

BAB I

PENDAHULUAN

1.1 Latar Belakang

Buku merupakan jendela dunia. Dengan membaca buku, banyak ilmu yang dapat diperoleh. Salah satu manfaat buku adalah menunjang proses belajar mengajar, baik itu di sekolah maupun perguruan tinggi (Setiawan, 2015). Begitu pentingnya buku, banyak penulis maupun mahasiswa tertarik untuk mengeluarkan pendapat dan berbagi ilmu dalam bentuk buku. Mahasiswa sudah mulai cenderung menyimpan informasi secara digital dan *E-book*, karena lebih mudah dalam penyimpanan dan cepat. Semakin banyak buku yang masuk tidak menutup kemungkinan memiliki kemiripan dalam judul buku. Judul buku yang memiliki kemiripan belum tentu berada dalam *class* golongan yang sama, sehingga petugas perpustakaan dapat melakukan kesalahan dalam pengelompokan buku tersebut. Untuk membantu memecahkan masalah pencarian tersebut, tentunya diperlukan sebuah solusi, dimana solusi tersebut nantinya diharapkan bisa membantu *user* dalam melakukan pencarian buku yang dibutuhkan.

Pengelompokan judul buku merupakan bagaimana data-data koleksi tersebut diolah untuk dikelompokkan yang terorganisir pada setiap penggunaanya dalam hal ini adalah peminjam dari berbagai jurusan maka dapat diketahui kelompok literatur apa saja yang paling sering dipinjam oleh mahasiswa dan dibentuk berdasarkan dokumen yang sesuai dengan kategori yang dibuat oleh pihak perpustakaan (Andyni, 2013). Permasalahan yang lainnya terjadi adalah buku yang memiliki kemiripan judul, sehingga buku tersebut bagaimana data-data judul buku tersebut diolah untuk dikelompokkan. Permasalahan yang timbul antara lain, dimana judul buku yang tersedia secara manual tersebut menjadi hilang, karena terlalu banyak berkas judul buku yang tersimpan manual dalam satu dokumen terutama yang ada di perpustakaan. Pengaruh dari permasalahan tersebut adalah proses mencari judul buku yang dibutuhkan menjadi tidak efisien sehingga membutuhkan waktu yang lama dan sulit. Masalah lain akan terjadi ketika setiap dokumen judul tersebut ingin dikelompokkan ke dalam kategori judul tertentu, karena harus dilihat, dibaca, dan dipahami isi tiap dokumen judul dalam data.

Setelah selesai membaca seluruh isi data, barulah bisa ditentukan kategori-kategori bagi dokumen dan membagi dokumen dalam kategori judul tersebut.

Untuk mengatasi permasalahan tersebut, maka perlu upaya sedini mungkin untuk membantu memecahkan masalah pencarian tersebut, tentunya diperlukan sebuah solusi, dimana solusi tersebut diharapkan bisa membantu *user* dalam melakukan mengelompokkan data dokumen pencarian buku yang dibutuhkan. Dengan adanya pengelompokkan dokumen ini, mahasiswa hanya perlu mengetikkan judul buku, lalu secara otomatis pencarian judul buku akan memberikan dokumen-dokumen yang mirip dengan judul yang dimasukkan agar tidak terjadi pencarian dengan judul yang sama. Hal ini dikarenakan di dalam pencarian tersebut dapat mengelompokkan data-data berdasarkan kesamaan antar dokumen atau berdasarkan kelompoknya. Dengan adanya sebuah sistem yang berfungsi untuk mengelompokkan dokumen pencarian judul buku ini, dapat mempermudah mahasiswa dalam pencarian judul buku dengan tingkat kemiripan yang paling sesuai dengan *query* yang diberikan oleh *user* dapat membantu administrator menentukan kategori yang tepat untuk setiap buku dan mempermudah dan mempercepat proses pencarian buku yang sesuai dengan kebutuhan.

Maka dari itu diperlukan sistem pencarian buku yang dapat dilakukan secara otomatis menggunakan suatu metode yang diterapkan dalam aplikasi yaitu *Text Mining Clustering* dengan menggunakan *Vector Space Model*. Pada *Text Mining* metode *Clustering* digunakan untuk mengelompokkan data tekstual berdasarkan kesamaan konten yang dimiliki ke dalam beberapa klaster, sehingga didalam setiap klaster akan berisi data tekstual dengan konten semirip mungkin (Feldman, 2007). Dengan keunggulan tersebut, *Text Mining* telah menjadi semakin efektif seiring dengan daya komputasi yang meningkat. *Text Mining* telah sukses dalam menemukan berbagai pola dan koneksi yang tidak terlihat oleh sekilas mata, seperti mengukur tingkat kebahagiaan melalui kata-kata dalam *twitter*, mengetahui hubungan senyawa tertentu dengan enzim tertentu, dan sebagainya (Belsky, 2012). *Text Mining* memiliki beberapa metode, salah satunya adalah *Vector Space Model*. Pada *Vector Space Model* masukan dari pengguna (*query*) dan kumpulan dokumen diterjemahkan menjadi vektor-vektor, kemudian vektor-vektor

