

Query Classification for Information Retrieval of Conference Papers in Digital Library

Ei Ei Kay Khaing, Khin Mo Mo Tun
University of Computer Studies, Yangon
eieikaykhaing.ekk@gmail.com

Abstract

Classification is one of the most widely used data mining techniques with a lot of extension. Query classification is crucial for web search and advertising. Classification of web queries is important to provide search result effectively. As more and more documents become electronically available, this system provides to find documents in large database that fit users' need. Information can be extracted from summaries for the words contained in the documents. This system presents the implementation of query classification system for information retrieval of conference papers in digital library. In this system, Multinomial Naïve Bayes algorithm is used in computing weights of terms for each classes and then combined the overall weights. Cosine similarity algorithm is used to find the relevant documents. Resultant documents are ranked according to their class weights and similarity value. By using the system, user can obtain more relevant result from digital library since the document can be viewed by their degree of similarity values.