

The daily grind of the rumor mill: Machine learning deciphers fake news

August 11 2022, by David Bradley



Credit: Pixabay/CC0 Public Domain

Research published in the *International Journal of Cloud Computing* looks at how machine learning might allow us to analyze the nature and characteristics of social media updates and detect which of those updates are adding grist to the rumor mill rather than being factual.

Fake news has been with us ever since the first gossip passed on a rumor

back in the day. But, with the advent of [social media](#), it is now so much easier to spread fake news, disinformation, and propaganda to a vast global audience with little constraint. A rumor can make or break a reputation. These days, that might happen the world over through the amplifying echo chamber of social media.

Mohammed Al-Sarem, Muna Al-Harby, Faisal Saeed, and Essa Abdullah Hezzam of Taibah University in Medina, Saudi Arabia have surveyed the different text pre-processing approaches for approaching the vast quantities of data that pour from social [media](#) on a daily basis. How well these approaches work in the subsequent rumor detection analysis is critical to how well [fake news](#) can be spotted and stopped. The team has tested various approaches on a dataset of political news-related tweets from Saudi Arabia.

Pre-processing can look at the three most relevant characteristics of an update before the [text analysis](#) is carried out and silo the different updates accordingly: First, it can look at the use of question marks and exclamation marks and the word count. Second, it can look at whether an account is verified or has properties more often associated with a fake or bot account, such as tweet count, replies, retweets, etc. Third, it can look at user-based features, such as the user name and the user's logo or profile picture.

The researchers found that pre-processing can improve analysis significantly when the output is fed to any of support vector machine (SVM), multinomial naïve Bayes (MNB), and K-nearest neighbor (KNN) classifiers. However, those classifiers do react differently depending on what combination of pre-processing techniques is used. For instance, removing stop words, and cleaning out coding tags, such as HTML, stemming, and tokenization.

More information: Mohammed Al Sarem et al, Machine learning

classifiers with pre-processing techniques for rumour detection on social media: an empirical study, *International Journal of Cloud Computing* (2022). [DOI: 10.1504/IJCC.2022.124797](https://doi.org/10.1504/IJCC.2022.124797)

Provided by Inderscience

Citation: The daily grind of the rumor mill: Machine learning deciphers fake news (2022, August 11) retrieved 5 May 2024 from <https://techxplore.com/news/2022-08-daily-rumor-mill-machine-deciphers.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.