tersebut dikenakan operasi perkalian titik dan hasilnya menjadi acuan dalam menentukan relevansi masukan pengguna (*query*) terhadap kumpulan dokumen (Witanti, et al., 2016). Clustering merupakan salah satu metode Data Mining yang bersifat unsupervised (tidak terawasi). Terdapat dua jenis data clustering yang sering digunakan untuk pengelompokan data yaitu hirarki data clustering dan non-hirarki data clustering (Redjeki, 2011).

Para peneliti telah mencoba menerapkan *Text Mining* di dalam berbagai kasus. Dwi Smaradahana Indraloka, melakukan penelitian tentang Penerapan *Text Mining* untuk Melakukan *Clustering Data Tweet Shopee* Indonesia (Indraloka, 2017). Sri Redjeki melakukan penelitian tentang *Clustering Terhadap Indeks Prestasi Mahasiswa STMIK Akakom Menggunakan K-Means* (Redjeki, 2011). Erene Gernaria Sihombing melakukan penelitian tentang *Klasifikasi Data Mining Pada Rumah Tangga Menurut Provinsi Dan Status Kepemilikan Rumah Kontrak/Sewa Menggunakan K-Means Clustering Method* (Sihombing, 2017). Dalam penelitian ini, diharapkan *Text Mining* juga dapat melakukan klasifikasi judul buku merancang dan membuat suatu sistem komputerisasi yang dapat mengklasifikasi buku sesuai dengan kategorinya serta membuat suatu sistem pencarian buku dari hasil klasifikasi buku tersebut.

Berdasarkan latar belakang yang telah dikembangkan diatas maka dilakukan penelitian yang berjudul ***“Text Mining Sebagai Rekomendasi Pencarian Judul Buku di Perpustakaan Universitas Andalas, Padang”***.

1.2 Rumusan Masalah

Berdasarkan latar belakang maka rumusan masalah pada tugas akhir ini adalah sebagai berikut :

1. Bagaimana implementasi proses pencarian judul buku yang dilakukan dengan metode *Text Mining* dan *Vector Space Model* dengan pembobotan *Term Frequency – Inverse Document Frequency* (TF-IDF).
2. Bagaimana proses *cluster* dengan mengukur akurasi pencarian dokumen dengan metode *Cosine Similarity*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini dapat ditentukan berdasarkan rumusan masalah yang telah dijelaskan sebelumnya yaitu sebagai berikut :

1. Dokumen teks uji dalam penelitian ini berupa dokumen teks judul buku.
2. Metode yang digunakan pada *Text Mining* yaitu metode *Vector Space Model*
3. Data yang digunakan beberapa data sampel dari judul buku yang terdiri dari masing-masing fakultas Universitas Andalas Padang.
4. Aplikasi dibangun berbasis web dengan menggunakan bahasa pemrograman PHP.
5. Database server menggunakan SQL Server

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah bagaimana membuat sebuah sistem pencarian judul buku di perpustakaan Universitas Andalas Padang, yang lebih cepat serta mampu mengklasifikasikan judul buku tersebut sesuai dengan letak fakultas yang ada di Universitas Andalas, Padang.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Menambah pengetahuan penulis mengenai aplikasi dari metode *Vector Space Model* untuk klasifikasi suatu kasus.
2. Memudahkan mahasiswa menemukan buku yang paling prioritas atau relevan dengan pencarian yang dilakukan oleh sistem, sehingga dapat menentukan klasifikasi terhadap buku tersebut.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan ini adalah sebagai berikut:

a. BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan.

b. BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan tentang kajian-kajian literatur yang digunakan untuk menunjang penelitian tugas akhir ini yang terdiri dari *Text Mining*, *Vector Space Model*, PHP dan SQL Server.

c. **BAB III METODOLOGI PENELITIAN**

Pada bab ini membahas langkah-langkah yang dilaksanakan dalam proses penelitian, yaitu identifikasi masalah, studi literatur, pengumpulan data, olah data, proses klasifikasi judul buku dengan menggunakan metode *Text Mining* dan *Vector Space Model*, pengujian, implementasi *Interface* dan kesimpulan dan saran.

d. **BAB IV ANALISIS**

Bab ini menjelaskan tentang implementasi dan pengujian yang dilakukan pada penelitian ini.

e. **BAB V PENUTUP**

Bab ini berisi tentang kesimpulan yang didapatkan pada penelitian ini dan saran untuk pengembangan penelitian selanjutnya.

