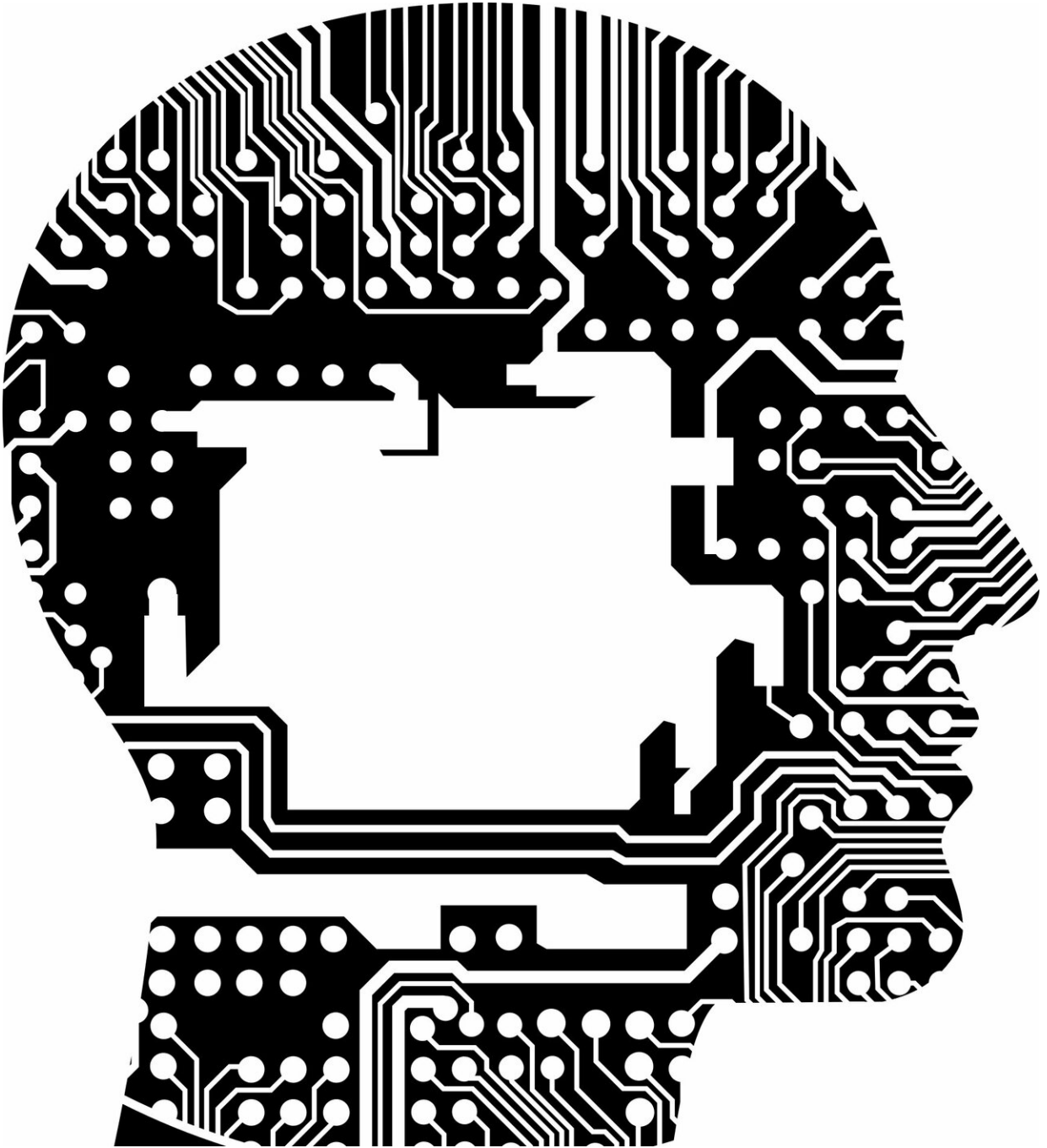


Want better AI? Get input from a real (human) expert

November 20 2023



Credit: CC0 Public Domain

Can AI be trusted? The question pops up wherever AI is used or

discussed—which, these days, is everywhere.

