

BAB I

PERSYARATAN PRODUK

1.1 PENDAHULUAN

Pada saat kita melakukan pencarian melalui *search engine* (*google.com*, *yahoo*, dsb), kita bisa mendapatkan beberapa hasil, yang berupa dokumen - dokumen yang sama atau hampir sesuai dengan kata atau *query* yang kita masukkan. Demikian pula jika kita melakukan pencarian dalam aplikasi sistem informasi, seperti halnya sistem pencarian dalam perpustakaan. Aplikasi yang dibuat adalah aplikasi yang menggunakan algoritma IR (*Information Retrieval*) dengan metode *Generalized Vector Space. Information Retrieval (IR) system* merupakan suatu sistem yang membantu pengguna dalam mencari informasi di dalam dokumen - dokumen. Beberapa sistem yang menggunakan IR *system* adalah aplikasi *search engine*, seperti *google.com* dan aplikasi sistem informasi, seperti perpustakaan.

Algoritma yang dibahas menggunakan konsep ruang vektor. Masukan dari pengguna dan dokumen – dokumen diterjemahkan menjadi vektor- vektor kemudian vektor – vektor tersebut dikenakan operasi perkalian titik dan hasilnya menjadi acuan dalam menentukan relevansi masukan pengguna (*query*) dengan dokumen – dokumen.

Aplikasi ini juga didukung dengan adanya *stemmer*, yaitu suatu program yang digunakan untuk memecah kata, artinya mengambil kata dasar. Program ini dibuat untuk mempermudah perhitungan kesamaan kata pada algoritma IR.

Aplikasi ini diimplementasikan dengan menggunakan bahasa JAVA. Apabila seorang programmer membuat software dengan berbasis bahasa Java, software tersebut dapat dijalankan sekaligus di berbagai sistem operasi, termasuk Windows dan Linux yang saat ini tengah bersaing merebut simpati para pengguna komputer.

1.1.1. Tujuan

Membuat aplikasi pencarian yang diharapkan dapat memudahkan *user* untuk mencari dan mendapatkan informasi sesuai dengan *query*. Aplikasi ini menerapkan *Generalized Vector Space Model* yang digunakan untuk membantu aplikasi dalam mencari informasi di dalam dokumen – dokumen berdasarkan *query* yang dimasukkan oleh *user*. Aplikasi ini juga menggunakan *stemmer*, yang digunakan untuk memecah kata, yaitu mengambil kata dasar, dengan memisahkan kata depan, kata penghubung, dan imbuhan.

1.1.2. Ruang Lingkup Proyek

Dalam tugas akhir ini, aplikasi pencarian akan dirancang dan dibuat dengan memanfaatkan algoritma *Information Retrieval*.

Kemampuan aplikasi yang akan dikembangkan tersebut adalah

- Menampilkan dokumen - dokumen yang dibutuhkan berdasarkan informasi atau *query* yang dimasukan oleh *user*.
- Melakukan pemecahan kata. Apabila *query* dimasukan, kata-kata yang bukan merupakan kata dasar, akan dipecah menjadi sebuah kata dasar. Dalam hal ini, aplikasi *stemmer* membuang kata depan, imbuhan, kata penghubung,dll dan hanya mengambil kata dasarnya saja.
- Menampilkan dokumen - dokumen berdasarkan kesamaan nilai yang didapat dari perhitungan hasil kali titik vektor – vektor, dan menjadi acuan untuk menampilkan urutan dokumen sesuai dengan *query* yang telah diinputkan.

1.1.3. Definisi, Akronim dan Singkatan

- ♦ *Query* : inputan atau masukan dari *user*
- ♦ *Index term* : kata yang dipilih dan mengacu pada isi dokumen

- ◆ *Recall* : ukuran performansi *Information Retrieval* berdasarkan banyak dokumen yang relevan dan ter-*retrieved* dibagi banyak semua dokumen relevan di koleksi dokumen
- ◆ *Precision* : ukuran performansi *Information Retrieval* berdasarkan banyak dokumen yang relevan dan ter-*retrieved* dibagi semua dokumen ter-*retrieved*
- ◆ *Stemming* : teknik membuang awalan dan akhiran dari sebuah kata untuk memperoleh kata dasar
- ◆ *User* : pengguna
- ◆ JAVA : nama salah satu bahasa pemrograman
- ◆ *input* : masukan
- ◆ *output* : keluaran / hasil
- ◆ *minterm* : mewakili pola *index term* yang mungkin muncul dalam suatu dokumen

1.1.4. Overview Laporan

Laporan ini berisi hal-hal yang berkaitan dengan pembuatan sebuah produk yaitu: tujuan pembuatan produk, gambaran antarmuka produk, fungsi produk, dsb. Selain itu juga dibahas mengenai fitur-fitur produk, batasan desain, rancangan antarmuka produk, serta rancangan *coding* yang telah disusun.

