

Science News

from research organizations

Using artificial intelligence to detect discrimination

Date: July 10, 2019

Source: Penn State

Summary: A new artificial intelligence (AI) tool for detecting unfair discrimination -- such as on the basis of race or gender -- has been created by researchers.

Share:     

FULL STORY

A new artificial intelligence (AI) tool for detecting unfair discrimination -- such as on the basis of race or gender -- has been created by researchers at Penn State and Columbia University.

Preventing unfair treatment of individuals on the basis of race, gender or ethnicity, for example, been a long-standing concern of civilized societies. However, detecting such discrimination resulting from decisions, whether by human decision makers or automated AI systems, can be extremely challenging. This challenge is further exacerbated by the wide adoption of AI systems to automate decisions in many domains -- including policing, consumer finance, higher education and business.

"Artificial intelligence systems -- such as those involved in selecting candidates for a job or for admission to a university -- are trained on large amounts of data," said Vasant Honavar, Professor and Edward Frymoyer Chair of Information Sciences and Technology, Penn State. "But if these data are biased, they can affect the recommendations of AI systems."

For example, he said, if a company historically has never hired a woman for a particular type of job, then an AI system trained on this historical data will not recommend a woman for a new job.

"There's nothing wrong with the machine learning algorithm itself," said Honavar. "It's doing what it's supposed to do, which is to identify good job candidates based on certain desirable characteristics. But since it was trained on historical, biased data it has the potential to make unfair recommendations."

The team created an AI tool for detecting discrimination with respect to a protected attribute, such as race or gender, by human decision makers or AI systems that is based on the concept of causality in which one thing -- a cause -- causes another thing -- an effect.

"For example, the question, 'Is there gender-based discrimination in salaries?' can be reframed as, 'Does gender have a causal effect on salary?', or in other words, 'Would a woman be paid more if she was a man?' said Aria Khademi, graduate student in information sciences and technology, Penn State.

Since it is not possible to directly know the answer to such a hypothetical question, the team's tool uses sophisticated counterfactual inference algorithms to arrive at a best guess.

"For instance," said Khademi, "one intuitive way of arriving at a best guess as to what a fair salary would be for a female employee is to find a male employee who is similar to the woman with respect to qualifications, productivity and experience. We can minimize gender-based discrimination in salary if we ensure that similar men and women receive similar salaries."

The researchers tested their method using various types of available data, such as income data from the U.S. Census Bureau to determine whether there is gender-based discrimination in salaries. They also tested their method using the New York City Police Department's stop-and-frisk program data to determine whether there is discrimination against people of color in arrests made after stops. The results appeared in May in Proceedings of The Web Conference 2019.

"We analyzed an adult income data set containing salary, demographic and employment-related information for close to 50,000 individuals," said Honavar. "We found evidence of gender-based discrimination in salary. Specifically, we found that the odds of a woman having a salary greater than \$50,000 per year is only one-third that for a man. This would suggest that employers should look for and correct, when appropriate, gender bias in salaries."

Although the team's analysis of the New York stop-and-frisk dataset -- which contains demographic and other information about drivers stopped by the New York City police force -- revealed evidence of possible racial bias against Hispanics and African American individuals, it found no evidence of discrimination against them on average as a group.

"You cannot correct for a problem if you don't know that the problem exists," said Honavar. "To avoid discrimination on the basis of race, gender or other attributes you need effective tools for detecting discrimination. Our tool can help with that."

Honavar added that as data-driven artificial intelligence systems increasingly determine how businesses target advertisements to consumers, how police departments monitor individuals or groups for criminal activity, how banks decide who gets a loan, who employers decide to hire, and how colleges and universities decide who gets admitted or receives financial aid, there is an urgent need for tools such as the one he and his colleagues developed.

"Our tool," he said, "can help ensure that such systems do not become instruments of discrimination, barriers to equality, threats to social justice and sources of unfairness."

The National Institutes of Health and National Science Foundation supported this research.

MAKE A DIFFERENCE: SPONSORED OPPORTUNITY



Support an Inclusive America

How many businesses in the U.S. are started by immigrants?

