

Abstract

In news article, generally has more than one label, known as *multi-label*. In multi-label itself there are two main issues, namely the labels are correlated and interdependent with each other and the data has high dimensionality. To address the problem of high dimensionality can be applied to feature selection, while for feature selection in this case must also consider the correlation between attributes is called the Correlated Multi-label Feature Selection (CMLFS), one of feature selection method. The Feature selection is done based on a correlation coefficient within attribute, divided into high and low correlation. To determine the effect of CMLFS, the classification process is required, using the method of classification Support Vector Machine (SVM). Assessment of the performance was evaluated using the testing accuracy and classification time. The result after feature selection based on high and low correlations generally showed a decrease in accuracy, but the classification time will be faster, because the number of attributes getting smaller.

Keyword: *feature selection, Support Vector Machine, multi-label, correlation, correlated multi-label feature selection.*