

## Abstract

*Document clustering is a very important problem at this time. This is caused by the increase of people that stores digital document but not yet clustered well enough. This cause difficulties in finding the document when we need to get information information in the document. Many algorithm can be used in document clustering. In this final assignment, document clustering is done by using Active Fuzzy Constrained Clustering.*

*In this final assignment, generalization is done to the build vector process before the document is clustered with Active Fuzzy Constrained Clustering. The build vector method that are going to be used is Vector Space Model (VSM) and Generalized Vector Space Model (GVSM) with semantic information addition. With the use of this two method, it can be seen that the difference and the factor that can affect the accuracy improvement in the Active Fuzzy Constrained Clustering Algorithm in document clustering.*

*The similarity calculation by using wordnet thesaurus (GVSM) is very dependant to wordnet, if the word in the testing document can detect well in the wordnet, then the accuration will also be good.*

**Keyword:** *Active Fuzzy Constrained Clustering, Vector Space Model (VSM), Generalized Vector Space Model (GVSM), wordnet.*