

Analisis kinerja information retrieval dengan menggunakan local weighted relevancy latent semantic indexing = Performance analysis of information retrieval with local weighted relevancy latent semantic indexing method

Norman Indradi Wirahardja, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20242718&lokasi=lokal>

Abstrak

LSI (Latent Semantic Indexing) adalah suatu metode yang digunakan dalam information retrieval. Menurut penelitian, LSI ini merupakan metode yang paling baik dalam mencari informasi seperti teks. Akan tetapi LSI kurang baik dalam melakukan pengelompokan dokumen. Dokumen-dokumen yang akan diproses harus dikelompokkan terlebih dahulu supaya proses LSI lebih efektif. Pada skripsi ini dianalisis sebuah metode baru dalam melakukan information retrieval, yaitu Local Weighted Relevancy LSI. (LWR-LSI). Metode ini menambahkan proses pembobotan (weighting) pada metode lokal LSI. Jadi, setelah dikelompokkan, dokumen-dokumen yang relevan akan dibobotkan. Pembobotan yang dilakukan terhadap dokumen pun bermacam-macam. Dengan adanya pembobotan ini, dokumen dapat diurutkan sesuai dengan tingkat relevansi dokumen tersebut dengan kata kunci pada query. Semakin relevan, maka ranking dokumen tersebut semakin tinggi. Aplikasi information retrieval yang dibuat ini merupakan aplikasi web-based dengan menggunakan bahasa pemrograman PHP. Kemudian, kinerja dari metode LWR-LSI pada aplikasi ini pun akan dianalisa. Analisis kinerja ini dilakukan berdasarkan waktu yang diperlukan untuk menyelesaikan proses information retrieval pada saat diletakkan dalam sistem localhost dan website di internet. Beragamnya data yang diperoleh disebabkan oleh beberapa faktor, antara lain faktor query yang dimasukkan oleh user, faktor digunakan atau tidaknya klasifikasi, faktor batas relevansi dari dokumen yang ingin ditampilkan, dan faktor pembobotan yang digunakan. Berdasarkan analisa yang dilakukan dapat disimpulkan bahwa metode LWRLSI ini lebih baik dari metode LSI biasa dalam aplikasi information retrieval. Kesimpulan tersebut diambil berdasarkan waktu yang dibutuhkan untuk menyelesaikan seluruh proses dan ketepatan yang dihasilkan. Metode LWR-LSI dapat mengurangi waktu yang diperlukan untuk menyelesaikan proses sampai 1,721% dari metode LSI biasa.

.....LSI (Latent Semantic Indexing) is known as one of the method that used in an information retrieval. According to the research, LSI is the best method for retrieving information like documents. But, LSI is not good enough to classify the documents. Before those documents are being processed, they have to be classified first. So, in this paper, the author use the new method, known as Local Weighted Relevancy LSI (LWR-LSI) for retrieving information. This method add a weighting process in the local LSI. So, after being classified, the relevant documents will be weighted. The process is using different formulas. The result, the documents can be ranked with the relevancy to the user query. More relevant the document, higher the rank it will get. The information retrieval application is an web-based application with PHP script. Then the performance of the method in this application will be analyzed. Analyzing the performance is based on the time that needed to finish the whole process in the localhost system and the internet network (website). The various received data is caused by some factors, like user query factor, classification factor, relevancy limit factor, and weighting factor. After analyzing the data, can be concluded that the LWR-LSI method in the information retrieval application is better than the former LSI method. This conclusion is based on the speed

of the process and the accuracy of the result. The new method can decrease 1,721% the needed time if it is compared to the old method.