Pada bab I akan dijelaskan mengenai tujuan, ruang lingkup proyek, definisi, akronim dan singkatan, *overview* laporan, selain itu terdapat pula gambaran keseluruhan dari produk antara lain perspektif produk, fungsi produk, karakteristik pengguna, batasan-batasan, asumsi dan ketergantungan serta penundaan persyaratan.

Pada bab II merupakan gambaran keseluruhan dari spesifikasi produk yang mencakup antarmuka pengguna, antarmuka perangkat keras, antarmuka

perangkat lunak, antarmuka komunikasi, fitur yang ada dalam aplikasi beserta penjelasannya.

Pada bab III akan berisi identifikasi kebutuhan sistem, rancangan sistem baru, desain perangkat lunak secara keseluruhan, desain arsitektur perangkat lunak, komponen perangkat lunak, konsep eksekusi, dan desain antarmuka aplikasi.

Pada bab IV akan berisi perencanaan tahap implementasi yang meliputi pembagian *Class* implementasi, keterkaitan antar *Class*, selain itu juga akan membahas perjalanan tahap implementasi meliputi jenis implementasi, *debugging* dan membahas ulasan realisasi fungsionalitas, dan ulasan realisasi user interface desain aplikasi.

Pada bab V akan berisi rencana pengujian sistem terimplementasi, test case, uji fungsionalitas *Class*, perjalanan metodologi pengujian, *white box*, *black box*, survey/wawancara, serta ulasan hasil evaluasi yang telah dilaksanakan.

Pada bab VI akan berisi kesimpulan dan saran yang meliputi keterkaitan antara kesimpulan dengan hasil evaluasi, keterkaitan antara saran dengan hasil evaluasi serta akan membahas rencana perbaikan/implementasi terhadap saran yang diberikan.

1.2 GAMBARAN KESELURUHAN

Aplikasi ini dirancang untuk membantu *user* dalam mencari informasi di dalam sekumpulan dokumen. Dengan menggunakan *Generalized Vector Space Model*, yang dapat mempermudah pencarian dalam sistem, dan dapat menilai kesamaan antara *query* dan dokumen. Hasil yang ditampilkan berdasarkan urutan kesamaan antara *query* dan dokumen.

1.2.1. Perspektif Produk

Aplikasi ini dibuat untuk membantu memudahkan pencarian informasi dalam kumpulan dokumen, dengan menggunakan algoritma *Generalized Vector Space Model* dengan metode vektor.

Term adalah query yang telah dikenai operasi teks, seperti pemotongan kata depan, kata penghubung, dan imbuhan. Dokumen - dokumen diurutkan dengan menggunakan rumus kesamaan. Nilai kesamaan didapat dari perkalian vektor antara *query* dan masing – masing dokumen.

1.2.2. Fungsi Produk

Aplikasi ini digunakan untuk memudahkan semua pengguna dalam mencari informasi dalam koleksi dokumen. Berbagai informasi bisa didapat oleh pengguna dengan memasukkan kata kunci informasi yang ingin diperoleh. Kata kunci yang dimasukkan dapat berupa kata benda atau kata sifat.

1.2.3. Karakteristik Pengguna

Untuk menjalankan aplikasi ini diperlukan *user* yang mengerti bagaimana mengoperasikan komputer, memiliki pengalaman dalam melakukan *input* data. Input data adalah memasukkan kata dalam sebuah *textbox* untuk mendapatkan informasi yang dibutuhkan dari sistem.

1.2.4. Batasan - Batasan

- Aplikasi yang dibuat adalah aplikasi pencarian informasi dalam koleksi dokumen dengan menggunakan metode vektor. Metode vektor yang digunakan adalah *Generalized Vector Space Model*.
- Aplikasi ini menggunakan Bahasa Indonesia.

- Koleksi dokumen yang digunakan untuk pengujian berupa dokumen teks tanpa format. Hal ini dimaksudkan untuk menghilangkan kebutuhan untuk mempelajari format dokumen seperti *Microsoft Word Document Format*, *Adobe Portable Document Format*, dan lain-lain.

Software yang menunjang pembuatan program:

- *Microsoft Windows XP Professional SP 1*
- *JDK1.6*
- *Netbeans 5.5*

1.2.5. Asumsi dan Ketergantungan

Untuk dapat menjalankan aplikasi ini diperlukan komputer yang mempunyai Java SE Runtime Environment (JRE) 1.6