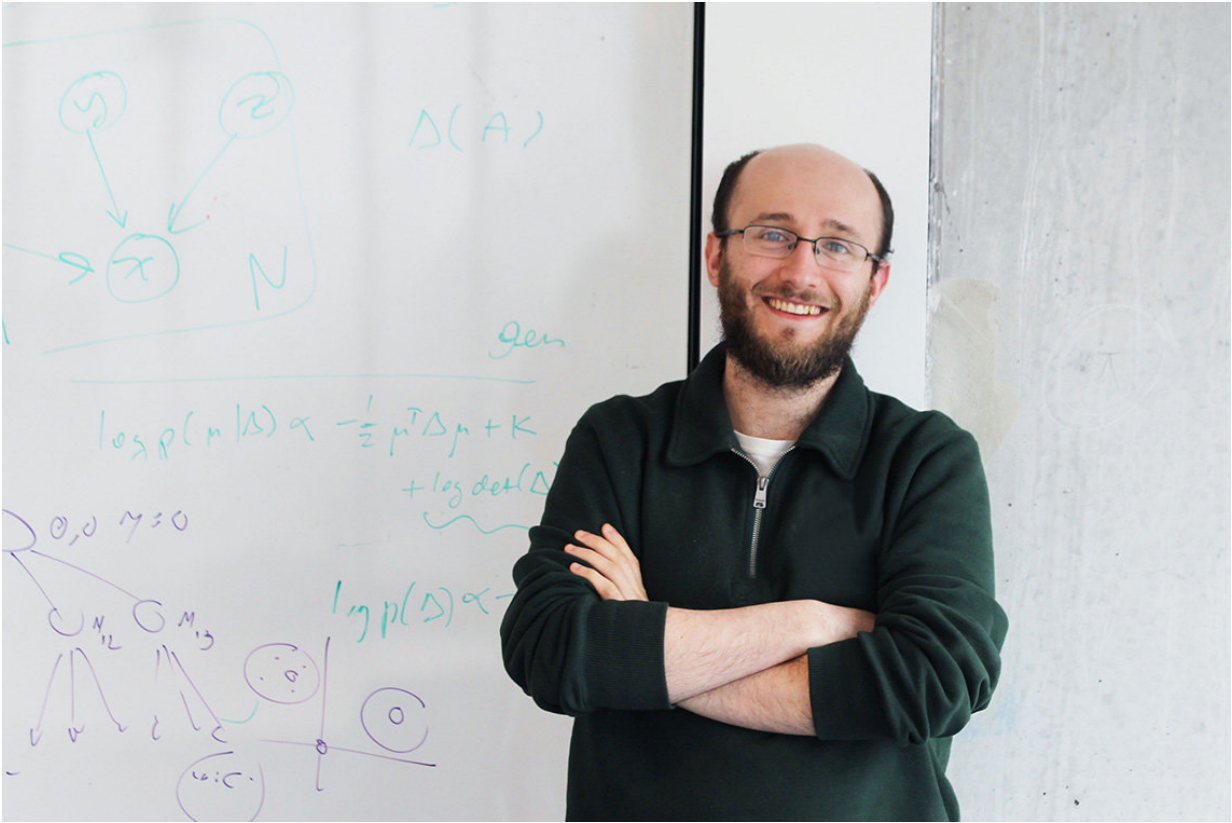


Fairness needed in algorithmic decision-making, experts say

May 2 2018, by Nina Haikara



David Madras, a PhD student in the department of computer science, says there are important questions to be asked about the fairness of decisions made by automated systems . Credit: Nina Haikara

University of Toronto Ph.D. student David Madras says many of today's

algorithms are good at making accurate predictions, but don't know how to handle uncertainty well. If a badly calibrated algorithm makes the wrong decision, it's usually very wrong.

"A human user might understand a situation better than a computer, whether it's information not available because it's very qualitative in nature, or something happening in the real world that didn't get inputted into the algorithm," says Madras, a [machine learning](#) researcher in the department of computer science who is also affiliated with the Vector Institute for Artificial Intelligence.

"Both can be very important and can have an effect on what predictions should be [made]."

Madras is presenting his research, "Predict Responsibly: Increasing Fairness by Learning to Defer," at the International Conference on Learning Representations (ICLR), in Vancouver this week. The conference is focused on the methods and performance of machine learning and brings together leaders in the field.

Madras says he and Toniann Pitassi, a professor in U of T's departments of computer science and mathematics and a computational theory expert who also explores computational fairness, as well as Richard Zemel, a U of T professor of computer science and research director of the Vector Institute, have developed their model with fairness in mind. Where there is a degree of uncertainty, an algorithm must have the option to answer, "I don't know" and defer its decision to a human user.

Madras explains if Facebook were to use an algorithm to automatically tag people in images, it's maybe not so important if the tagging is done wrong. But when individual outcomes are high-impact, the risk can be greater. He says the model hasn't been applied yet to any specific application, but rather the researchers are thinking of the types of ways it

could be used in [real world](#) cases.

"In medical settings, it can be important to output something that can be interpreted – there's some amount of uncertainty surrounding its prediction – and a doctor should decide if treatment should be given."

Madras' graduate supervisor Zemel, who will take up an NSERC Industrial Research Chair in Machine Learning this summer, is also examining how machine learning can be made to be more expressive, controllable and fair.

Zemel says machine learning based on historical data, such as whether a bank loan was approved or the length of prison sentences, will naturally pick up on biases. And the biases in the data set can play out in a machine's predictions, he says.

"In this paper, we think a lot about an external decision-maker. In order to train up our model, we have to use historical decisions that are made by decision-makers. The outcomes of those decisions, created by existing decision-makers, can be themselves biased or in a sense incomplete."

Madras believes the increased focus on algorithmic fairness alongside issues of privacy, security and safety, will help make machine learning more applicable to high-stakes applications.

"It's raising important questions about the role of an automated system that's making important decisions, and how to get [them] to predict in ways that we want them to."

Madras says he's continuing to think about issues of fairness and related areas, like causality: Two things can be correlated – because they occur together frequently – but that doesn't mean that one causes the other.

"If an [algorithm](#) is deciding when to give someone a loan, it might learn that people who live in a certain postal code are less likely to pay back loans. But that can be a source of unfairness. It's not like living in a certain postal code makes you less likely to pay back a loan," he says.

"It's an interesting and important set of problems to be working on."

More information: Predict Responsibly: Increasing Fairness by Learning to Defer: openreview.net/forum?id=HyjVKLJwz

Provided by University of Toronto

Citation: Fairness needed in algorithmic decision-making, experts say (2018, May 2) retrieved 23 April 2024 from <https://techxplore.com/news/2018-05-fairness-algorithmic-decision-making-experts.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.