

Abstract

Covid-19 was declared a pandemic by WHO in March 2020. The virus has been spread to all around the world, including Indonesia. Vaccination is one of the government's efforts to prevent the spread of the virus. However, there are various opinions from social media platforms users, including Twitter users. On Twitter, opinions are called tweets. For each tweet, sentiment analysis will be given which is classified into neutral, positive and negative classes. By using the crawling method author obtained 3,000 tweet data. To improve the model performance, the weighting process on the data is given using the TF-IDF method. The weighting is divided into three parts, such as unigram, bigram, and unigram+bigram. The weighting results are classified using two models, that is Random Forest and Support Vector Machine. The best accuracy between the two models are SVM model using a polynomial kernel with 91% accuracy score and unigram +bigram weighting.

Keywords: SVM, random forest, sentiment analysis, TF-IDF