BAB I
PENDAHULUAN

I.1 Latar Belakang Masalah

Unit Pelaksana Teknis (UPT) Balai Pengembangan Instrumentasi (BPI) Lembaga Ilmu Pengetahuan Indoensia (LIPI) (selanjutnya dikenal dengan UPT BPI LIPI) Bandung bekerjasama dengan beberapa universitas dalam bentuk praktek kerja lapangan (PKL) untuk membina kompetensi sumber daya manusia yang menunjang sistem pendidikan perguruan tinggi. PKL merupakan salah satu mata kuliah wajib pada program studi strata 1. Kegiatan ini diharapkan dapat menambah pengetahuan dan kemampuan mahasiswa teknik elektro dalam memecahkan kasus-kasus,dan menemukan solusi yang efektif dan tepat.

Berdasarkan hasil wawancara terhadap kepala bagian salah satu unit kerja di UPT Balai Pengembangan Instrumentasi LIPI yaitu Seksi Rekayasa Instrumentasi (Reki) menyatakan bahwa adanya pengumpulan informasi oleh mahasiswa teknik elektro dari berbagai universitas berupa dokumen dengan salah satu tujuannya yaitu untuk peninjauan dan penelusuran terhadap perkembangan-perkembangan yang terjadi di dalam materi pembelajaran yang bersangkutan dengan praktek kerja lapangan (PKL).

Setelah dilakukan penelusuran terhadap kumpulan dokumen teknik elektro dengan menggunakan fungsi pencarian file (*search file*) yang tersedia, maka dapat disimpulkan bahwa pencarian file tersebut adalah pencarian dokumen terhadap judul dokumen sedangkan penelusuran berdasarkan isi dan subab-subab di dalam dokumen sulit dilakukan menggunakan pencarian file dokumen yang tersedia. Maka dari itu dibutuhkannya suatu *Information Retrieval System* yang mampu membantu pencarian dokumen berdasarkan judul dan isi yang dikandung dari dokumen.

Information Retrieval (IR) merupakan suatu metode untuk menemukan kembali data tidak terstruktur yang tersimpan pada sekumpulan dokumen, dan

mampu menyediakan informasi yang sesuai dengan subyek yang dibutuhkan pengguna. *Information Retrieval* (IR) bertujuan untuk memenuhi kebutuhan informasi pengguna dalam mendapatkan semua dokumen yang relevan dengan kebutuhan pengguna dan pada waktu yang sama mendapatkan sesedikit mungkin dokumen yang tidak relevan. Hal tersebut tidak praktis dan efisien karena pada umumnya manusia tidak memiliki banyak waktu untuk membaca seluruh dokumen satu per satu dari sekian banyak dokumen yang ada sehingga diperlukan suatu metode untuk dapat menemukan kembali dokumen yang dibutuhkan. Terdapat beberapa metode dalam *Information Retrieval* (IR) diantaranya adalah metode *Gravitation Based Model, Latent Semantic Model*, *Vector Space Model*, dan *Generalized Vector Space Model* [5].

Generalized Vector Space Model (GVSM) merupakan salah satu model sistem *Information Retrieval* (IR) yang termasuk dalam *Algebraic Model* yang merupakan perluasan dari *Vector Space Model* (VSM) dengan memberikan tambahan jenis informasi, proses yang terjadi pada GVSM terbagi menjadi dua yaitu proses *preprocessing* dan proses yang kedua adalah menghitung relevansi antara kumpulan dokumen yang telah di-*preprocess* dengan *query* yang diinginkan pengguna. Proses *preprocessing* yaitu *reading text* (*.pdf), *tokenization*, *filtration,stemming* dan *parse query*. Banyaknya kemunculan kata dalam kumpulan dokumen yang sesuai dengan *query* akan dihitung. *Query* yang dimasukan oleh pengguna dan dokumen diterjemahkan menjadi vektor-vektor yang sudah dinormalisasi. Kemudian vektor-vektor tersebut dikenakan operasi perkalian titik dan hasilnya menjadi acuan dalam menentukan relevansi masukan pengguna (*query*) terhadap kumpulan dokumen [14].

Berdasarkan permasalahan tersebut maka penelitian memanfaatkan Metode *Generalized Vector Space Model* pada aplikasi *Information Retrieval* untuk membangun sebuah aplikasi pencarian informasi pada kumpulan dokumen teknik elektro berdasarkan *query* yang dimasukan oleh pengguna.

I.2 Perumusan Masalah

Berdasarkan uraian latar belakang masalah yang dikemukakan maka rumusan masalah dapat disimpulkan yaitu bagaimana mengimplementasikan metode *generalized vector space model* (GVSM) pada aplikasi *Information Retrieval* sehingga memperoleh dokumen yang relevan dan sudah terurut sesuai dengan *query* yang dimasukan pengguna.

I.3 Maksud dan Tujuan

Adapun maksud dari penelitian ini adalah untuk melakukan analisis metode *Generalized Vector Space Model* (GVSM) pada *Information Retrieval System* dan membangun Aplikasi *Information Retrieval*.

