

Abstract

One of Data Mining Task is a classification of text as a way to classify documents into particular topic or category. Currently, text classification can be used to classify a text containing homograph words. This text classification makes the user easier to determine pronunciation of a homograph word.

Naive Bayes and K-Nearest Neighbour Methods are two of several classification methods. The accuracy of these two methods has not been measured yet both in direct comparison and with the same data. Naive Bayes method represents the document into the form of probability model. That probability model used as a model to classify a document. While the method of K-Nearest Neighbour keep a group of training data as reference to classify a document.

This final task discusses Naive Bayes and K-Nearest Neighbor Method in the case of text classification of homograph words. Besides, there is also analysis of the performance results of two classification methods. Performance parameters used are precession, recall, and f-measure. In final Naive Bayes method resulted f-measure value higher than k-nearest noighbor method.

Keywords: *probability model, training set, precission, recall, f-measure.*