

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dengan semakin pesatnya perkembangan dalam penggunaan teknologi komputer baik di perusahaan maupun di bidang pendidikan, maka semakin banyak pula dokumen-dokumen yang berbentuk digital. Dokumen dengan jumlah yang sangat banyak tersebut tentunya memerlukan suatu mekanisme agar pengguna dapat melakukan pencarian atau mendapatkan kembali dokumen yang sesuai dengan kebutuhan secara cepat dan mudah. Tanpa hal tersebut maka informasi yang didapatkan tidak sesuai dengan informasi yang dicari.

Salah satu cara untuk mendapatkan kembali informasi dari dokumen yang diinginkan dalam suatu kumpulan dokumen yang jumlahnya sangat banyak adalah dengan melakukan perankingan. *Ranking* merupakan salah satu cabang dari *information retrieval*. *Ranking* biasanya digunakan dalam pengambilan dokumen, penyaringan informasi, penempatan iklan *online*, dan lain-lain.

Salah satu metode untuk melakukan perankingan dokumen adalah dengan menggunakan metode *Vector Space Model* (VSM). *Vector Space Model* adalah suatu model yang digunakan untuk mengukur kemiripan antara suatu dokumen dengan suatu *query*. VSM menggunakan konsep yang terdapat pada aljabar linear yaitu ruang vektor. Dengan konsep ini kita bisa mengetahui kedekatan antara dua buah vektor dengan cara menghitung besarnya sudut yang terbentuk antara dua vektor dan kemudian diurutkan dari data yang memiliki besar sudut yang terkecil hingga yang terbesar yang menandakan urutan data hasil *ranking* dari yang paling relevan hingga yang tidak relevan.

Vector space model dapat diterapkan pada pencarian artikel-artikel singkat maupun berita-berita yang jumlahnya sangat banyak. Salah satu instansi yang sering diliput dalam berita-berita harian adalah Kantor Imigrasi. Kantor Imigrasi merupakan salah satu kantor pelayanan publik yang bekerja dibawah naungan Kementerian Hukum dan Ham (Kemenkumham) Indonesia. Kantor imigrasi tersebar di seluruh wilayah di Indonesia. Di wilayah Nusa Tenggara Timur (NTT), terdapat 3 kantor cabang Imigrasi yaitu Kantor Imigrasi Kelas I Kupang,

Kantor Imigrasi Klas II Maumere, dan Kantor Imigrasi Kelas II Atambua. Sebagai salah satu kantor pelayanan publik, kantor imigrasi tentunya memiliki banyak kegiatan yang telah dilakukan. Kegiatan-kegiatan tersebut biasanya didokumentasikan dan dipublikasikan dalam sebuah berita. Berita-berita tersebut ada yang dipublikasikan melalui *website* resmi kantor imigrasi, ada juga yang dipublikasikan oleh media massa yang ada.

Oleh karena berita-berita tentang Kantor Imigrasi sangat banyak, maka dapat diterapkan metode *vector space model* dalam pencarian berita-berita tentang keimigrasian untuk dapat divisualisasikan.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang ada, maka masalah yang diangkat adalah bagaimana memvisualisasikan metode *vector space model* untuk pencarian berita-berita keimigrasian di wilayah Nusa Tenggara Timur.

1.3. Tujuan Penelitian

Tujuan penelitian tugas akhir ini adalah memvisualisasikan *metode Vector Space Model* dalam pencarian berita-berita keimigrasian di wilayah Nusa Tenggara Timur (NTT).

1.4. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Untuk membantu dan mempermudah pengguna dalam mencari dokumen yang diinginkan secara tepat.
2. Menghasilkan informasi yang lebih relevan dengan hasil ketepatan (*precision*) tinggi dan perolehan (*recall*) rendah.

1.5. Batasan Masalah

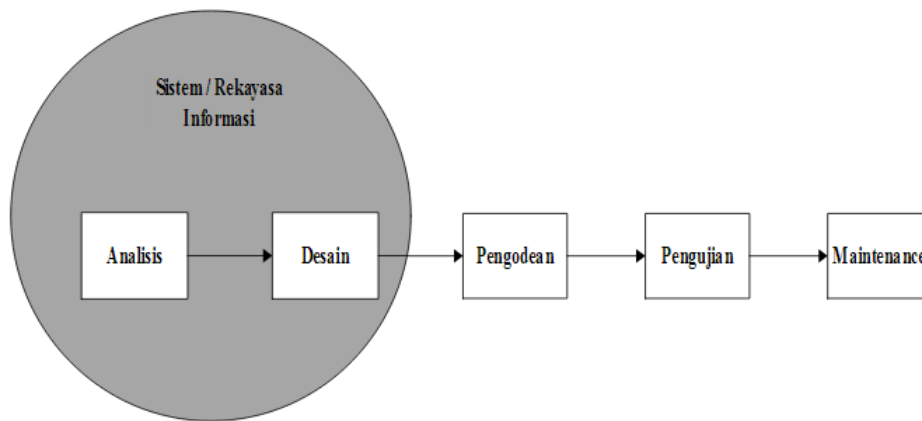
Batasan masalah dalam penelitian ini adalah :

