

Tech Talk | Robotics | Artificial Intelligence

15 Mar 2021 | 16:00 GMT

Can AI Stop People From Believing Fake News?

Machine learning algorithms provide a way to detect misinformation based on writing style and how articles are shared

By Brian T. Horowitz



Illustration: iStockphoto/IEEE Spectrum

Machine learning algorithms provide a way to detect misinformation based on writing style and how articles are shared.

On topics as varied as climate change and the safety of vaccines, you will find a wave of misinformation all over social media. Trust in conventional news sources may seem lower than ever, but researchers are working on ways to give people more insight on whether they can believe what they read. Researchers have been testing artificial intelligence (AI) tools that could help filter legitimate news. But how trustworthy is AI when it comes to stopping the spread of misinformation?

Researchers at the [Rensselaer Polytechnic Institute \(RPI\)](#) and the University of Tennessee collaborated to study the role of AI in helping people identify whether the news they're reading is legitimate or not.

The [research paper](#), “Tailoring Heuristics and Timing AI Interventions for Supporting News Veracity Assessments,” was published in *Computers in Human Behavior Reports*. It discussed how crowdsourcing marketplace [Amazon Mechanical Turk \(AMT\)](#) can be used to identify misinformation for fresh news and specific heuristics, which are rules of thumb used to process information and consider its veracity. In other words, heuristics are essentially “shortcuts for decisions,” explained [Dorit Nevo](#), an associate professor at RPI’s Lally School of Management and a lead author for the paper.

The study found that AI would be successful in flagging false stories only if the reader did not already have an opinion on the topic, Nevo said. When study subjects were set in their beliefs, confirmation bias kept them from reassessing their views.

Nevo said the first part of the project focused on whether subjects could detect misinformation around climate change and vaccines like the one designed to prevent chicken pox. Then, beginning in April 2020, her team studied how people responded to news related to COVID-19.

“With COVID-19, there was a significant difference,” Nevo said. They found that about 72 percent of respondents could identify misinformation about the coronavirus without heuristic clues, and roughly 93 percent were able to be convinced by the researcher’s heuristics that the content was fake.

Examples of heuristic clues include text with too many capital letters or the use of strong language, Nevo said.

There were two types of heuristics mentioned in the team’s paper: objective heuristics and source heuristics. They put a statement at the top of each article the subjects read; it instructed them to read the article and indicate whether they believed its central thesis.

“We either put a statement that says the AI finds this article reliable and accurate based on the objective heuristics, or we said the AI finds the source reliable,” Nevo said. “So that’s the source heuristic.”

In her research on heuristics, Nevo found that people’s thinking takes one of two paths: The first path is to read the article, think about it and decide if they believe it; the second is to consider the source and what others think about the news, and decide whether to believe it before reading it.

Image: Dorit Nevo/RPI/IEEE Spectrum

Researchers at RPI researched the role of heuristics and AI in detecting whether people thought news was credible

Another research paper, “[Timing Matters When Correcting Fake News](#),” published in the *Proceedings of the National Academy of Science* by researchers at Harvard University, differed from the RPI researchers in its findings. While Nevo and her collaborators found that it’s easier to convince people that a story is fake news before reading it, the Harvard researchers, led by Nadia M. Brashier, a psychologist and neuroscientist, discovered that a fact-check can convince people of misinformation even after reading headlines. When study subjects read true or false labels after reading a headline, that resulted in a 25.3 percent reduction in “subsequent misclassification,” when compared to headlines with no tag, Brashier and her team found.

In the end, fighting misinformation will require both computing and human efforts such as policy changes, says Benjamin D. Horne, an assistant professor of Information Sciences at the University of Tennessee and one of Nevo’s co-authors. He says the RPI-Tennessee work was inspired by AI tools he designed previously. Horne was previously a research assistant at RPI, where he developed machine learning (ML) algorithms that can detect partial truths as well as decontextualized truths and out-of-date information.

“Our algorithms are trained on source-level behavior, both when using the textual content of an article and the network of other news sources that it draws news from,” Horne said. “We have found that these two types of features together are quite good at distinguishing between sources labeled as reliable or unreliable by external news source ratings.”

The machine learning algorithms analyze the writing style and the content-sharing behavior of news outlets, Horne said. Researchers trained a supervised ML algorithm called Random Forest, a classification algorithm that uses decision trees.

AI for Detecting Fake News

Everyday News		
Source	Title	Ground Truth
Natural News	Climate change HOAX has literally convinced a member of Congress that “the world is going to end in 12 years”	Not Credible
Freedom Bunker	Fight illness with this ancient immune booster	Not Credible
Jew World Order	Greenpeace Founder: Global Warming is a Hoax Pushed by Corrupt Scientists ‘Hooked on Government Grants’	Not Credible
The Gateway Pundit	NOAA Ruins Assertions by Unhinged Democrats that Global Warming Has Caused Increase in Hurricane Activity	Not Credible
Natural News	World Health Organization declares <i>anti-vax</i> movement to be a top “global health threat” just like the climate change hoax ... the vaccine deep state grows desperate	Not Credible
BBC	‘Completely avoidable’ measles outbreak hits 25-year high in US	Credible
NPR	Climate Change Was the Engine That Powered Hurricane Maria’s Devastating Rains	Credible
Chicago-Sun Times	Kentucky governor exposed his kids to chickenpox instead of getting vaccine	Credible
NPR	New U.S. Measles Cases Break 25-Year-Old Record, Health Officials Say	Credible
Fortune	U.S. Carbon Emissions Soared in 2018. Here’s Why	Credible
Emerging News		
Source	Title	Ground Truth
The New York Times	Open Windows. Don’t Share Food. Here’s the Government’s Coronavirus Advice.	Credible
Reuters	World Faces Chronic Shortage of Coronavirus Protective Equipment: WHO	Credible
The Guardian	Can a face mask stop coronavirus? Covid-19 facts checked	Credible
Breitbart	Hillary Clinton Falsely Claims to Jimmy Fallon That Trump Called Coronavirus Outbreak a ‘Hoax’	Not Credible
Natural News	Vitamin C infusions being studied in China as possible treatment for coronavirus-related pneumonia	Not Credible
Natural News	Spirulina found to boost the body’s type 1 interferon response to fight RNA viral infections “including coronavirus,” new science finds	Not Credible
The Russophile	CORONAVIRUS HOAX: Fake Virus Pandemic Fabricated to Cover-Up Global Outbreak of 5G Syndrome	Not Credible
The Russophile	CORONAVIRUS SPECIAL REPORT: Worldwide Outbreaks of 5G Syndrome and 5G Flu Driving Pandemic	Not Credible
Th Liberty Daily	Coronavirus: Chinese Espionage Behind Wuhan Bioweapon?	Not Credible

So, what's the potential for AI to be successful in detecting misinformation?

"The tools we have developed, and other tools developed in this area, have fairly high accuracy in lab settings," says Horne. "For example, our most recent technical work showed around 83% accuracy in predicting when the source of a news article is reliable or unreliable."

Despite the effectiveness of algorithms, old-fashioned fact-checking by journalists will still be required to combat fake news. AI could filter the information for fact-checkers to verify, according to Horne.

"AI tools are great at dealing with high quantities of information at fast speeds but lack

The Tech Alert Newsletter

Receive latest technology science and technology news & analysis from IEEE Spectrum every Thursday.

About the Tech Talk blog

IEEE Spectrum's general technology blog, featuring news, analysis, and opinions about engineering, consumer electronics, and technology and society, from the editorial staff and freelance contributors.

Follow @IEEE_Spectrum

Subscribe to RSS Feed

x