

Abstract

Public opinion on important services for owners of the company for the repair of existing services as well as predict what services can be added in the future and knowing the position of the company among its competitors. The sentiment of the community can serve as one of the indicators by the company to get information and feedback from consumers. On twitter can do the process of extracting information regarding community sentiment against GO-JEK excavation technique and information on twitter sentiment analysis is.

The process begins with the preprocessing continued with the weighting of words, then the designation which consists of calculating cosine similarity and classification of sentiment. Preprocessing including cleansing, case folding, tokenizing, stopword removal, and stemming. Method of weighting using TF-IDF and its classification method using Improved K-Nearest Neighbor. A method of Improved K-Nearest Neighbor has advantages in the form of stability on any variation of the value of k.

The results obtained through implementation and testing of the system is the amount of training data and test data as well as the value of k to the precision of the results of the analysis of sentiment. The average Precision obtained the system amounted to 79%, the average Recall of 88%, average F-Measure amounted to 83% and an average Accuracy of 81% at $k = 1$ with the composition of the training data and test data by 90%: 10% so that it can be concluded the effectiveness the system has been running quite well.

Keywords: twitter, GO-JEK, sentiment analysis, Improved K-Nearest Neighbor