

ABSTRACT

Detection of Fake News using different Machine Learning and Natural Language Processing Algorithms

The amount of information shared on the internet, primarily via web-based networking media, grows day by day. Because of the simple availability and exponential expansion of data through social media networks, distinguishing between fake and real information. Most smartphone users tend to read news on social media rather than on the internet. The information published on news websites often needs to authenticate. The simple spread of reports by way of sharing has included the exponential development of its misrepresentation. So, fake news has been a major issue ever since the web developed and expanded it to the general mass. This paper demonstrates several models and techniques for detecting false news by using different machine learning and natural language processing (NLP) models such as Logistic Regression, Decision Tree, Naïve Bayes, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), Bidirectional Encoder Representation from Transformers (BERT). We tried to combine the news, then find out if the information was authentic or fake. Various feature engineering methods such as Regex, Tokenization, stop words, Lemmatization, Term Frequency- Inverse Document Frequency (TF-IDF) generate feature vectors in this paper. Every Machine Learning and NLP model was evaluated with test data. For the machine learning model Logistic Regression, Decision Tree, Naïve Bayes, and SVM, we got 73.75%, 89.66%, 74.19%, and 76.65%, respectively. But the highest accuracy we get is for the NLP method, which is 95% for LSTM and 98% for the BERT language model.