

BAB I

PENDAHULUAN

1.1. Latar Belakang

Semakin pesat dan murah nya perkembangan teknologi media penyimpanan digital telah mendorong terjadinya ledakan jumlah dokumen elektronik yang tersimpan dalam *repository* perpustakaan Universitas. Berbagai karya ilmiah dari sivitas akademika seperti skripsi, laporan penelitian, laporan kerja praktik dan lain sebagainya telah tersedia dalam versi digital (Prilianti dan Wijaya, 2014). Namun, pada umumnya fenomena ini tidak disertai dengan pertumbuhan jumlah informasi atau pengetahuan yang dapat digali dari dokumen-dokumen elektronik tersebut (Gupta, 2011). Melimpahnya informasi teks tidak terstruktur telah mendorong munculnya disiplin baru dalam analisis teks, yaitu *text mining* yang mencoba menemukan pola-pola informasi yang dapat digali dari suatu teks yang tidak terstruktur tersebut (Hamzah, 2012). Algoritma-algoritma dalam *text mining* dibuat untuk dapat mengenali data yang sifatnya semi terstruktur misalnya isi dari dokumen-dokumen (Gupta dan Lehal, 2009). Salah satu dokumen tersebut adalah proposal skripsi.

Proposal skripsi di Fakultas Sains dan Teknologi Universitas Katolik Musi Charitas program studi informatika memiliki topik-topik saat pendaftaran seminar proposal. Topik-Topik tersebut terdiri dari Algoritma dan Struktur data, Arsitektur Komputer, Rekayasa Perangkat Lunak, *Artificial Intelligence* dan Robotik, Interaksi Manusia dan Komputer, Bahasa Pemrograman, Sistem Operasi dan Jaringan, *Database* dan Sistem Retrieval Informasi, Grafika Komputer, Ilmu Komputasi. Namun, dalam membagi proposal skripsi ke dalam topik-topik tersebut isi dari proposal skripsi perlu dibaca agar dapat mengklasifikasikan ke topik yang sebenarnya dimana dalam membaca isi proposal skripsi tersebut terkadang dilakukan secara berulang-ulang dan hal tersebut tentu saja memakan waktu. Dalam mendaftar sidang seminar mahasiswa diminta untuk mengisi formulir yang telah disediakan yang berisikan nama, nim, judul, dan topik yang diambil mahasiswa

tersebut. Formulir tersebut nantinya akan dikumpulkan ke bagian koordinator tugas akhir pada jadwal yang telah ditentukan, selanjutnya koordinator tugas akhir menggunakan data tersebut untuk merekap data mahasiswa dan menyusun jadwal sidang seminar. Oleh karena itu, perlu adanya aplikasi untuk dapat mengklasifikasikan proposal skripsi secara otomatis agar sesuai dengan topik-topik proposal skripsi yang ada, dimana akan diimplementasikan saat mahasiswa akan mendaftar sidang seminar dan menggunakan algoritma dalam klasifikasinya.

Menurut Prasetyo (2012) Ada berbagai macam algoritma yang digunakan untuk klasifikasi seperti *Artificial Neural Network (ANN)*, *Support Vector Machine*, *Decision Tree*, *Bayesian* dan sebagainya. Dalam klasifikasinya, *Naive Bayes* menggunakan data latih untuk mengklasifikasikan sesuai dengan label kelas keluarannya. Menurut penelitian Hamzah (2012) algoritma *Naive Bayes Classifier* memiliki beberapa kelebihan antara lain sederhana, cepat, dan berakurasi tinggi.

Berdasarkan hal yang telah diuraikan di atas maka dibangunlah sebuah aplikasi yang dapat mengklasifikasikan proposal skripsi secara otomatis. Pembuatan aplikasi menggunakan algoritma *Naive Bayes Classifier* dalam pengklasifikasian topik-topik tersebut. Aplikasi klasifikasi yang dibuat berbasis web sehingga bisa diakses dimana dan kapan saja serta menggunakan bahasa PHP dan *database MySQL*.

1.2. Perumusan Masalah

Masalah yang ingin diselesaikan dalam penelitian ini adalah mengklasifikasikan dokumen proposal skripsi secara otomatis dengan menggunakan *Naive Bayes Classifier* dan bagaimana kinerja dari algoritma tersebut.

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Dokumen yang digunakan adalah dokumen proposal skripsi di Fakultas Sains dan Teknologi khususnya program studi informatika Universitas Katolik Musi Charitas.

2. Topik-Topik yang digunakan hanya sebatas topik yang ada pada formulir pendaftaran sidang seminar pada program studi Informatika.
3. Dokumen yang digunakan merupakan Bahasa Indonesia.
4. Dokumen hanya berupa teks.
5. Pengklasifikasian hanya pada bagian pendahuluan sampai dengan metodologi penelitian.

1.4. Tujuan dan Manfaat

Tujuan dan manfaat yang ingin dicapai dalam melakukan penelitian ini secara jelas dapat dilihat sebagai berikut :

1.4.1. Tujuan Penelitian

Berdasarkan latar belakang dan permasalahan yang ada, maka tujuan penelitian ini untuk :

1. Membangun aplikasi yang dapat mengklasifikasikan proposal skripsi secara otomatis menggunakan klasifikasi *Naive Bayes Classifier*.
2. Melihat kinerja klasifikasi *Naive Bayes Classifier* pada klasifikasi topik proposal skripsi

1.4.2. Manfaat Penelitian

Adapun manfaat yang bisa dirasakan dari hasil penelitian adalah sebagai berikut:

1. Dapat mengklasifikasikan topik proposal skripsi dengan cepat.
2. Menambah wawasan para peneliti yang tertarik pada bidang klasifikasi terutama klasifikasi teks.

