Abstract

Evaluation of text summarization automatic machine is a process to determine the quality of an automatic text summarization. This process produces value, obtained with comparing the summaries generated by text summarization automatic machine (also known as a candidate summary) with a summary (ideally) made by other people (also known as a reference summary) originating from the same source.

In this final project, the author implements the development of Longest Common Subsequences (LCS) and Skip Bigram Co Occurrences Statistics (SB) methods in ROUGE techniques for evaluating multi-document text summarization automatic machine. The objective of this development of methods is to produce a way of evaluation that has a good level of corelation with the human judgment, because the information becomes more and more drawn

Testing is done by using Pearson's Coefficient of Correlation to measure the linear relationship between the value given by the application and the average value of human judgment of the same summary. Test results showed that the evaluation by using LCS+ SB algorithm is capable of producing higher evaluation score than LCS or SB algorithms, with the correlation result that sometimes the is higher than the other two algorithms. In addition, the correlation can be better by adding the percentage of the number of used summary reference

Key word: *Summarization Evaluation*, *Longest Common Subsequences*, *Skip Bigram Co Occurrences Statistics*, ROUGE*, Pearson's Coefficient of Correlation*