

Abstract

Twitter is a social network which is growing in the world, including Indonesia. Twitter users in Indonesia reached 9.9 million with accessing rate 22% of all Indonesian netizen population. This shows that Twitter is one of the media to be reckoned with to analyze Indonesian public sentiment.

This undergraduate thesis will use multinomial naive bayes to classify tweets by their sentiment, that are positive, negative or neutral. The data collection that was used was the data collection with Indonesian language, because most of Indonesian people are tweeting in Indonesian language. And for the classification were performed using unigrams, bigrams, trigrams, and pos tag as the attributes.

After testing with several scenarios, it showed that the multinomial naive bayes is relevant to be used for classifying Indonesian language tweets by their sentiment. This is shown by the system accuracy which is achieved 95.539%.

Keywords: sentiment analysis, twitter, multinomial naive bayes.