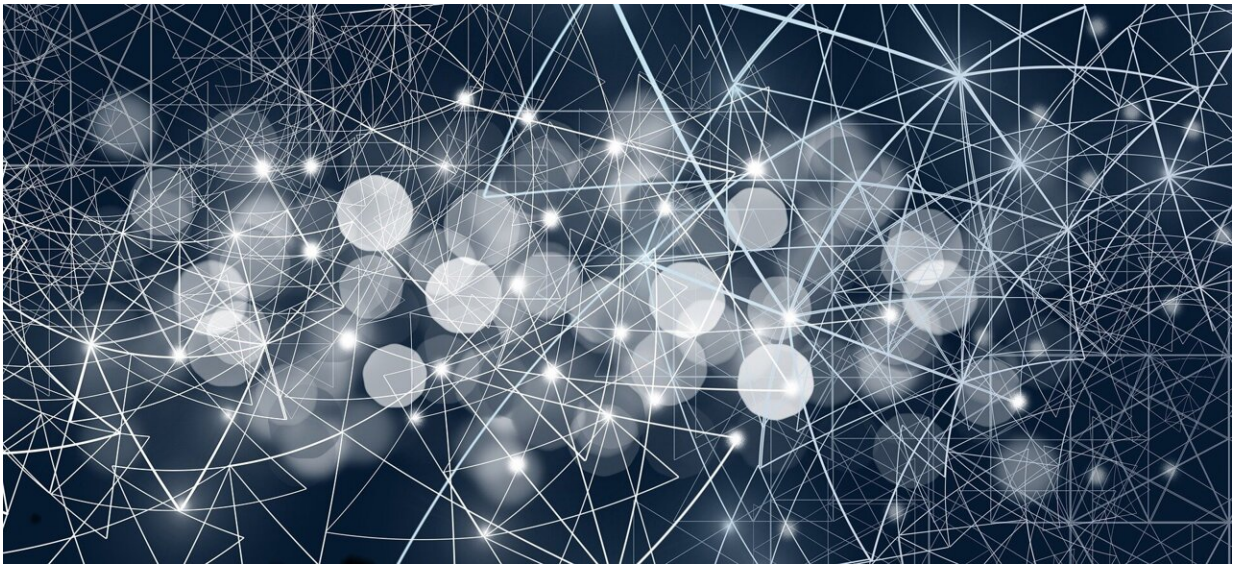# Scientists tap AI betting agents to help solve research reproducibility concerns

March 17 2022

Credit: Pixabay/CC0 Public Domain

Scientists are increasingly concerned that the lack of reproducibility in research may lead to, among other things, inaccuracies that slow scientific output and diminished public trust in science. Now, a team of researchers reports that creating a prediction market, where artificially intelligent—AI—agents make predictions—or bet—on hypothetical replication studies, could lead to an explainable, scalable approach to estimate confidence in published scholarly work.

Replication of experiments and studies, a critical step in the scientific process, helps provide confidence in the results and indicates whether they can generalize across contexts, according to Sarah Rajtmajer, assistant professor in information sciences and technology, Penn State. As experiments have become more complex, costly and time consuming, scientists increasingly lack the resources for robust replication efforts—what is often referred to now by them as the "replication crisis."

"As scientists, we want to do work, and we want to know that our work is good," said Rajtmajer. "Our approach to help address the replication crisis is to use AI to help predict whether a finding would replicate if repeated and why."

Crowdsourced prediction markets can be likened to betting parlors to help forecast real world events, rather than horse races or football game outcomes. These markets have already been used to help anticipate everything from elections to infectious disease spreads.

"What inspired us was the success of prediction markets in precisely this task—that is, when you place researchers in a market and give them some cash to bet on outcomes of replications, they're pretty good at it," said Rajtmajer, who is a research associate at the Rock Ethics Institute and an Institute for Computational and Data Sciences associate. "But human-run prediction markets are expensive and slow. And ideally, you should run replications in parallel to the market so there is some ground truth on which researchers are betting. It just doesn't scale."

A bot-based approach scales and offers some explanation of its findings based on trading patterns and the features of the papers and claims that influenced the bots' behavior. In the team's approach, bots are trained to recognize key features in academic research papers—such as the authors and institutions, statistics, and linguistic cues, downstream mentions, and

similar studies in the literature—and then make assessments regarding the confidence that the study is robust enough to replicate in future studies. Just like a human betting on the outcomes of a sporting event, the bot then bids on its level of confidence. The AI-powered bots' results are compared to the bets made in human predictions.

C. Lee Giles, the David Reese Professor at the College of Information Sciences and Technology, said that while prediction markets based on human participants are well-known and have been used successfully in a number of fields, prediction markets are novel in examining research results.

"That's probably the interesting and unique thing we're doing here," said Giles, who is also an ICDS associate. "We have already seen that humans are pretty good at using prediction markets. But, here, we're using bots for our market, which is a little unusual and sort of fun."

According to the researchers, who presented their results at a recent meeting for the Association for the Advancement of Artificial Intelligence, the system provided confidence scores for about 68 of the 192 papers—or about 35%—of the papers that were eventually reproduced, or ground truth replication studies. On that set of papers, the accuracy was approximately 90%.

Because humans tend to perform better at predicting research reproducibility, but bots can perform at scale, Giles and Rajtmajer suggest that a hybrid approach—human and bots working together—may deliver the best of both worlds: a system that would feature higher accuracy but still be scalable.

"Maybe we can train the bots in the presence of human traders every so often, and then deploy them offline when we need a quick result, or when we need replication efforts at scale," said Rajtmajer. "Moreover,

we can create bot markets that also leverage that intangible human wisdom. That is something we are working on right now."

More information: Results: [www.cs.odu.edu/~jwu/downloads/ … tmajer-2022-aaai.pdf](www.cs.odu.edu/~jwu/downloads/)

Provided by Pennsylvania State University