

Abstraksi

Information Retrieval (IR) merupakan bagian dari *computer science* yang berhubungan dengan pengambilan informasi dari dokumen-dokumen yang didasarkan pada isi dan konteks dari dokumen-dokumen itu sendiri. Proses dalam *Information Retrieval* dapat digambarkan sebagai sebuah proses untuk mendapatkan *relevant documents* dari *collection documents* melalui pencarian *query* yang diinputkan *user*.

Parameter uji untuk menilai relevansi sebuah dokumen yang digunakan dalam Tugas Akhir ini adalah *precision*, *recall* dan IAP. *Precision* adalah parameter untuk menghitung nilai tingkat ketepatan antara *query* dengan dokumen koleksi. *Precision* merupakan hasil dari perbandingan antara dokumen relevan dengan seluruh dokumen yang berhasil diambil oleh sistem. *Recall* adalah parameter untuk menghitung nilai tingkat kelengkapan antara *query* dengan dokumen koleksi. *Recall* merupakan hasil perbandingan antara dokumen relevan dengan dokumen relevan yang ada dalam seluruh dokumen koleksi. IAP menghitung nilai keterurutan dokumen relevan dalam sistem. Nilai IAP didapatkan dari perhitungan nilai *precision* dan *recall*

Dalam IR terdapat model untuk mendapatkan nilai similiarity dokumen yang relevan dengan *query* yang diinputkan oleh *user*. Salah satu model tersebut adalah *Vector Space Model*. Terdapat 3 metode pembobotan dalam *Vector Space Model*, yaitu *Term Frequency* (TF), *Inverse Document Frequency* (IDF) dan *Term Frequency-Inverse Document Frequency* (TF-IDF). Dari hasil pengujian didapatkan bahwa dengan parameter IAP metode TF-IDF lebih unggul dibandingkan dengan dua metode yang lain. Hal ini menunjukkan bahwa metode pembobotan TF-IDF lebih baik dibandingkan dengan 2 metode lainnya, yaitu TF dan IDF.

Kata Kunci: *Information Retrieval, Information Retrieval System, Vector Space Model, precision, recall, IAP, TF, IDF, TF-IDF, document collection dan query.*