

Abstract

Information overloaded is a popular term used to describe the difficulty in understanding an issue and making decisions, because too much information is obtained. Summarization is one of the solution to overcome the information overloaded problem. Summarization can be defined as translation from a language that has more words to a language that has less words. Extraction summary is a style of summary that reusing the words from original text, and then shrink the portion by deleting the words which are not essential, become one of summarization's style that the researchers commonly used for article summary. Hybrid Statistic and Grammar Hidden Markov Model will be implemented with extraction summary, by calculating the weight from the connection between the words and the connection between the POS tagging in HMM's evaluation phase. In decoding phase, summary is done by choosing the maximum probability path from bag of words which have connection's weight from evaluation phase. Finally, by using ROUGE-2, by calculating the total of the same bigram from the result from automatic summarization with human's summarization and divide by the number of existing bigram in the summary, then we get the accuracy of our summary.

Keywords : information overloaded, summary, extraction, Hidden Markov Model, decoding, evaluation, ROUGE-2, accuracy.