

Preference based term weighting for arabic fiqh document ranking

Khadijah Fahmi Hayati Holle, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20448135&lokasi=lokal>

Abstrak

In document retrieval, besides the suitability of query with search results, there is also a subjective user assessment that is expected to be a deciding factor in document ranking. This preference aspect is referred at the fiqh document searching. People tend to prefer on certain fiqh methodology without rejecting other fiqh methodologies. It is necessary to investigate preference factor in addition to the relevance factor in the document ranking. Therefore, this research proposed a method of term weighting based on preference to rank documents according to user preference. The proposed method is also combined with term weighting based on documents index and books index so it sees relevance and preference aspect. The proposed method is Inverse Preference Frequency with ω value (IPF ω). In this method, we calculate preference value by IPF term weighting. Then, the preference values of terms that is equal with the query are multiplied by ω . IPF ω combined with the existing weighting methods become TF.IDF.IBF.IPF ω . Experiment of the proposed method uses dataset of several Arabic fiqh documents. Evaluation uses recall, precision, and f-measure calculations. Proposed term weighting method is obtained to rank the document in the right order according to user preference. It is shown from the result with recall value reach 75%, precision 100%, and F-measure 85.7% respectively.

Dalam pencarian, selain kesesuaian query dengan hasil pencarian, terdapat penilaian subjektif pengguna yang diharapkan menjadi faktor penentu dalam perankingan dokumen. Aspek preferensi tersebut tampak pada pencarian dokumen fiqh. Seseorang cenderung mengutamakan metodologi fiqh tertentu meskipun tidak mengabaikan pendapat metodologi fiqh lain. Faktor preferensi menjadi hal yang diperlukan selain relevansi dalam perankingan dokumen. Oleh karena itu, pada penelitian ini diajukan metode pembobotan kata berbasis preferensi untuk merankingkan dokumen sesuai dengan preferensi pengguna. Metode yang diajukan digabungkan dengan pembobotan kata berbasis indeks dokumen dan buku sehingga mampu memperhatikan aspek kesesuaian (relevance) dan keutamaan (preference). Metode pembobotan yang diusulkan disebut dengan Invers Preference Frequency with ω value (IPF ω). Langkah pembobotan yang diusulkan yaitu dengan perhitungan nilai preferensi term dengan pembobotan IPF. Kemudian nilai preferensi dari term dokumen yang sama dengan term query dikalikan dengan ω sebagai penguat. IPF ω digabungkan dengan metode pembobotan yang telah ada menjadi TF.IDF.IBF.IPF ω . Pengujian metode yang diusulkan menggunakan dataset dari beberapa dokumen fiqh berbahasa Arab. Evaluasi menggunakan perhitungan recall, precision, dan F-measure. Hasil uji coba menunjukkan bahwa dengan pembobotan TF.IDF.IBF.IPF ω diperoleh perankingan dokumen dengan urutan yang tepat dan sesuai dengan preferensi pengguna. Hal ini ditunjukkan dengan nilai maksimal recall mencapai 75%, precision 100%, dan F-measure 85.7%.