Machine learning system aims to determine if a news outlet is accurate or biased

October 4 2018

A machine learning system aims to determine if a news outlet is accurate or biased. Credit: Public domain

Lately the fact-checking world has been in a bit of a crisis. Sites like Politifact and Snopes have traditionally focused on specific claims, which is admirable but tedious—by the time they've gotten through verifying or debunking a fact, there's a good chance it's already traveled across the globe and back again.

Social media companies have also had mixed results limiting the spread of propaganda and misinformation: Facebook plans to have 20,000 human moderators by the end of the year, and is spending many millions developing its own fake-news-detecting algorithms.

Researchers from MIT's Computer Science and Artificial Intelligence Lab (CSAIL) and the Qatar Computing Research Institute (QCRI) believe that the best approach is to focus not on the factuality of individual claims, but on the news sources themselves. Using this tack, they've demonstrated a new system that uses machine learning to determine if a source is accurate or politically biased.

"If a website has published fake news before, there's a good chance they'll do it again," says postdoctoral associate Ramy Baly, lead author on a new paper about the system. "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."

Baly says the system needs only about 150 articles to reliably detect if a news source can be trusted—meaning that an approach like theirs could be used to help stamp out fake-news outlets before the stories spread too widely.

The system is a collaboration between computer scientists at MIT CSAIL and QCRI, which is part of the Hamad Bin Khalifa University in Qatar. Researchers first took data from Media Bias/Fact Check (MBFC), a website with human fact-checkers who analyze the accuracy

and biases of more than 2,000 news sites, from MSNBC and Fox News to low-traffic content farms.

They then fed that data to a machine learning algorithm called a Support Vector Machine (SVM) classifier, and programmed it to classify news sites the same way as MBFC. When given a new news outlet, the system was then 65 percent accurate at detecting whether it has a high, low or medium level of "factuality," and roughly 70 percent accurate at detecting if it is left-leaning, right-leaning or moderate.

The team determined that the most reliable ways to detect both fake news and biased reporting were to look at the common linguistic features across the source's stories, including sentiment, complexity and structure.

For example, fake-news outlets were found to be more likely to use language that is hyperbolic, subjective, and emotional. In terms of bias, left-leaning outlets were more likely to have language that related to concepts of harm/care and fairness/reciprocity, compared to other qualities such as loyalty, authority and sanctity. (These qualities represent the 5 "moral foundations," a popular theory in social psychology.)

Co-author Preslav Nakov says that the system also found correlations with an outlet's Wikipedia page, which it assessed for general length—longer is more credible—as well as target words like "extreme" or "conspiracy theory." It even found correlations with the text structure of a source's URLs: those that had lots of special characters and complicated subdirectories, for example, were associated with less reliable sources.

"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," says Sibel Adali, a professor of computer science at Rensselaer Polytechnic Institute who was not involved in the project.

Nakov is quick to caution that the system is still a work-in-progress, and that, even with improvements in accuracy, it would work best in conjunction with traditional fact-checkers.

"If outlets report differently on a particular topic, a site like Politifact could instantly look at our 'fake news' scores for those outlets to determine how much validity to give to different perspectives," says Nakov, a senior scientist at QCRI.

Baly and Nakov co-wrote the new paper with MIT senior research scientist James Glass alongside master's students Dimitar Alexandrov and Georgi Karadzhov of Sofia University. The team will present the work later this month at the 2018 Empirical Methods in Natural Language Processing (EMNLP) conference in Brussels, Belgium.

The researchers also created a new open-source dataset of more than 1,000 news sources, annotated with factuality and bias scores—the world's largest database of its kind. As next steps, the team will be exploring whether the English-trained system can be adapted to other languages, as well as to go beyond the traditional left/right bias to explore region-specific biases (like the Muslim World's division between religious and secular).

"This direction of research can shed light on what untrustworthy websites look like and the kind of content they tend to share, which would be very useful for both web designers and the wider public," says Andreas Vlachos, a senior lecturer at the University of Cambridge who was not involved in the project.

Nakov says that QCRI also has plans to roll out an app that helps users step out of their political bubbles, responding to specific news items by offering users a collection of articles that span the political spectrum.

"It's interesting to think about new ways to present the news to people," says Nakov. "Tools like this could help people give a bit more thought to issues and explore other perspectives that they might not have otherwise considered."

Provided by Massachusetts Institute of Technology