

Nama : m Dimas Rizky Tunggono

Nim : 141221030

Prodi : Teknik Informatika

- 1) * Query formulation, yaitu tahap ketika pengguna merumuskan kebutuhan informasinya ke dalam kata kunci yang sesuai. Misalnya, seorang mahasiswa mengetik kata kunci "deep learning untuk pengolahan citra" pada mesin pencari.
* retrieval, yaitu proses ketika sistem mencari dokumen atau data relevan berdasarkan query yang diberikan. Contohnya, mesin pencari menampilkan artikel atau jurnal terkait.
* evaluation, yaitu tahap ketika pengguna mengevaluasi hasil pencarian untuk menentukan relevansi dokumen dengan kebutuhannya. Misalnya, pengguna memilih artikel yang benar-benar sesuai dengan topik penelitian yang diinginkan.
- 2) Jika hasil pencarian tidak relevan, seorang penelusur sebaiknya memperbaiki query yang digunakan. Hal ini dapat dilakukan dengan menambahkan kata kunci yang lebih spesifik, menggunakan operator logika seperti AND, OR, atau NOT, serta memanfaatkan filter bawaan sistem pencarian seperti rentang tahun publikasi atau jenis dokumen. Dengan cara ini, hasil pencarian dapat lebih terarah dan sesuai dengan kebutuhan informasi pengguna.
- 3) mahasiswa yang hanya mengetik kata kunci "AI" akan memperoleh hasil pencarian yang sangat luas karena istilah tersebut mencakup berbagai aspek kecerdasan buatan, mulai dari robotika, NLP, hingga sistem pakar. Akibatnya, banyak hasil yang tidak relevan dengan topik spesifik yang diinginkan. Agar hasil lebih relevan, strategi yang sebaiknya digunakan adalah mempersempit kata kunci dengan menambahkan istilah khusus, seperti "deep learning untuk pengolahan citra medis" atau "deep learning untuk tugas akhir". Selain itu, mahasiswa dapat menggunakan tanda kutip ("...") untuk mencari frasa yang tepat, serta menambahkan kata pendukung seperti journal, thesis, atau reference.
- 4) Dalam kasus ketika sistem IR menampilkan 100 dokumen hasil pencarian dan hanya 20 di antaranya yg relevan, maka nilai precision adalah $20/100$ atau sebesar 20%. Precision menggambarkan seberapa banyak dokumen relevan dibandingkan jumlah dokumen yang berhasil diambil. Sedangkan recall mengukur berapa banyak dokumen relevan yg berhasil ditemukan dibandingkan total dokumen relevan yang sebenarnya ada, meskipun jumlah pastinya tidak diketahui dalam kasus ini. Untuk meningkatkan precision, pengguna dapat mempersempit query agar lebih spesifik, sedangkan untuk meningkatkan recall, pengguna dapat menambahkan sinonim/istilah lain yg berkaitan sehingga lebih banyak dokumen relevan terjaring.

5) Perbedaan utama antara text clustering dan text categorization terletak pada cara pengelompokan dokumen. Text clustering dilakukan secara otomatis tanpa adanya kategori yang ditentukan sebelumnya, sehingga dokumen dikelompokkan berdasarkan kemiripan konten (unsupervised learning). Contohnya, sekumpulan artikel secara otomatis dikelompokkan ke dalam topik seperti "machine learning", "computer vision", dan "natural language processing". Sementara itu, text categorization menggunakan kategori yang sudah ditentukan sebelumnya (supervised learning), di mana dokumen akan dimasukkan ke dalam kategori tertentu sesuai model yang telah dilatih. Contohnya adalah pengelompokan email ke dalam folder "spam" atau "inbox".