

Can AI Get Smart Enough to Know Fake News When It Sees it?

MIT researchers launch new AI program to rate the accuracy of news postings.

[Kevin McCaney](#)

Wed, 11/21/2018 - 09:27



Photo Credit: Sean Pavone/iStock

Artificial intelligence has contributed plenty to the flood of fake news spreading around the globe, with AI bots generating, amplifying and distributing false or misleading stories to targeted audiences. And it's likely to help [make things worse](#) in the years ahead, manipulating text, images, videos and voices to create a worldwide house of mirrors in which it could be hard to believe your own eyes and ears.

So it only seems fair that the power of AI be employed to try to correct the problem. Research teams in academia and the government have launched a number of projects aimed at using AI to combat the spread of false information, by identifying phony content, studying the human element of why people buy into fake news, and using AI to help human fact-checkers speed up their time-consuming work. The Massachusetts Institute of Technology, working with the Qatar Computing Research Institute (QCRI), have come up with another approach: [rating the sources producing it](#).

Researchers at MIT's Computer Science and Artificial Intelligence Lab (CSAIL) and QCRI developed a machine learning system that examines a range of sources — major outlets like CNN or Fox to low-traffic content suppliers — and rates them on factors such as language, sentence structure, complexity, and emphasis on features like fairness or loyalty. The project used data from Media Bias/Fact Check (MBFC), whose human fact-checkers rate the accuracy of about 2,000 large and small sites, and so far, has built an open-source database of more than 1,000 sources with ratings for accuracy and bias.

Ballpark Estimates

But while the project attacks a new angle in the fake news fight and moves the ball forward, it also shows what this and other projects are up against, with or without AI in their arsenals. The best model the researchers tested proved to be about 65 percent accurate in gauging a site's factuality and about 70 percent accurate in determining if it leans right, left or moderate. Those results are promising, but they also create a sizable margin of error. Can you trust the accuracy rating of a website if the rating itself has a 35 percent chance of being wrong?

Researchers, who presented their results in October at the 2018 Empirical Methods in Natural Language Processing conference in Brussels, were thorough in building their ratings. Among the best clues to bias, for instance, were linguistic traits such as sentiment, use of hyperbolic language and sentence structures. A focus on fairness or reciprocity tended to indicate a left-leaning outlet, while a focus on sanctity or authority was indicative of a rightward position. With regard to accuracy, the team also found that sites such as Wikipedia pages were useful (more credible sources had longer entries) as were the structure of a site's URL (sites with a lot of special characters or complicated subdirectories tended to be less reliable).

The team tested more than 900 variations for predicting a site's trustworthiness and trained a machine learning model to try various combinations of those variations to see which produced the most on-the-mark ratings. That 1,000-plus-source database, now available to fact-checkers, is the largest of its kind in the world, the researchers say. The team also plans to adapt the system for languages other than English and expand its ratings beyond merely right or left-leaning to more specific socio-political topics.

Consider the Source

And while the results are far from perfect at this point, the researchers say rating the sources could be the best way to help the fact-checking process, because of the difficulty in checking every fact published across the internet in enough time to make a difference. "If a website has published fake news before, there's a good chance they'll do it again," Ramy Baly, a post-doctoral MIT researcher and lead author of the team's paper, said in MIT News.

"By automatically scraping data about these sites, the hope is that our system can help figure out which ones are likely to do it in the first place."

A key to accurate ratings is starting with reliable training data, what researchers call "ground truth," which in this case came from MBFC. "Since it is much easier to obtain ground truth on sources [than on articles], this method is able to provide direct and accurate predictions regarding the type of content distributed by these sources," Sibel Adali, a professor of computer science at Rensselaer Polytechnic Institute who was not involved in the project, told MIT News.

Efforts to use AI to combat fake news are making progress, as the MIT project and other efforts have shown. If nothing else, they can be an invaluable aid to human fact checkers at Politifact, Snopes and other organizations doing the hard work of verifying the claims being made. But with fake news so widespread — and aided by AI — countering its impact is still an uphill battle.

[View PDF](#)

[AI](#)

[MIT](#)

[bots](#)

[natural language processing](#)

[machine learning](#)