World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:02, 2023

Sentiment Analysis of Fake Health News Using Naive Bayes Classification Models

Authors : Danielle Shackley, Yetunde Folajimi

Abstract : As more people turn to the internet seeking health-related information, there is more risk of finding false, inaccurate, or dangerous information. Sentiment analysis is a natural language processing technique that assigns polarity scores to text, ranging from positive, neutral, and negative. In this research, we evaluate the weight of a sentiment analysis feature added to fake health news classification models. The dataset consists of existing reliably labeled health article headlines that were supplemented with health information collected about COVID-19 from social media sources. We started with data preprocessing and tested out various vectorization methods such as Count and TFIDF vectorization. We implemented 3 Naive Bayes classifier models, including Bernoulli, Multinomial, and Complement. To test the weight of the sentiment analysis feature on the dataset, we created benchmark Naive Bayes classification models without sentiment analysis, and those same models were reproduced, and the feature was added. We evaluated using the precision and accuracy scores. The Bernoulli initial model performed with 90% precision and 75.2% accuracy, while the model supplemented with sentiment labels performed with 90.4% precision and stayed constant at 75.2% accuracy. Our results show that the addition of sentiment analysis did not improve model precision by a wide margin; while there was no evidence of improvement in accuracy, we had a 1.9% improvement margin of the precision score with the Complement model. Future expansion of this work could include replicating the experiment process and substituting the Naive Bayes for a deep learning neural network model.

Keywords : sentiment analysis, Naive Bayes model, natural language processing, topic analysis, fake health news classification model
Conference Title : ICNLP 2023 : International Conference on Natural Language Processing
Conference Location : Paris, France
Conference Dates : February 06-07, 2023