

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan pengetahuan dan kehidupan manusia sungguh di percepat dengan kemudahan akses terhadap begitu banyak informasi .Proses pencarian dokumen yang sangat banyak, tentunya membutuhkan sebuah teknik khusus dalam mengambil / mencari dokumen yang relevan dengan keinginan pengguna. Terdapat beberapa Teknik yang dapat digunakan, tetapi yang sangat umum digunakan dalam proses pencarian dokumen adalah *Information Retrieval* (IR). IR memberikan informasi tentang subjek yang diperlukan. Data termasuk teks, audio, video, dan beberapa bentuk dokumen lainnya. Tujuan dari penggunaan IR dalam pencarian dokumen salah satunya yaitu menghasilkan dokumen yang relevan dengan permintaan pengguna dalam waktu yang singkat dan tepat [1]

Pada proses pencarian dokumen dapat dilakukan dengan menggunakan sebuah *query*. *Query* adalah sebuah kata kunci yang digunakan untuk menampilkan data dari *database* kemudian akan diproses lebih lanjut. Data tersebut diambil dari tabel-tabel dalam *database* yang memudahkan pengguna dalam mengolah banyak data. Untuk mendapatkan dokumen yang relevan dengan permintaan pengguna, maka *query* harus relevan dengan dokumen yang diinginkan dan hal ini sangat dibutuhkan dalam proses pencarian dokumen [2]

Seiring dengan perkembangan informasi, banyak pihak menyadari bahwa masalah utama telah bergeser dari cara mengakses informasi menjadi memilih informasi yang berguna secara selektif. Usaha untuk memilih informasi ternyata lebih besar dari sekedar mendapatkan akses terhadap informasi. Pemilihan atau penemuan kembali informasi ini tidak mungkin dilakukan secara manual karena kumpulan informasi yang sangat besar dan terus bertambah besar. [3]

Suatu sistem otomatis diperlukan untuk membantu pengguna dalam menemukan informasi. Sistem temu kembali informasi (*information retrieval system*) merupakan sistem yang digunakan untuk menemukan informasi yang relevan dengan kebutuhan dari penggunanya secara otomatis dari suatu koleksi informasi. Permasalahan di tersebut dapat di selesaikan dengan cara membangun sitem peningkatan performansi sistem temu-kembali informasi dengan melakukan perluasan *query* awal dari pengguna secara otomatis. Perluasan *query* tersebut dilakukan berdasarkan umpan-balik relevansi dari pengguna. Umpan balik relevansi dari pengguna tersebut akan dipergunakan oleh sistem untuk melakukan pembobotan ulang ataupun untuk memformulasi ulang *query* awal dari pengguna secara otomatis, sehingga akan menaikkan performansi pada proses pencarian berikutnya.[3]

Pada penelitian sebelumnya Nana Nofiana [4] Pengujian *query expansion* menggunakan metode *ide-dec-hi* pada penambahan 1 kata memiliki nilai rata-rata akurasi untuk penambahan 1 kata bernilai 0,9506 dan 0,9462 untuk penambahan 2 kata. Sedangkan pengujian *query expansion* menggunakan metode *ide-regular* didapatkan nilai rata-rata akurasi untuk penambahan 1 kata sebesar 0,9488 dan 0,9482 untuk tambahan 2 kata. Banyaknya jumlah kata yang ditambahkan pada *query expansion* sehingga mempengaruhi hasil dari nilai *precision* dan akurasi yang dihasilkan.

Berdasarkan uraian-uraian diatas, penerapan metode *Ide-dec-ci* memberikan hasil yang cukup bagus jika dipakai dalam *query* , *ide- dec-hi* lebih baik daripada *ide-regular* dengan kenaikan hingga 0,18% [4][5]. Penelitian ini mencoba untuk membangun suatu sistem Peningkatan Performasi Sistem Temu-Kembali Berita Olahraga . Algoritma yang digunakan adalah algoritma *Ide-dec-ci* karena beberapa hasil penelitian sebelumnya memperlihatkan bahwa algoritma *Ide-dec-ci* teruji mampu meningkatkan kinerja sistem temu kembali informasi. dengan metode pembobotan kata *TF-IDF* dan metode *Improved sqrt cosine similarity* untuk menentukan kemiripan term dengan dokumen.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Permasalahan yang umum terjadi adalah penggunaan *query* yang terlalu singkat dan ambigu. Sehingga *query* tersebut mendapatkan hasil pencarian yang tidak sesuai dengan yang diinginkan .

1.2.2 Usulan Solusi

Solusi dari permasalahan tersebut adalah peningkatan performansi sistem temu-kembali informasi dilakukan menggunakan metode *ide-dec-hi* dengan dataset berita olah raga .

1.2.3 Pertanyaan Penelitian

Seberapa besar rata-rata presisi metode *ide-dec-hi* dalam peningkatan performansi sistem temu-kembali informasi ?

1.3 Tujuan & Mamfaat

1.3.1 Tujuan

Bertujuan untuk mengukur nilai rata-rata presisi metode *ide-dec-ci* dengan dataset Berita olah raga jika diterapkan dalam peningkatan performansi sistem temu-kembali informasi.

1.3.2 Mamfaat

kita dapat menentukan metode yang sesuai untuk peningkatan performansi sistem temu-kembali informasi. Jika dengan penggunaan metode *ide-dec-ci* dengan dataset berita olah raga memiliki nilai rata-rata presisi yang sangat besar, maka dapat di terapkan dalam peningkatan performansi sistem temu-kembali informasi. Dengan harapan dapat menjadi referensi untuk rekan-rekan Mahasiswa lain maupun untuk kebutuhan umum.

1.4 Batasan-Batasan Masalah

1. Hanya menghitung nilai presisi metode *Ide-dec-hi* dengan dataset Berita olahraga dalam peningkatan performansi sistem temu-kembali informasi.
2. Data berita olahraga yang digunakan adalah isi dari artikel berita di situs berita olahraga
3. Hasil akhir dalam penelitian skripsi ini adalah menemukan dokumen penelitian yang cocok dan sesuai dengan kata kunci
4. Data dokumen penelitian yang akan digunakan hanya mengambil isi dari bagian **judul** dan **isi berita**.

1.5 Sistematika Laporan

Sistematika penulisan Proposal ini terdiri dari beberapa bab, yaitu :

BAB I PENDAHULUAN

Pada bab ini membahas tentang latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, serta sistematika penulisan laporan.

BAB II KAJIAN PUSTAKA

Pada bab ini membahas tentang landasan teori dan penelitian sebelumnya.

BAB III METODE USULAN

Pada bab ini membahas tentang metode yang diusulkan (algoritma), arsitektur (Input, Proses, Output), Dataset, tahapan penelitian, serta rencana pengujian.

BAB IV HASIL DAN PEMBAHASAN Pada bab ini membahas tentang proses pembuatan sistem, penjelasan beberapa kode program, dan hasil pengujian

BAB V PENUTUP Pada bab ini membahas tentang kesimpulan dan saran dari penelitian ini.