Abstract

Sentiment analysis has become an important approach in understanding public responses and opinions towards various things, such as products, services, or certain issues. In the digital and social media era, text data generated by social media users has great potential as a source of information for sentiment analysis. This research aims to develop a sentiment analysis model using the *Naive Bayes* method on text data from social media. In the development stage, text data is taken from the twitter_samples dataset provided by *NLTK* (*Natural Language Toolkit*). The application of *NLP* in the form of text processing is carried out, including *cleaning* (removing punctuation, numbers, symbols), *tokenization* (separating words in sentences), *stopword removal* (removing words that do not have too meaningful words), and *stemming* (changing words to English root words). After that, the *Naive Bayes* classification model was trained using the processed training data. The results of this research include the development of a sentiment analysis model that can classify English comment text from social media as positive or negative sentiment. The results of this study show that the Naive Bayes algorithm can be used to classify sentiment on English text using a division ratio of 80% training data and 20% test data has an accuracy of ± 79%-80% on the dataset used. The result of this research is to successfully implement *NLP* in sentiment analysis applications on English text using the *naive bayes* algorithm.