Sedangkan tujuan penelitian ini diharapkan dapat membantu pengguna dalam pencarian dokumen tertentu (relevan) di dalam kumpulan dokumen teknik elektro di UPT BPI LIPI.

I.4 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

1. Dokumen yang digunakan pada penelitian ini adalah dokumen teks yang berbahasa Indonesia berformat *.pdf.
2. Kumpulan dokumen teknik elektro yang akan diuji pada penelitian ini adalah dokumen mengenai bidang sistem dan sinyal.
3. Aplikasi yang dikembangkan dalam menerapkan metode *Generalized Vector Space Model* berbasis desktop.
4. Algoritma Stemming yang digunakan pada penelitian ini adalah algoritma Nazief dan Adriani.

I.5 Metodologi Penelitian

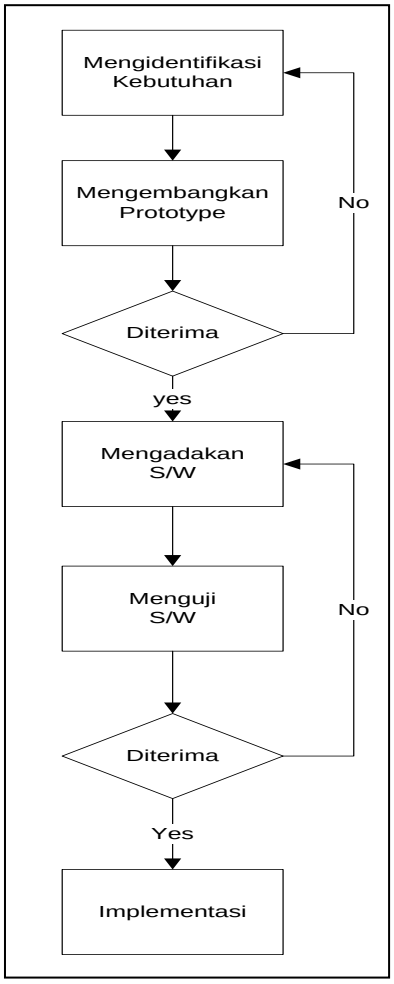
Metodologi penelitian yang digunakan adalah kuantitatif. Adapun metode penelitian yang dilakukan meliputi :

I.5.1 Metode Pengumpulan Data

Pengumpulan data yang dilakukan yaitu observasi ke Unit Pelaksana Teknis Balai Pengembangan Instumentasi Lembaga Ilmu Pengetahuan Indonesia (UPT

BPI LIPI) untuk mendapatkan kumpulan dokumen teknik elektro, studi kepustakaan yaitu dengan mengumpulkan informasi mengenai IR dan metode GVSM dari berbagai sumber literatur berupa buku-buku, skripsi, atau jurnal yang berkaitan untuk dijadikan referensi.

I.5.2 Metode Pengembangan Sistem



Gambar I.1 Diagram Prototype Model [11]

Metodologi pengembangan sistem pada penelitian ini menggunakan Metodologi *Prototype*, seperti yang dijelaskan pada gambar diatas yang meliputi :

1. Mengidentifikasi kebutuhan sistem meliputi bagaimana sistem membaca dokumen berformat *.pdf, melakukan pemecahan kalimat menjadi kata tunggal, melakukan proses filtrasi kata, melakukan *stemming* kata,

melakukan pencarian dan pengurutan dokumen menggunaka metode GVSM.

2. Mengembangkan *prototype*, meliputi pengembangan *prototype tokenization, filtration, stemming* dan *retrieval process.*
3. Mengadakan *software*.
4. Menguji *software*.
5. Implementasi.

I.6 SISTEMATIKA PENULISAN

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini menguraikan tentang latar belakang permasalahan, mencoba merumuskan inti permasalahan yang dihadapi, menentukan tujuan dan kegunaan penelitian, yang kemudian diikuti dengan pembatasan masalah, asumsi, serta sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan topic penelitian yang dilakukan.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini menganalisis masalah dari objek penelitian untuk mengetahui hal atau masalah apa yang timbul dan mencoba memecahkan permasalahan tersebut dengan memperangkat lunakan perangkat-perangkat yang digunakan hal-hal yang berguna dalam proses analisis permasalahan serta tinjauan terhadap penelitian-penelitian serupa yang telah pernah dilakukan sebelumnya. Sedangkan perancangan adalah suatau tahap pada penelitian dimana perangkat yang digunakan ditentukan, mengidentifikasi data yang diperlukan dan cara pengumpulannya, serta menentukan cara kerja perangkat yang digunakan terhadap objek penelitian yang dirumuskan dari awal persiapan hingga penarikan kesimpulan.

BAB IV PERANCANGAN DAN IMPLEMENTASI

Bab ini menjelaskan proses dimana analisis dan perancangan yang telah dipersiapkan untuk selanjutnya diimplementasikan menjadi perangkat lunak dan dilakukan pengujian terhadap perangkat lunak tersebut.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran yang sudah diperoleh dari hasil penulisan penelitian.