It's a question that even some [AI](#) systems ask themselves.

Many machine-learning systems create what experts call a "confidence score," a value that reflects how confident the system is in its decisions. A low score tells the human user that there is some uncertainty about the recommendation; a high score indicates to the human user that the system is at least sure of its decisions. Savvy humans know to check the confidence score when deciding whether to trust the recommendation of a machine-learning system.

Scientists at the Department of Energy's Pacific Northwest National Laboratory have put forth a new way to evaluate an AI system's recommendations. They bring human experts into the loop to view how the ML performed on a set of data. The expert learns which types of data the machine-learning system typically classifies correctly and which data types lead to confusion and system errors. Armed with this knowledge, the experts then offer their own confidence score on future system recommendations.

The result of having a human look over the shoulder of the AI system? Humans predicted the AI system's performance more accurately.

Minimal human effort—just a few hours—evaluating some of the decisions made by the AI program allowed researchers to vastly improve on the AI program's ability to assess its decisions. In some analyses by the team, the accuracy of the confidence score doubled when a human provided the score.

The PNNL team presented its results at a recent meeting of the [Human Factors and Ergonomics Society](#) in Washington, D.C., part of a session on human-AI robot teaming.

"If you didn't develop the [machine-learning algorithm](#) in the first place, then it can seem like a black box," said Corey Fallon, the lead author of the study and an expert in human-machine interaction. "In some cases, the decisions seem fine. In other cases, you might get a recommendation that is a real head-scratcher. You may not understand why it's making the decisions it is."

The grid and AI

It's a dilemma that power engineers working with the electric grid face. Their decisions based on reams of data that change every instant keep the lights on and the nation running. However, power engineers may be reluctant to turn over decision-making authority to machine-learning systems.

"There are hundreds of [research papers](#) about the use of machine learning in power systems, but almost none of them are applied in the real world. Many operators simply don't trust ML. They have domain experience—something that ML can't learn," said co-author Tianzhixi "Tim" Yin.

The researchers at PNNL, which has a world-class team modernizing the grid, took a closer look at one machine-learning algorithm applied to power systems. They trained the SVM (support-vector machine) algorithm on real data from the grid's Eastern Interconnection in the U.S. The program looked at 124 events, deciding whether a generator was malfunctioning or whether the data was showing other types of events that were less noteworthy.

The algorithm was 85% reliable in its decisions. Many of its errors occurred when there were complex power bumps or frequency shifts. Confidence scores created with a human in the loop were a marked improvement over the system's assessment of its own decisions. The

human expert's input predicted the algorithm's decisions with much greater accuracy.

More human, better machine learning

Fallon and Yin call the new score an "Expert-Derived Confidence" score, or EDC score.

They found that, on average, when humans weighed in on the data, their EDC scores predicted model behavior that the algorithm's confidence scores couldn't predict.

"The human expert fills in gaps in the ML's knowledge," said Yin. "The human provides information that the ML did not have, and we show that that information is significant. The bottom line is that we've shown that if you add human expertise to the ML results, you get much better confidence."

"This is the type of research needed to prepare and equip an AI-ready workforce," said Fallon. "If people don't trust the tool, then you've wasted your time and money. You've got to know what will happen when you take a machine-learning model out of the laboratory and put it to work in the real world."

"I'm a big fan of human expertise and of human-machine teaming. Our EDC scores allow the human to better assess the situation and make the ultimate decision."

Provided by Pacific Northwest National Laboratory

Citation: Want better AI? Get input from a real (human) expert (2023, November 20) retrieved 12 May 2024 from <https://techxplore.com/news/2023-11-ai-real-human-expert.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.