

Abstract

The increasing of internet's using has made the growth and exchanging of informations become higher than before. The volume of Indonesian electronic news become bigger and its save valuable information in it. The grouping of Indonesian news is one of solution which can be used to catch valuable information easier. *Clustering* can be used to help analizing news by grouping news which have the similarity automatically.

Text *clustering* has a problem, that is high dimension of features. *Feature selection's* method is needed to reduce this problem. *Feature selection* has the ability to reduce data dimension so it can improve *clustering's* performance. There are some approaches as the technique of *feature selection's* implementation, one of them is filter based feature selecion.

On this final project, the analysis of *feature selection's* method between *Document Frequency* and *Term contribution* is done. These methods are implemented by filter *feature selection*. At the end of testing, can be proved that *Term contribution* is better than *Document Frequency*, because it considers *term* frequency in a document and the amount of *document frequency*, so the choosen *terms* are unique or discriminatory. It can improve *clustering's* performance with precision and entropy as the points to measure the performance.

Key Word : *clustering, filter feature selection, Term contribution, Document Frequency.*