Meaningful standards for auditing high-stakes artificial intelligence

March 14 2022

Credit: Pixabay/CC0 Public Domain

When hiring, many organizations use artificial intelligence tools to scan resumes and predict job-relevant skills. Colleges and universities use AI to automatically score essays, process transcripts and review

extracurricular activities to predetermine who is likely to be a "good student." With so many unique use-cases, it is important to ask: can AI tools ever be truly unbiased decision-makers? In response to claims of unfairness and bias in tools used in hiring, college admissions, predictive policing, health interventions, and more, the University of Minnesota recently developed a new set of auditing guidelines for AI tools.

The auditing guidelines, published in the *American Psychologist*, were developed by Richard Landers, associate professor of psychology at the University of Minnesota, and Tara Behrend from Purdue University. They apply a century's worth of research and professional standards for measuring personal characteristics by psychology and education researchers to ensure the fairness of AI.

The researchers developed guidelines for AI auditing by first considering the ideas of fairness and bias through three major lenses of focus:

- How individuals decide if a decision was fair and unbiased
- How societal, legal, ethical and moral standards present fairness and bias
- How individual technical domains—like computer science, statistics and psychology—define fairness and bias internally

Using these lenses, the researchers presented psychological audits as a standardized approach for evaluating the fairness and bias of AI systems that make predictions about humans across high-stakes application areas, such as hiring and college admissions.

There are twelve components to the auditing framework across three categories that include:

- Components related to the creation of, processing done by, and predictions created by the AI

- Components related to how the AI is used, who its decisions affect and why
- Components related to overarching challenges: the cultural context in which the AI is used, respect for the people affected by it, and the scientific integrity of the research used by AI purveyors to support their claims

"The use of AI, especially in hiring, is a decades-old practice, but recent advances in AI sophistication have created a bit of a 'wild west' feel for AI developers," said Landers. "There are a ton of startups now that are unfamiliar with existing ethical and legal standards for hiring people using algorithms, and they are sometimes harming people due to ignorance of established practices. We developed this framework to help inform both those companies and related regulatory authorities."

The researchers recommend the standards they developed to be followed both by internal auditors during the development of high-stakes predictive AI technologies, and afterward by independent external auditors. Any system that claims to make meaningful recommendations about how people should be treated should be evaluated within this framework.

"Industrial psychologists have unique expertise in the evaluation of high-stakes assessments," said Behrend. "Our goal was to educate the developers and users of AI-based assessments about existing requirements for fairness and effectiveness, and to guide the development of future policy that will protect workers and applicants."

AI models are developing so rapidly, it can be difficult to keep up with the most appropriate way to audit a particular kind of AI system. The researchers hope to develop more precise standards for specific use cases, partner with other organizations globally interested in establishing auditing as a default approach in these situations, and work toward a

better future with AI more broadly.

  More information: Richard N. Landers et al, Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models., *American Psychologist* (2022). [DOI: 10.1037/amp0000972](#)

Provided by University of Minnesota