Almost none

10%

25%

[Sponsored by Postmates](#)

Story Source:

Materials provided by **Penn State**. *Note: Content may be edited for style and length.*

Cite This Page:

MLA

APA

Chicago

Penn State. "Using artificial intelligence to detect discrimination." ScienceDaily. ScienceDaily, 10 July 2019. <www.sciencedaily.com/releases/2019/07/190710121649.htm>.

RELATED STORIES

Artificial Intelligence Could Use EKG Data to Measure Patient's Overall Health Status

Aug. 27, 2019 — Researchers applying artificial intelligence to electrocardiogram data estimated the age group of a patient and predicted their gender. Artificial intelligence could more accurately track overall ... **read more »**

Artificial Intelligence Helps Banana Growers Protect the World's Most Favorite Fruit

Aug. 12, 2019 — Using artificial intelligence, scientists created an easy-to-use tool to detect banana diseases and pests. With an average 90 percent success rate in detecting a pest or a disease, the tool can help ... **read more »**

Artificial Intelligence Could Be 'Game Changer' in Detecting, Managing Alzheimer's Disease

June 25, 2019 — Could artificial intelligence be the solution for detecting and managing a complex condition like Alzheimer's disease? A team of researchers thinks so. They employed a novel application of ... **read more** »

AI Tool Automatically Reveals How to Write Apps That Drain Less Battery

Oct. 9, 2018 — Researchers have created a new tool, called 'DiffProf,' that uses artificial intelligence to automatically decide for the developer if a feature should be improved to drain less battery and ... **read more** »

FROM AROUND THE WEB

Below are relevant articles that may interest you. ScienceDaily shares links with scholarly publications in the TrendMD network and earns revenue from third-party advertisers, where indicated.

Helomics Inks Separate Pacts With UPMC, Interpace Dx for AI-Driven Oncology Decision Making 

Precision Oncology News, 2019

Japanese Foundation, Fronteo Developing AI-Based System for Cancer Dx 

staff reporter, 360Dx, 2017

Inspirata, DeePathology Partner on AI-Based H. Pylori Pathology Testing 

staff reporter, 360Dx, 2019

Hospitals, Vendors Try to Figure Out AI's Role in Precision Oncology 

Precision Oncology News, 2019

Anticipating Approvals in China, VoxelCloud Leveraging AI Technology for Diagnostics 

Tony Fong, 360Dx, 2017

Fujitsu Laboratories, Aichi Cancer Center to Develop AI-Based Tech for Cancer Genome Data Analysis



staff reporter, 360Dx, 2019

Ibex Medical Gets CE-IVD Mark for AI Prostate Cancer Detection System 

staff reporter, 360Dx, 2020

European Committee Recommends Approval of Tecentriq-Chemo Combination in Triple Negative Breast Cancer 

Precision Oncology News, 2019

Powered by **TREND MD**



Free Subscriptions

Get the latest science news with ScienceDaily's free email newsletters, updated daily and weekly. Or view hourly updated newsfeeds in your RSS reader:

 **Email Newsletters**

 [RSS Feeds](#)

Follow Us

Keep up to date with the latest news from ScienceDaily via social networks:

 [Facebook](#)

 [Twitter](#)

 [LinkedIn](#)

Have Feedback?

Tell us what you think of ScienceDaily -- we welcome both positive and negative comments. Have any problems using the site? Questions?

 [Leave Feedback](#)

 [Contact Us](#)

[About This Site](#) | [Staff](#) | [Reviews](#) | [Contribute](#) | [Advertise](#) | [Privacy Policy](#) | [Editorial Policy](#) | [Terms of Use](#)

Copyright 2020 ScienceDaily or by other parties, where indicated. All rights controlled by their respective owners.

Content on this website is for information only. It is not intended to provide medical or other professional advice.

Views expressed here do not necessarily reflect those of ScienceDaily, its staff, its contributors, or its partners.

Financial support for ScienceDaily comes from advertisements and referral programs, where indicated.