

RESEARCH NEWS

RESEARCH NEWS

February 1, 2019 || Page 1 | 3

Using machine learning to analyze social media data

Software that can automatically detect fake news

Invented stories, distorted facts: fake news is spreading like wildfire on the internet and is often shared without thought, particularly on social media. In response, Fraunhofer researchers have developed a system that automatically analyzes social media posts and deliberately filters out fake news and disinformation. To do this, the tool analyzes both content and metadata, classifying it using machine learning techniques and drawing on user interaction to optimize the results as it goes.

Fake news is designed to provoke a specific response or incite agitation against an individual or a group of people. Its aim is to influence and manipulate public opinion on targeted topics of the day. Fake news can spread like wildfire on the internet, particularly on social media such as Facebook or Twitter. What is more, identifying it can be a tricky task. That is where a classification tool developed by the Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE comes in, automatically analyzing social media posts and processing vast quantities of data.

As well as processing text, the tool also factors metadata into its analyses and delivers its findings in visual form. "Our software focuses on Twitter and other websites. Tweets are where you find the links pointing to the web pages that contain the actual fake news. In other words, social media acts as a trigger, if you like. Fake news items are often hosted on websites designed to mimic the web presence of news agencies and can be difficult to distinguish from the genuine sites. In many cases, they will be based on official news items, but their wording has been altered," explains Prof. Ulrich Schade of Fraunhofer FKIE, whose research group developed the tool.

Schade and his team begin the process by building libraries made up of serious news pieces and also texts that users have identified as fake news. Together, they form the learning sets used to train the system. To filter out fake news, the researchers employ machine learning techniques that automatically search for specific indicators, called features, in texts and metadata. For instance, in a political context, it could be formulations or combinations of words that rarely occur in everyday language or in journalistic reporting, such as "the current female chancellor of Germany." Linguistic errors are

Editorial Notes

Janis Eitner | Fraunhofer-Gesellschaft, Munich | Communications | Phone +49 89 1205-1333 | presse@zv.fraunhofer.de

Silke Wiesemann | Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE | +49 228 9435-103 | silke.wiesemann@fkie.fraunhofer.de
Fraunhoferstrasse 20 | 53343 Wachtberg | www.fkie.fraunhofer.de

also a red flag. This is particularly common when the author of the fake news was writing in a language other than their native tongue. In such cases, incorrect punctuation, spelling, verb forms or sentence structure are all warnings of a potential fake news item. Other indicators might include out-of-place expressions or cumbersome formulations.

“When we supply the system with an array of features, the tool will teach itself to select the markers that work. Another decisive factor is choosing the machine learning algorithm that will deliver the best results. It’s a very time-consuming process, because you have to run the various algorithms with different combinations of features,” says Schade.

Metadata yields vital clues

Metadata is also used as a feature. It plays a crucial role in differentiating between authentic sources of information and fake news: For instance, how often are posts being issued, when is a tweet scheduled, and at what time? The timing of a post can be very telling. For instance, it can reveal the country and time zone of the originator of the news. A high send frequency suggests bots, which increases the probability of fake news. Social bots send their messages to a huge number of users to spread uncertainty among the public. An account’s connections and followers can also prove to be fertile ground for analysts. This is because it allows researchers to build heat maps and graphs of send data, send frequency and follower networks. These network structures can be used to calculate which node in the network circulated an item of fake news or initiated a fake news campaign.

Another feature of the automated tool is its ability to detect hate speech. Posts that pose as news but also include hate speech often link to fake news. “The important thing is to develop a classifier capable of identifying clear cases of hate speech. Examples include expressions such as ‘political scum’ or ‘nigger’,” says the linguist and mathematician.

The researchers are able to adapt their system to various types of text in order to classify them. Both public bodies and businesses can use the tool to identify and combat fake news. “Our software can be personalized and trained to suit the needs of any customer. For public bodies, it can be a useful early warning system,” finishes Schade.



Picture 1: To identify fake news, Fraunhofer FKIE's new machine learning tool analyzes both text and metadata.

© Fraunhofer FKIE

RESEARCH NEWS

February 1, 2019 || Page 3 | 3
