

# Abstract

The 2016 U.S. presidential campaigns were rocked by various scandals, many of which were ignited by fake news articles that spread like wildfire on social media platforms like Facebook and Twitter. When it came to light that many of these articles were purposefully constructed by foreign actors to influence the presidential election, it became apparent that social media platforms needed more safeguards to prevent individuals from deceiving the public for personal gain.

This study attempts to fulfill this need by building an automated system capable of detecting the malicious content published during the 2016 presidential campaign season. Using a set of articles flagged as false by Snopes, and a set of articles from leading news organizations, select machine learning algorithms (k-nearest neighbors, support vector machines, and long short-term memory networks) are trained using only each article's textual content. Other instances of these models are also given the sentiment-related features of each article to take advantage of any possible correlation between the overall sentiment of an article and its factual accuracy. The results of this study show that a long short-term memory network is capable of obtaining an overall accuracy and average F1-score of 90% with this dataset.