

Abstract

Text Summarization is a process to produce a summary from articles but it has accurate main points from the content of the articles. Objective of this process is to take the source of information by citing most of the important content and show it to the readers in a simple form that appropriate with reader's need.

This system implements semantic method Lexical Chain [1] which is one of the text summarization method that determine semantic relation between words (terms) and sentences. As a comparison of Lexical Chain method used another text summarization method is TF-IDF (Term Frequency*Inverse Document Frequency) modified using keyphrase extraction [10] concept. Thus, hopefully this system can help readers to get informations from articles through a summary. Text Summarization will produce a text that still has the main points from the original articles.

Test result of 40 single documents showed that Lexical Chain Method had better performance on 50%-20% summary length especially for documents which have length less than 29 sentences (average length sentence of 40 documents). This method gave precision score ≥ 0.502 and F-Measure score ≥ 0.312 , so it's much better than TF-IDF with Keyphrase Extraction method.

Multidocuments test of 6 folder documents, showed that TF-IDF with Keyphrase Extraction method had better performance on summary length 20%-10%. This method gave precision score ≥ 0.524 and F-Measure score ≥ 0.191 , so it's much better than Lexical Chain Method and TIPSTER SUMMAC Evaluation.

Keyword: Text Summarization, Lexical Chains, TF-IDF, Keyphrase Extraction.