'll get biased results.

A team from MIT CSAIL is working on a solution, with an [algorithm](#) that can automatically "de-bias" data by resampling it to be more balanced.

The algorithm can learn both a specific task like face detection, as well as the underlying structure of the training data, which allows it to identify and minimize any hidden biases. In tests the algorithm decreased "categorical bias" by over 60 percent compared to state-of-the-art facial detection models—while simultaneously maintaining the overall precision of these systems. The team

evaluated the algorithm on the same [facial-image dataset](#) that was developed last year by researchers from the MIT Media Lab.

A lot of existing approaches in this field require at least some level of human input into the system to define specific biases that researchers want it to learn. In contrast, the MIT team's algorithm can look at a [dataset](#), learn what is intrinsically hidden inside it, and automatically resample it to be more fair without needing a programmer in the loop.

"Facial classification in particular is a technology that's often seen as 'solved,' even as it's become clear that the datasets being used often aren't properly vetted," says Ph.D. student Alexander Amini, who was co-lead author on a related paper that was presented this week at the Conference on Artificial Intelligence, Ethics and Society (AIES). "Rectifying these issues is especially important as we start to see these kinds of algorithms being used in security, [law enforcement](#) and other domains."

Amini says that the team's system would be particularly relevant for larger datasets that are too big to vet manually and also extends to other computer vision applications beyond facial detection.

**More information:** [Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure. www.aies-conference.com/wp-con... IES-19\\_paper\\_220.pdf](#)

Provided by Massachusetts Institute of Technology

APA citation: An AI that 'de-biases' algorithms (2019, January 29) retrieved 20 September 2019 from <https://techxplore.com/news/2019-01-ai-de-biases-algorithms.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*