

# 1. PENDAHULUAN

## 1.1 Latar Belakang

Pada saat sekarang ini, informasi sangat cepat berkembang. Bahkan informasi menjadi suatu kebutuhan yang sangat penting bagi pengguna. Berbagai jenis media dipenuhi beragam informasi salah satunya dalam bentuk dokumen dan dengan berkembangnya teknologi yang tinggi juga sangat mendukung dalam memberikan kemudahan dan kecepatan akses dalam memperolehnya. Namun dibalik kemudahan tersebut muncul sebuah kesulitan baru, pengguna mengalami kesulitan untuk memilih dokumen yang mengandung informasi yang dibutuhkan. Pengguna bisa saja memperoleh dokumen-dokumen tersebut dengan membaca semua dokumen dalam tempat penyimpanan kemudian menyimpan informasi yang relevan dan membuang yang tidak relevan. Pemilihan informasi ini mengalami kesulitan jika dilakukan secara manual karena kumpulan informasi yang sangat besar dan terus bertambah. Untuk mengatasi kesulitan inilah *Information Retrieval(IR)* dikembangkan.

Sistem temu kembali informasi(*information retrieval system*) merupakan sistem yang digunakan untuk menemukan kembali(*retrieve*) dokumen-dokumen yang sesuai dengan kata kunci yang diberikan pengguna dari suatu koleksi informasi[13].

*Relevance feedback* adalah mekanisme untuk mendapatkan *feedback* dari pengguna, dimana pengguna akan memasukkan kata kunci ke dalam sistem, kemudian sistem akan mengembalikan sejumlah dokumen yang dianggap relevan berdasarkan kata kunci yang disebut dengan hasil pencarian awal. *Relevance feedback* merupakan proses formulasi ulang kata kunci awal berdasarkan informasi umpan balik relevansi dari pengguna terhadap dokumen-dokumen hasil pencarian awal. *Feedback* digunakan untuk menentukan kata kunci baru. Metode *Ide dec hi* adalah metode *relevance feedback* yang menggunakan sebuah dokumen tidak relevan pada rangking teratas sebagai *feedback*[11]. Berdasarkan *feedback*, sistem secara otomatis akan menentukan kata kunci baru dan melakukan pencarian berdasarkan kata kunci baru tersebut.

## 1.2 Perumusan Masalah

Permasalahan yang dijadikan sebagai objek penelitian pada tugas akhir ini adalah

1. Bagaimana menerapkan *relevance feedback* pada *information retrieval* menggunakan metode *Ide dec hi*.
2. Bagaimana pengaruh *relevance feedback* pada *information retrieval* menggunakan metode *Ide dec hi* dalam memberikan informasi yang relevan dengan kebutuhan pengguna.
3. Bagaimana penggunaan *relevance feedback* dengan metode *Ide dec hi* dapat meningkatkan performansi *Information Retrieval System(IRS)*.

Batasan masalah dari tugas akhir ini adalah:

1. Koleksi dokumen dan kata kunci yang digunakan adalah dokumen dalam teks bahasa Inggris.

2. Koleksi dokumen menggunakan koleksi dokumen yang diambil dari <ftp://ftp.cs.cornell.edu/SMART> yang terdiri dari tiga koleksi dokumen yaitu koleksi dokumen CRAN, dan MED, dan pada koleksi dokumen tersebut sudah terdapat kumpulan kata kunci beserta *relevance judgement* untuk setiap kata kunci tersebut.
3. Model sistem temu kembali informasi yang digunakan adalah model ruang vektor, dimana dokumen maupun kata kunci direpresentasikan sebagai vektor berdimensi  $n$  dengan  $n$  adalah jumlah kata atau *term* pada kata kunci. *Relevance feedback* pada model ruang vektor dapat dijelaskan sebagai penggeseran vektor kata kunci mendekati vektor dokumen relevan dan menjauhi vektor dokumen tidak relevan[2].
4. Pembobotan kata kunci akan menggunakan *term frequency* , sedangkan pembobotan dokumen akan menggunakan *term frequency inverse document frequency*.
5. Mekanisme penggunaan *relevance feedback* yang digunakan adalah secara manual

### 1.3 Tujuan

Adapun tujuan Tugas Akhir ini adalah :

1. Merancang dan membangun perangkat lunak untuk penerapan *relevance feedback* pada IR menggunakan metode Ide dec hi.
2. Melakukan analisis performansi terhadap hasil dari penerapan metode Ide dec hi, dimana performansi sistem diukur berdasarkan nilai dari *precision*, *recall*, *Interpolated Average Precision(IAP)*.