1. Dokumen yang digunakan dalam penelitian ini adalah berita-berita tentang Kantor Imigrasi yang ada di wilayah Nusa Tenggara Timur.
2. Pembobotan yang digunakan adalah pembobotan *Term Frequency – Inverse Document Frequency* (TF-IDF).

1.6. Metodologi Penelitian

Dalam aplikasinya, pengembangan suatu produk memerlukan suatu model pengembangan sedangkan tahapan-tahapan dalam pengembangan perangkat lunak sering disebut dengan SDLC (*Software Development Life Cycle*). Model yang digunakan dalam sistem temu kembali informasi ini adalah SDLC model *waterfall*.

Model ini digambarkan seperti pada gambar berikut :



Gambar 1.1 Ilustrasi model *waterfall* [1]

1.6.1. Analisis

Pada tahap analisis dilakukan analisis terhadap konsep *information retrieval*, menganalisis metode atau algoritma yang digunakan, bahasa pemrograman yang digunakan, serta data yang cocok digunakan untuk melakukan pencarian dengan menggunakan metode *vector space model*. Pada tahap ini juga dikumpulkan data-data yang digunakan dalam sistem yaitu berita-berita tentang keimigrasian yang ada di wilayah Nusa Tenggara Timur. Berita-berita ini akan digunakan dalam proses pencarian dengan menggunakan metode *vector space model*.

Selain data-data, dikumpulkan juga bahan-bahan referensi yang menunjang proses penelitian yaitu yang berhubungan dengan pengenalan *information retrieval*, penentuan indeks dan pembobotan *tf-idf*, dan visualisasi yang digunakan dalam metode *vector space model*.

1.6.2. Desain

Pada tahap ini dilakukan perancangan interface *input* dan *output* untuk sistem yang digunakan dalam uji coba pencarian dengan metode *vector space model*. Selain itu juga dilakukan perancangan *database* yang digunakan didalam sistem.

1.6.3. Pengodean

Pada tahap ini dibuat kode program yang akan digunakan untuk melakukan uji coba pencarian berita menggunakan metode *vector space model*.

1.6.4. Pengujian

Pada tahap ini dilakukan uji coba sistem yang dibuat untuk melakukan pencarian terhadap berita yang telah diinput ke dalam *database*.

1.6.5. Maintenance (Perawatan)

Tidak menutup kemungkinan sebuah perangkat lunak mengalami perubahan ketika sudah dikirimkan kepada *user*. Perubahan bisa terjadi karena adanya kesalahan yang muncul karena tidak terdeteksi pada tahap pengujian. Tahap pemeliharaan dapat mengulangi proses pengembangan mulai dari analisis spesifikasi untuk perubahan perangkat lunak yang sudah ada, tapi tidak untuk membuat perangkat lunak yang baru.

1.7. Sistematika Penulisan

Sistematika penulisan tugas akhir ini merupakan gambaran umum tentang seluruh isi penulisan yang terdiri atas 6 (enam) bab, sebagai berikut:

BAB I PENDAHULUAN

Dalam bab ini akan dibahas mengenai latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Dalam bab ini akan dibahas tentang penelitian terdahulu dan teori-teori tentang *information retrieval*, *vector space model*, *indexing*, dan pembobotan tf-idf.

BAB III PERANCANGAN SISTEM

Pada bab ini akan dibahas tentang tempat penelitian, analisis sistem dan perancangan sistem.

BAB IV IMPLEMENTASI SISTEM

Dalam bab ini membahas tentang implementasi sistem sesuai dengan hasil perancangan pada bab sebelumnya.

BAB V PENGUJIAN DAN ANALISIS HASIL

Pada tahap bab ini akan dibahas tentang pengujian hasil sistem yang telah dibangun.

BAB VI PENUTUP

Pada bab ini berisi kesimpulan yang diperoleh dari hasil pengembangan sistem serta saran terhadap sistem untuk perkembangan selanjutnya.