

Abstract

Along with the increasing development of the internet, the growth of textual information on the internet continues to increase. With the increase of such information, then the need for automatic classification of news is needed to find the desired information or news. One way to classify a news into a particular category based on the information contained in the news is text classification.

One method in text classification is Bayesian Network. Bayesian Network is one of the reasoning methods that model the relationship between variables in Probabilistic Graphical Model (PGM). Bayesian Network advantages compared to other methods that are, suitable for small and incomplete datasets, can handle uncertainty and decision making, and rapid computation. In addition, feature selection is made using the Mutual Information method to reduce the number of dimensions and to improve performance.

The results of this classification are expressed in F1-measure micro-average with performance value 75,34%.

Keywords: Text classification, Bayesian Network, Mutual Information, F1-measure.