

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi memberikan kemudahan dalam memperoleh informasi-informasi yang dibutuhkan. Jumlah permintaan terhadap buku selalu mengalami peningkatan setiap tahunnya, sehingga menyebabkan banyaknya beredar buku-buku baru sesuai dengan kebutuhan masyarakat dan perkembangan ilmu pengetahuan dan teknologi, tetapi untuk mencari buku yang diharapkan atau dibutuhkan tidak semudah yang dikira sebagai contoh banyak sekolah yang melakukan perubahan pada penggunaan pedoman atau panduan dalam proses belajar mengajar agar tidak “tertinggal” dengan ilmu pengetahuan dan teknologi yang ada. Hal ini mendorong para pelajar untuk mencari buku sesuai dengan apa yang dibutuhkan.

Sistem pencarian kemiripan buku berdasarkan keyword terkadang memberikan hasil yang tidak sesuai dalam hal pencarian buku yang sejenis, dikarenakan dalam sebuah pencarian terdapat pilihan kata yang digunakan pengguna dalam query sering kali berbeda dengan kata yang digunakan penulis, bahkan ketika beberapa user memiliki kebutuhan informasi yang sama, mereka jarang menggunakan pilihan kata yang sama untuk menggambarannya dalam query.

Aplikasi pencarian buku berdasarkan keyword ini dibangun dengan metode Vector Space Model dengan menggunakan pengukuran dalam menentukan nilai kemiripan antar dokumen yang diawali dari deskripsi buku dimana parameter yang digunakan adalah kecocokan keyword buku yang sejenis. *Cosine Similarity* menggunakan dua vektor yang mempresentasikan pencarian buku dimana nilai sudut kosinus dari kedua vektor tersebut adalah nilai kemiripan dari dua kata tersebut. Batasan nilai yang dihasilkan mulai dari 0 sampai dengan 1 atau biasa disebut *Binary Term*.

Aplikasi pencarian buku pada toko online dibangun dengan menggunakan suatu ilmu yang mengacu pada pengolahan data teks. *Text Mining* adalah

proses menganalisa teks untuk mengekstrak informasi yang berguna untuk tujuan tertentu. Tahap-tahap pada *text mining* secara umum terdiri dari *case folding, tokenizing, filtering, stemming, tagging, analyzing*. Terdapat beberapa metode dalam melakukan tahapan terakhir *text mining* yaitu *analyzing* atau biasa disebut analisa data untuk menentukan kemiripan dua objek yang sedang dibandingkan. *Cosine Similarity* adalah salah satu metode dalam menentukan nilai kemiripan antar dua objek. Salah satu contoh penerapan adalah penentuan kemiripan pada sidik jari manusia. *Cosine Similarity* dapat diterapkan dalam menentukan nilai kemiripan pada dua berkas dokumen teks. Parameter yang digunakan adalah jumlah kata-kata pada dua dokumen teks yang dibandingkan. *Cosine Similarity* menggunakan dua vektor yang mempresentasikan dua dokumen teks dimana nilai sudut kosinus dari kedua vektor tersebut adalah nilai kemiripan dari dua dokumen teks tersebut. Batasan nilai yang dihasilkan mulai dari 0 sampai dengan 1 atau biasa disebut *Binary Term*.

Penelitian ini bertujuan dapat memberikan hasil penerapan metode Vector Space Model untuk mendeteksi kemiripan buku yang sejenis dan hasil pencarian terkait query yang dicari oleh pengguna. Serta bermanfaat bagi pengguna dalam memberikan kemudahan mencari kemiripan buku yang sejenis dengan praktis dan sesuai. Dengan melakukan beberapa penerapan metode yang berbeda dalam tahapan *text mining*. Pada sistem ini akan diaplikasikan *text mining* untuk tahap *preprocessing* dan *cosine similarity* sebagai *string matching*. Dengan mengekstrak dokumen yang berisi teks kemudian dilakukan suatu perhitungan dengan metode *cosine similarity* agar menghasilkan suatu keluaran *similarity* dokumen dari jumlah *string* kata dengan batasan *threshold* tertentu untuk mengetahui kemiripan dokumen tersebut. Selanjutnya akan dibahas bagaimana *text mining* dan *cosine similarity* bekerja dan sekaligus implementasi pada sebuah aplikasi dalam mendeteksi kemiripan.

1.2 Rumusan Masalah

1. Bagaimana mencari referensi buku sejenis yang mempunyai kemiripan dengan buku yang dicari?

1.3 Tujuan Penelitian

Adapun penelitian ini bertujuan dapat memberikan referensi buku sejenis yang mempunyai kemiripan dengan buku yang dicari dengan metode vector space model.

1.4 Batasan Masalah

1. Pada sistem ini teks yang dianalisa berdasarkan isi deskripsi buku

1.5 Manfaat Penelitian

Adapun manfaat yang ingin dicapai dari penelitian ini adalah :

Bagi pengguna aplikasi dapat memberikan kemudahan dalam mencari buku yang dibutuhkan dan membantu pengguna dalam menganalisa kemiripan kata berdasarkan keyword melalui nilai similarity dan dibuktikan dengan letak kemiripan yang sama antar buku.

1.6 Metodologi Penelitian

Metodologi dalam penelitian skripsi ini sebagai berikut:

1. Studi Literatur

Dalam tahap ini dilakukan pembelajaran sejumlah literatur dalam hal konsep dan teknologi yang akan digunakan. Melalui suatu literature meliputi referensi paper, buku dan dokumentasi internet.

2. Tahap Pengumpulan Data

Pada Tahap ini dilakukan pengumpulan data buku yang akan dibandingkan.

3. Tahap Preprocessing Data

Dalam Tahap ini data yang sudah diperoleh akan diproses untuk dipilah kembali sesuai kebutuhan sistem yang nantinya akan dijadikan sebagai data yang akan digunakan.

4. Analisis Sistem

Pada tahap ini dilakukan sebuah analisis sistem untuk menentukan konsep dan teknologi yang digunakan untuk menyelesaikan permasalahan terkait.

5. Implementasi

Dalam tahap ini dilakukan sebuah proses penerjemah dari perancangan sistem yang telah di analisis sebelumnya untuk dilakukan pengkodean menggunakan Bahasa pemrograman PHP dan Mysql.

6. Pengujian

Tahap selanjutnya merupakan uji coba terhadap program yang telah dibuat untuk mengetahui sejauh mana kinerja sistem dalam melakukan perhitungan kemiripan kata buku yang dibandingkan.

7. Tahap Penyusunan Laporan

Melakukan penyusunan laporan berdasarkan sistematika penulisan.

1.7 Sistematika Penulisan

Adapun sistematika penulisan pada tugas akhir ini sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, metodologi penelitian, sistematika penulisan dan jadwal kegiatan yang direncanakan.

BAB II LANDASAN TEORI

Bab ini membahas tentang teori serta konsep dari permasalahan terkait kemiripan kata buku menggunakan *Text Mining* dengan metode *Cosine Similarity*.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini menjelaskan tentang penganalisaan kebutuhan dan perancangan sistem. Meliputi dari analisa kebutuhan sistem dan desain perancangan sistem

