

PSEUDO-RELEVANCE FEEDBACK USING SENTENCE SEGMENTATION

Woro Indriyani, Julio Adisantoso

Department of Computer Science

Faculty of Mathematics and Natural Sciences

Bogor Agricultural University

julioipb@gmail.com

August 13, 2011

Abstract

A very large amount of information has stimulated the development of search engine to help users in finding information they need. To retrieve information that relevant to the user's needs, users should be able to formulate queries correctly. Pseudo-relevance feedback is an automatically local analysis technique for improving queries. This technique takes the top n-ranked documents and takes the top x-ranked terms from relevant documents. The purpose of this research is to implement query expansion with pseudo-relevance feedback using sentence segmentation. There are two groups of documents, 1.000 agriculture documents and 93 medicine plants documents which are used. The test result shows that the use of medicine plants documents is better than agriculture documents. This is due to agriculture documents have a high enough similarity between documents. The performance of information retrieval with pseudo-relevance feedback using sentence segmentation gave good result with around 89% average precision for medicine plants documents and 56% for agriculture documents.