

Abstract

One of the biggest social media in the world that is in great demand is Twitter. Twitter is usually used by the public to convey messages and opinions in the form of opinions about certain topics and topics that become the latest trending topic. One of the topics which was widely discussed was the news that Mr. Erick Thohir appointed and announced Mr Ahok as President Commissioner of PT Pertamina (Persero). In social media Twitter, opinions expressed in the form of positive opinions and negative opinions with a relatively large amount of data. Therefore, manual assessment of sentiments is somewhat difficult because of the large amount of data that automation needs to be implemented that can facilitate this. In this study using the Feature Selection Query Expansion Ranking method and the K-Nearest Neighbor classification model with data crawling technique taken from Twitter using the Twitter API and using the term weighting method with the Term Frequency method. Based on the results of the study, the best accuracy is obtained with a scenario of 100% features at $k = 3$ with an accuracy value of 84% and the lowest accuracy value at a scenario of 60% features at $k = 9$ with an accuracy value of 71%. With the results of these studies it can be concluded that the value of k in the K-Nearest Neighbor classification affects the accuracy value% of the features generated by the Feature Selection Query Expansion Ranking on Twitter data.

Keywords: Twitter, Feature Selection QER, KNN, % feature