

Working out how to catch sarcasm on the Internet

24 January 2016, by Nancy Owano



JD Scott
@MrJDScott



My favorite thing to do at 4am is go to the airport. How about you? #Sarcasm #AutoPilot

User self-reporting of sarcasm. Credit: David Bamman and Noah A. Smith

"Oh, lovely. Now there's a surprise. How charitable of her. What could possibly go wrong? What a genius." These are remarks you usually read on the Internet and most of the time they are uttered with a full package of snark intended. Most of the time.

Researchers are now trying to teach computers to recognize sarcasm in an effort to improve computers' ability to make sense of human communications, said Will Knight, senior editor for AI at *MIT Technology Review*.

Two scientists have been working on a system capable of recognizing instances of sarcasm on Twitter. They [authored a paper about their work](#) for the Proceedings of the Ninth International AAAI (Association for the Advancement of Artificial Intelligence) Conference on Web and Social Media.

"Contextualized Sarcasm Detection on Twitter" is by David Bamman and Noah A. Smith; they describe their effort as a series of experiments to understand the effect of "extra-linguistic information" on the detection of sarcasm. Who are the speakers? Who is the audience? Context rules.

Will Knight commented on the significance of their research. "Previous efforts to automatically

[recognize](#) sarcasm in text relied entirely on linguistic cues. What's interesting here is that the researchers tried to include some wider context, such as who the author was and what they were tweeting about. And they found it to be noticeably better than existing approaches, correctly guessing 85 percent of the time if a post was sarcastic."

Their main findings were that "Including any aspect of the environment (features derived from the communicative context, the author, or the audience) leads to improvements in prediction accuracy."

They also found that when users are less familiar with their audience, they are more likely to tag their message with the explicit hashtag #sarcasm.

This work interestingly involved computing resources made available by the Open Science Data Cloud (OSDC). This is a petabyte-scale scientific community cloud which has become a "data science ecosystem." Researchers can house and share their scientific data, access complementary public [datasets](#), build and share customized virtual machines with whatever tools necessary to analyze their data and perform analysis to answer their research questions. "It is a one-stop shop for making scientific research faster and easier."

(With datasets growing larger and larger, researchers have found that the bottleneck to discovery is less lack of data and more an inability to manage, analyze, and share large datasets. The goal of the group has been to remove the bottleneck.)

Discussing their results, the authors said that "while tweet-only information yields an average accuracy of 75.4 percent across all ten folds, adding response features pushes this to 77.3 percent, audience features to 79.0% and author features to 84.9 percent. Including all features together yields

the best performance at 85.1 percent, but most of these gains come simply from the addition of author information."

What if there is no audience familiarity, no shared common ground in geography, for example?

"In the absence of shared common ground required for their interpretation, explicit illocutionary markers are often necessary to communicate intent." They wrote that "the #sarcasm hashtag is not a natural indicator of sarcasm expressed between friends, but rather serves an important communicative function of signaling the author's intent to an audience who may not otherwise be able to draw the correct inference about their message."

Caitlin Dewey, writing in *The Washington Post* back in August, quoted computational linguist Bamman on [sarcasm](#) on the Internet. "Sarcasm detection is a very difficult computational [problem](#)," said Bamman.

More information:

www.aaai.org/ocs/index.php/ICW...per/view/10538/10445

github.com/kshreyas91/trueratr

© 2016 Tech Xplore

APA citation: Working out how to catch sarcasm on the Internet (2016, January 24) retrieved 12 June 2021 from <https://techxplore.com/news/2016-01-sarcasm-internet.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.