

# New AI tool that improves fake news detection accuracy

Praveen Natarajan

February 2, 2020

Researchers at the UW developed an automated fact-checking AI tool which can detect fake news with an astounding accuracy. Amidst the growing concerns over the dissemination of false information, it is now more important than ever to curb the propagation of falsehoods. The developed screening tool sets a new benchmark for accuracy, in that it can flag fake news correctly nine out of ten times.

The tool uses a technique called “Stance Detection”, which is an emerging area of research that determines the degree of relationship between the principal claim and other news stories. In other words, the learning algorithm compares the content in the claim with a news from a variety of different sources on the same topic and estimates whether these stories substantiate or refute the claim.

Alexander Wong, a professor of systems design engineering at UW, said the automated fact-checking system consists of a series of sequential sub-tasks starting from the *Document Retrieval* stage, that gathers news stories pertaining to the claim, followed by the aforementioned *Stance Detection* process, that aims to identify the stance or relative position of each article with respect to the claim, then the penultimate step, *Reputation Assessment*, which as the name suggests determine the trustworthiness or reputation of each news source, and subsequently culminating in *Claim Verification* step that determine the veracity of the claim in question.

Wong pointed out that UW’s AI screening tool could be used by fact-checkers as a swift and efficient means to flag fake news, which otherwise would be laborious when the task is done manually.

The developed learning algorithm leverages the features of a highly optimized open source bidirectional language model, which allows the neural network to gain a deeper understanding of the language and context used in the principal claim and the reference articles compared to earlier unidirectional models.

“Garnering reliable sources of dataset for training the model is one of the major challenges the research team had to deal with as bias in the data leads to undesirable outcomes. Therefore, the intended primary end-users of the developed automated tool will be journalists and fact checkers. That said, a simple user interface in the form of browser plug-in will be rolled out in the near future”, Wong said.

All in all, it is worth highlighting that although there have been some fascinating improvements in the algorithm’s capability to flag fake news setting news benchmark levels for accuracy, it continues to be a challenge to keep up with the barrage of disinformation masquerading as real news in many forms, thereby prompting subject matter experts to work toward sophisticated calibration of the tool’s fake news detection. Researchers at UW are determined to elevate the algorithm’s performance further by training their model with much larger dataset targeting media outlets that report news in non-English form.

