

Abstract

Technological advances are developing very rapidly, especially in the field of computing, where at first the work that computers do is now done by humans in order to help ease the work of humans themselves. As in the case study, this research is a system that can classify abstract text in journals into subject groups. Subjects in journals are the focus of themes written by researchers in appropriate research and there are several subjects such as religion, agriculture, industry, social, health, etc. By knowing the subject of the journal, the reader will know about the contents of the journal. The method used in this research is to carry out the TF-IDF (Term Frequency Inverse Document Frequency) weighting method and Multinomial Naïve Bayes. On the TF method- IDF is used to weight each word contained in the document by integrating the POS Tag in the tokenization process before classification by the system can affect the results of accuracy which, if done without a POS Tag and there is a POS Tag, the results will be used into the Naïve Bayes Multinomial method in order to. Based on the evaluation using the confusion matrix using 640 training data and 160 test data, the researcher got the accuracy of the data using Tagger POS at 81.37% and testing on the data without using POS Tagger was 75.12%.

Keywords: Multinomial, Naïve Bayes, TF-IDF, accuracy