Abstract

The film industry in recent years has become one of the industries that people are most interested in. The convenience of watching movies through streaming services is one of the reasons why watching movies is so popular. This ease of access resulted in a large selection of available movies and encouraged the public to look for movie reviews to find out whether the movies was good or bad. Freedom of expression on the internet has resulted in many movie reviews being spread. Therefore, sentiment analysis was conducted to see the positive or negative of these reviews. The method used in this research is Random Forest and Word2Vec skip-gram as feature extraction. The Random Forest classification was chosen because Random forest is a highly flexible and highly accurate method, while Word2Vec Skip-Gram is used as a feature extraction because it is an efficient model that studies a large number of word vectors in an irregular text. The best model obtained from this experiment is a model built with stemming, Word2Vec with 300 dimensions, and a max_depth value of 23, achieving an f1-score of 83.59%.

Keywords: Sentiment Analysis; Random Forest; Word2Vec; Movie Review