

Nama : Calvin

NIM : 162112233003

Resume Information Retrieval

Model pencarian Boolean mengalami dilema "feast or famine" di mana hasil yang didapat sering kali terlalu berlebihan atau bahkan tidak ada. Ini terjadi karena kesulitan dalam menggunakan operator AND dan OR untuk mendapatkan hasil yang relevan dan sesuai dengan keinginan pengguna

Model pencarian dengan peringkat menawarkan sebuah solusi dengan mengurutkan dokumen-dokumen berdasarkan tingkat kepentingannya terhadap query yang dimasukkan. Daripada harus menggunakan bahasa query yang kompleks dengan operator-operator, pengguna dapat memasukkan query dalam bentuk teks bebas menggunakan bahasa sehari-hari. Esensi dari model ini adalah memberi skor dan ranking pada dokumen berdasarkan kesamaannya dengan istilah-istilah dalam query.

Untuk menghitung relevansi antara dokumen dengan query, digunakan beberapa metrik seperti frekuensi kemunculan istilah query dalam dokumen (TF) dan seberapa unik istilah tersebut muncul di seluruh koleksi dokumen (IDF). Skema tf-idf menggabungkan kedua aspek ini.

Dokumen dan query dapat diwakili sebagai vektor dalam ruang vektor istilah yang berdimensi tinggi. Skor relevansi dihitung dengan mengukur kesamaan kosinus antara vektor query dan dokumen. Sistem pembobotan seperti tf-idf atau BM25 digunakan untuk membentuk vektor-vektor tersebut.

Pada akhirnya, sistem pencarian berperingkat menampilkan sejumlah dokumen yang paling relevan dalam jumlah terbatas, menghindari kelebihan informasi atau ketiadaan hasil sama sekali bagi pengguna. Hal ini memberikan pengalaman pencarian yang lebih intuitif dibandingkan dengan model Boolean yang lebih kaku.