

Abstract

Intrinsically a text document is a kind of unstructured data base because it doesn't has fields such as conventional database. Difference in topic of a document with another document mean that those documents contain different information.

Text categorization is task to assign document into a predefined set of category. Nowadays there are a lot of text categorization method such as *K-Nearest Neighbor (KNN)*, *Bayesian Classifier*, *Decision Tree* and *Rocchio method* applicable to classification problem.

Hierarchical categorization problems are found a lot in the real life, such as storing digital files into a hierarchically structured folders. One of implementation in categorization is by utilize *Fuzzy Relational Thesaurus (FRT)* as a class hierarchy structure in categorization.

This final project has implemented a text classification method that uses FRT as topic hierarchy and *Rocchio* method as an assist method to build its classifier. Test result showed that training process to select the best feature in FRT method could produce a better classifier than *Rocchio* method's.

Keywords : *hierarchical text categorization, Rocchio method, Fuzzy Relational Thesaurus.*