

New study offers a better way to make AI fairer for everyone

June 5 2024, by Carnegie Mellon UniversitySheila Davis



Credit: Unsplash/CC0 Public Domain

In a new [paper](#), researchers from Carnegie Mellon University and Stevens Institute of Technology show a new way of thinking about the fair impacts of AI decisions.

They draw on a well-established tradition known as social welfare optimization, which aims to make decisions fairer by focusing on the overall benefits and harms to individuals. This method can be used to evaluate the industry standard assessment tools for AI fairness, which look at approval rates across protected groups.

"In assessing fairness, the AI community tries to ensure equitable treatment for groups that differ in economic level, race, [ethnic background](#), gender, and other categories," explained John Hooker, professor of operations research at the Tepper School of Business at Carnegie Mellon, who coauthored the study and presented the paper at the [International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research](#) (CPAIOR) on May 29 in Uppsala, Sweden. The paper received the Best Paper Award.

Imagine a situation where an AI system decides who gets approved for a mortgage or who gets a job interview. Traditional fairness methods might only ensure that the same percentage of people from different groups get approved.

But what if being denied a mortgage has a much bigger negative impact on someone from a disadvantaged group than on someone from an advantaged group? By employing a social welfare optimization method, AI systems can make decisions that lead to better outcomes for everyone, especially for those in disadvantaged groups.

The study focuses on "alpha fairness," a method of finding a balance between being fair and getting the most benefit for everyone. Alpha fairness can be adjusted to balance fairness and efficiency more or less, depending on the situation.

Hooker and his co-authors show how social welfare optimization can be

used to compare different assessments for group fairness currently used in AI. By using this method, we can understand the benefits of applying different group fairness tools in different contexts. It also ties these group fairness assessment tools to the larger world of fairness-efficiency standards used in economics and engineering.

Derek Leben, associate teaching professor of business ethics at the Tepper School, and Violet Chen, assistant professor at Stevens Institute of Technology, who received her Ph.D. from the Tepper School, coauthored the study.

"Our findings suggest that social welfare optimization can shed light on the intensely discussed question of how to achieve group fairness in AI," Leben said.

The study is important for both AI system developers and policymakers. Developers can create more equitable and effective AI models by adopting a broader approach to fairness and understanding the limitations of fairness measures. It also highlights the importance of considering [social justice](#) in AI development, ensuring that technology promotes equity across diverse groups in society.

The paper is published in CPAIOR 2024 Proceedings.

More information: Violet Chen et al, Assessing Group Fairness with Social Welfare Optimization, *Integration of Constraint Programming, Artificial Intelligence, and Operations Research* (2024). [DOI: 10.1007/978-3-031-60597-0_14](#)

Provided by Tepper School of Business, Carnegie Mellon University

Citation: New study offers a better way to make AI fairer for everyone (2024, June 5) retrieved 29 June 2024 from <https://techxplore.com/news/2024-06-ai-fairer.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.