Abstract

The development of information and communication technology is currently increased, especially related to social media. Nowadays, many people get information through social media, especially Twitter, because of its easy access and it doesn't cost much. However, it has a negative impact in the form of spreading fake news or hoaxes that are difficult to detect. In this research, the authors developed a hoax news detection model using the Convolutional Neural Network and the TF-IDF weighting method. Feature selection is performed using Information Gain with various features, such as unigram, bigram, trigram and a combination of the three. Testing is done with 3 scenarios, classification, classification by weighting, classification by weighting and feature selection. The parameter used in the information gain feature selection is the threshold 0.8. The results showed that the classification by weighting and feature selection produced the highest accuracy that is equal to 95.56% on the unigram + bigram features with a comparison of training data and test data 50:50.