

Term weighting based on positive impact factor query for arabic fiqh document ranking

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20447836&lokasi=lokal>

Abstrak

Query becomes one of the most decisive factor on documents searching. A query contains several words, where one of them will become a key term. Key term is a word that has higher information and value than the others in query. It can be used in any kind of text documents, including Arabic Fiqh documents. Using key term in term weighting process could led to an improvement on result's relevancy. In Arabic Fiqh document searching, not using the proper method in term weighting will relieve important value of key term. In this paper, we propose a new term weighting method based on Positive Impact Factor Query (PIFQ) for Arabic Fiqh documents ranking. PIFQ calculated using key term's frequency on each category (mazhab) on Fiqh. The key term that frequently appear on a certain mazhab will get higher score on that mazhab, and vice versa. After PIFQ values are acquired, TF.IDF calculation will be done to each words. Then, PIFQ weight will be combine with the result from TF.IDF so that the new weight values for each words will be produced. Experimental result performed on a number of queries using 143 Arabic Fiqh documents show that the proposed method is better than traditional TF.IDF, with 77.9%, 83.1%, and 80.1% of precision, recall, and F-measure respectivel.

Query menjadi salah satu faktor penentu dalam pencarian dokumen. Dalam sebuah query terdiri dari beberapa kata, dimana salah satunya menjadi key term. Key term adalah kata yang memiliki nilai informasi dan bobot lebih tinggi dibandingkan kata lain. Hal tersebut berlaku untuk semua jenis dokumen teks, termasuk dokumen fiqh berbahasa Arab. Penitik beratan pada key term dalam proses pembobotan kata memungkinkan terjadinya peningkatan relevansi pencarian. Di dalam pencarian dokumen fiqh berbahasa Arab, jika metode pembobotan kata yang digunakan tidak tepat, key term tidak akan memberikan pengaruh berarti. Oleh karena itu diusulkanlah sebuah metode pembobotan baru pada kata berbasis Positive Impact Factor Query (PIFQ) untuk perangkaan dokumen fiqh berbahasa arab. PIFQ dihitung menggunakan frekuensi kemunculan key term pada setiap kategori (mazhab) dalam fiqh. Semakin tinggi frekuensi key term tersebut pada suatu mazhab semakin tinggi pula nilainya pada mazhab tersebut, begitu pula sebaliknya. Setelah didapat nilai PIFQ, kemudian dilakukan perhitungan TF.IDF untuk setiap kata. Selanjutnya bobot PIFQ akan dikombinasikan dengan TF.IDF sehingga menghasilkan bobot baru untuk masing-masing kata. Hasil dari pengujian yang dilakukan pada sejumlah query dengan 143 dokumen fiqh berbahasa Arab menunjukkan bahwa metode usulan dapat lebih unggul jika dibandingkan metode TF.IDF, dengan nilai precision, recall, dan F-measure masing-masing sebesar 77,9%, 83,1%, dan 80,1%.