Abstract

Information retrieval system is a system consisting of several stages in an attempt to find relevant documents in a collection of documents. In general, information retrieval systems use a stemmer to reduce various morphological variation found in the source document. Since the document related to the language used, it is generally used stemmer is based on the rules of a particular language. While the language used in documents that could have been more than one language.

This system implements a n-gram stemmer stemming which is a method that is not bound by the rules of any language. As test material, the language used in the document consists of English and French. Also as a comparison with the commonly used stemmer in one language is used porter stemmer.

The test results that have been done on this final show the application of n-gram stemmer can be done to document more than one language. Compression strength and accuracy of the resulting n-gram stemmer if compared with the porter, the results vary depending on the value of n on n-gram is used.

Keyword: *information retrieval*, *stemming*, n-gram *stemmer*