1.5. Metodologi Penelitian

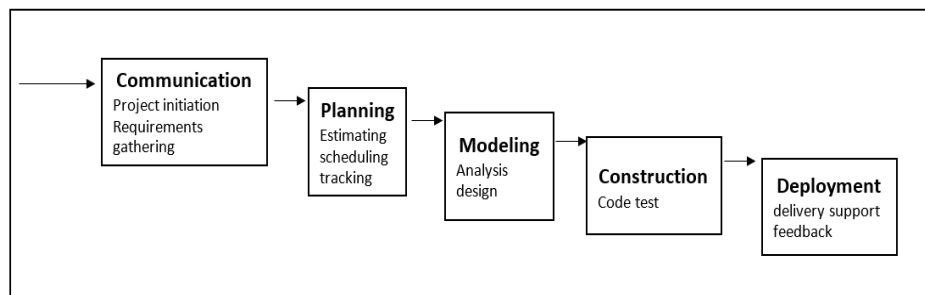
Metodologi penelitian yang digunakan dalam penelitian ini dapat dilihat sebagai berikut:

1.5.1. Jenis Penelitian

Berdasarkan klasifikasi penelitian berdasarkan tujuan dan manfaatnya, penelitian ini merupakan jenis penelitian terapan. Penelitian ini dilakukan untuk menerapkan algoritma *Naive Bayes Classifier* pada aplikasi pengklasifikasian topik proposal skripsi, sehingga dapat mengklasifikasikan topik proposal skripsi secara otomatis dengan cepat.

1.5.2. Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah model *waterfall*. Model ini melakukan pendekatan kepada pengembangan perangkat lunak yang sistematis dan sekuensial yang mulai pada tingkat dan kemajuan sistem pada *communication*, *planning*, *modelling*, *construction* dan *deployment* (Pressman, 2010). Adapun ilustrasinya dapat dilihat pada gambar 1.1.



Gambar 1.1 Model Waterfall

(Sumber : Pressman, 2010)

Adapun penjelasan mengenai fase-fase tersebut menurut Pressman (2010) :

a. *Communication*

Langkah ini untuk mengumpulkan fitur dan fungsi dari perangkat lunak yang akan dibuat untuk pelanggan atau pihak berkepentingan lainnya. Pada fase ini peneliti akan menganalisis kebutuhan perangkat lunak, seperti mengumpulkan data proposal skripsi dan juga data rekap pengumpulan topik yang akan digunakan dalam membuat aplikasi klasifikasi proposal skripsi.

b. *Planning*

Pada tahap ini menjelaskan resiko yang mungkin terjadi, sumber daya yang diperlukan, kerangka kerja produk yang akan dibuat. Pada tahap peneliti melakukan perencanaan yaitu membuat jadwal pelaksanaan.

c. *Modelling*

Pada tahap ini menciptakan model-model yang untuk lebih memahami pesyaratan perangkat lunak dan desain yang akan diimplementasikan. Pada tahap ini peneliti membuat design dari sistem aplikasi klasifikasi.

d. *Contruction*

Construction merupakan proses membuat kode. *Coding* atau pengkodean merupakan penerjemahan desain dalam bahasa yang bisa dikenali oleh komputer. *Programmer* akan menerjemahkan transaksi yang diminta oleh *user*. Tahap ini direalisasikan ke dalam bentuk kode program.

e. *Deployment*

Tahapan ini bisa dikatakan final dalam pembuatan sebuah *software* atau sistem. Setelah melakukan analisis, desain dan pengkodean maka sistem yang sudah jadi akan digunakan oleh *user*. Kemudian *software* yang telah dibuat harus dilakukan pemeliharaan secara berkala.

Pemilihan metode *waterfall* dalam pengembangan sistem ini dikarenakan dalam pengerjaan proyek ini dilakukan secara disiplin, sehingga membuat spesifikasi kebutuhan pada sistem ini jelas karena tidak adanya campur tangan pihak lain dan juga membuat penelitian ini akan berjalan secara linier.

1.6. Sistematika Penulisan

Penulisan laporan dilakukan secara sistematis menggunakan beberapa bab. Gambaran umum mengenai isi laporan penelitian secara keseluruhan akan diuraikan secara singkat dengan sistematika sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisikan latar belakang permasalahan yang dibahas penelitian, perumusan masalah, tujuan dan manfaat penelitian, metodologi penelitian dan sistematika penulisan laporan

BAB II LANDASAN TEORI

Bab ini berisikan teori-teori dasar yang digunakan landasan dan referensi dalam penelitian. Teori tersebut mencakup teori mengenai klasifikasi, *Naive bayes Classification*, *Information Retrieval*, Praproses teks, dan *Unified Modeling Language (UML)*, Pengujian Sistem yang digunakan dalam penelitian.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini memberikan analisis kebutuhan sistem, permodelan UML, desain antarmuka sistem dan arsitektur sistem pada sistem pengklasifikasian proposal skripsi yang akan dibangun.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini memberikan pembahasan dari proses pembuatan sistem, seperti hasil penerapan desain dalam sistem dan juga pengujian sistem.

BAB V KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan yang diambil setelah melakukan penelitian serta saran-saran yang dapat digunakan untuk pengembangan sistem dalam penelitian berikutnya.