### 1.4 Metodologi Penyelesaian Masalah

Metodologi penyelesaian yang digunakan dalam menyelesaikan tugas akhir ini adalah:

1. Studi literatur  
Mempelajari sumber-sumber pustaka yang ada, yang dapat dijadikan referensi mengenai *information retrieval* khususnya model ruang vektor, *relevance feedback*, metode Ide dec hi serta sumber-sumber lain yang relevan untuk menunjang penyelesaian tugas akhir ini. Sumber-sumber pustaka dapat berupa buku, *paper*, maupun halaman *web*.
2. Analisis  
Melakukan analisis lebih lanjut mengenai metode Ide dec hi dalam *relevance feedback*. Tujuan analisis adalah untuk mengetahui karakteristik metode Ide dec hi
3. Analisis kebutuhan Perangkat Lunak  
Yaitu melakukan analisis kebutuhan mengenai perangkat lunak yang akan digunakan dan perangkat lunak yang akan dikembangkan, sehingga didapat gambaran umum mengenai perangkat lunak yang akan dikembangkan
4. Perancangan perangkat lunak  
Melakukan perancangan perangkat lunak yang akan dikembangkan. Perancangan yang dilakukan adalah perancangan arsitektur perangkat lunak

5. Implementasi  
Melakukan pembangunan perangkat lunak yang telah dirancang berdasarkan analisis dan perancangan yang telah dilakukan.
6. Pengujian dan analisis hasil.  
Melakukan pengujian perangkat lunak yang telah dikembangkan, dan kemudian menganalisis hasil performansi yang didapatkan. Pengukuran performansi adalah *precision*, *recall*, *Interpolated Average Precision*(IAP) dan waktu yang digunakan selama pemrosesan. Tujuan pengujian adalah untuk mengetahui apakah terjadi peningkatan performansi pada penggunaan metode Ide dec hi dalam *relevance feedback*.

## 1.5 Sistematika Penulisan

Struktur Pembahasan Tugas Akhir ini terdiri dari lima bab dengan disertai lampiran terkait pelaksanaan tugas akhir :

- |         |  |
|---------|--|
| BAB I   | <b>PENDAHULUAN</b><br>Bab ini berisi latar belakang masalah, perumusan masalah, batasan masalah, tujuan pembahasan, metodologi pemecahan masalah dan sistematika penulisan.  |
| BAB II  | <b>DASAR TEORI</b><br>Membahas dasar teori yang berhubungan dengan tugas akhir ini. Secara garis besar dasar teori terdiri dari <i>information retrieval</i> , <i>information retrieval</i> model ruang vektor, <i>relevance feedback</i> dalam model ruang vektor, metode ide dec hi, dan penggunaan metode ide dec hi dalam <i>information retrieval</i> . |
| BAB III | <b>ANALISIS DAN PERANCANGAN PERANGKAT LUNAK</b><br>Bab ini berisi analisis domain masalah, analisis perangkat lunak, dan perancangan perangkat lunak. Pada bab ini akan dilakukan analisis pada pemilihan metode <i>relevance feedback</i> . Berdasarkan hasil analisis tersebut, dilakukan analisis kebutuhan perangkat lunak.                              |
| BAB IV  | <b>IMPLEMENTASI DAN PENGUJIAN</b><br>Bab ini berisi implementasi perangkat lunak berdasarkan hasil perancangan analisis dan perancangan pada bab sebelumnya dan berisi skenario pengujian dan hasil pengujian dari perangkat lunak yang dikembangkan dalam tugas akhir ini   |
| BAB V   | <b>KESIMPULAN &amp; SARAN</b><br>Bab ini akan menjelaskan kesimpulan dan saran sebagai hasil dari analisis dan implementasi Tugas Akhir.   |