

Detection of Fake News Using Machine Learning and Natural Language Processing: Addressing the Modern Information Crisis

WPTM Pathirana^{1#}, MG Saranga¹, PR Withange¹, PDSS Pitiyage¹, and KH Wickramathunga¹

¹Faculty of Management, Social Sciences and Humanities, General Sir John Kotelawala Defence University, Sri Lanka

#40-adc-0021@kdu.ac.lk

Fake news is spreading exponentially due to increased social media use, which is causing serious problems for society worldwide. Although a lot of studies have been done on the topic of identifying fake news, real-time detection of dynamic and quickly changing information environments has not received as much attention. This work attempted to close this crucial gap by creating a machine-learning model that employs cutting-edge natural language processing (NLP) methods to identify and categorize bogus news in real-time. Deep learning models like LSTM and BERT as well as more conventional machine learning models like logistic regression, support vector machines, and decision trees were used. This was done utilizing Python's sci-kit-learn module for text tokenization and feature extraction. The models were enhanced for real-time performance by including streaming data processing capabilities after being trained on a sizable and varied dataset. The outcomes show that, in real-time detection settings, the BERT-based NLP model achieves up to 98% accuracy, outperforming other models in terms of accuracy. To evaluate the models' performance in real-world settings, they were also evaluated using measures like recall, precision, and latency. This model is an essential resource for news organizations and social media platforms because it not only improves the timeliness and accuracy of fake news identification, but also offers a proactive way to stop the spread of false information as soon as it starts.

Keywords: *fake news, social media, machine learning, natural language processing, Python, feature extraction, accuracy metrics*