BAB I

PENDAHULUAN

1.1.Latar Belakang Masalah

Tugas Akhir (TA) atau Skripsi merupakan suatu karya tulis ilmiah, berupa paparan tulisan hasil penelitian yang membahas suatu masalah dalam bidang ilmu tertentu dengan menggunakan kaidah-kaidah yang berlaku dalam bidang ilmu tersebut. Tugas akhir mahasiswa akan sesuai dengan kajian-kajian yang ada di jurusan yang mereka ambil. Sebagai contoh mahasiswa yang mengambil jurusan Teknik Informatika kajian tugas akhirnya pasti tentang keinformatikaan. Tugas akhir ini terdiri dari berbagai judul sesuai dengan bidang kajian dari masing-masing kajian ilmu pengetahuan mahasiswa. Hal ini sesuai dengan minat dan bakat mahasiswa.

Sebagai salah satu jurusan yang paling banyak diminati di Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Sunan Gunung Djati Bandung, dengan jumlah mahasiswa sebanyak 805 (delapan ratus lima) mahasiswa dan jumlah dosen sebanyak 9 (sembilan) dosen tetap dan 54 (lima puluh empat) dosen tidak tetap (data ini diambil dari data mahasiswa dan dosen di jurusan teknik informatika UIN Sunan Gunung Djati Bandung sampai bulan Agustus 2015), jurusan teknik informatika UIN Sunan Gunung Djati Bandung memiliki banyak mahasiswa yang harus memiliki dosen pembimbing dan penguji saat mengerjakan tugas akhirnya. Mahasiswa yang sedang mengerjakan tugas akhir membutuhkan dosen pembimbing, sedangkan mahasiswa yang akan sidang tugas akhir juga membutuhkan dosen penguji. Untuk menetukan dosen

pembimbing dan dosen penguji sebuah judul tugas akhir, diperlukan data kelas keahlian dari judul tugas akhir tersebut agar kajian penelitian dari tugas akhir tersebut sesuai dengan bidang keahlian dari dosen penguji maupun dosen pembimbing. Oleh karena itu pengklasifikasian judul tugas akhir dan penentuan pembimbing dan penguji tugas akhir mahasiswa merupakan hal yang sangat penting.

Mahasiswa yang akan mengajukan proposal tugas akhir harus menghubungi pihak jurusan untuk memastikan metode yang akan digunakan sudah dipergunakan oleh mahasiswa lain atau belum, setelah itu mahasiswa harus mencari tugas akhir yang memiliki kajian serupa untuk lebih memahami masalah dan kajian yang akan diteliti dengan cara melihat satu persatu judul tugas akhir di lemari perpustakaan jurusan. Dalam hal ini jurusan Teknik Informatika UIN Sunan Gunung Djati Bandung sebagai pihak yang menjadi tempat penelitian ini. Hal ini merupakan masalah teknis mendasar yang harusnya bisa dilakukan lebih mudah dengan adanya sistem klasifikasi data judul.

Pengklasifikasian dokumen teks sangatlah diperlukan di tengah-tengah perkembangan teknologi. Efektifitas dan biaya dalam pengelompokan manual menjadi faktor utama diperlukannya pengelompokan secara otomatis. Pengelompokan dokumen adalah proses pengkategorian dokumen-dokumen teks berdasarkan kategori isi dokumen. Metode *Naïve Bayes Classifier* merupakan salah satu metode pengklasifikasian yang sering digunakan, dan memiliki akurasi yang cukup tinggi. Menurut Cahyo Darujati, dkk. (2012) dalam penelitiannya menyatakan presentase keberhasilan klasifikasi teks

dengan metode NBC adalah lebih dari 87%. Presentase ini lebih tinggi jika dibandingkan dengan metode lain seperti metode K-Nearest Neighbor (KNN) pada penelitian yang dilakukan Ridok, A., dkk. (2009) dengan presentase keberhasilan sebesar 71, 58%.

Berdasarkan hal tersebut di atas, peneliti berusaha menggunakan metode *Naïve Bayes Classifier* untuk mengklasifikan judul-judul tugas akhir mahasiswa ke dalam kelas-kelas tertentu sesuai dengan kajiannya masing-masing dan memberikan rekomendasi dosen pembimbing dan penguji yang keahliannya sesuai dengan kajian judul tugas akhir tersebut.

1.2. Rumusan Masalah

Berdasarkan pemaparan di latar belakang, permasalahan yang muncul adalah sebagai berikut:

1. Bagaimana mengklasifikasikan data judul tugas akhir mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Sunan Gunung Djati Bandung dengan menggunakan metode *Naïve Bayes Classifier*?
2. Berapakah persentase keberhasilan klasifikasi judul tugas akhir dengan metode *Naïve Bayes Classifier?*

1.3. Batasan Masalah

a. Data yang digunakan dalam penelitian ini adalah judul-judul tugas akhir mahasiswa jurusan teknik informatika UIN Bandung. Jumlah judul yang digunakan dalam penelitian ini adalah 200 (dua ratus) judul tugas akhir yang diambil dari *website e-library (http://e-library.if.uinsgd.ac.id/)*

Teknik Informatika Universitas Islam Negeri Sunan Gunung Djati Bandung.

b. Berdasarkan hasil observasi yang telah dilakukan peneliti di jurusan Teknik Informatika Universitas Islam Negeri Sunan Gunung Djati Bandung terdapat empat kelas judul tugas akhir, yaitu kelas Manajemen data dan Sistem Informasi, kelas Sistem Komputer dan Komputasi Terdistribusi, kelas Pemrograman dan Rekayasa Perangkat Lunak, dan kelas Visi Komputer dan Sistem Berintelegensia. Hal ini sesuai dengan kelompok keahlian jurusan teknik informatika yang dibentuk pada rapat internal jurusan teknik informatika pada tanggal 28 Juni 2016.

c. Metode yang digunakan adalah algoritma *Naïve Bayes Classifier* untuk klasifikasi teks dan algoritma TF-IDF untuk pembobotan kata dalam teks judul.

d. Penentuan dosen pembimbing dan penguji masih dilakukan secara manual dengan mencocokkan kelas judul dengan kelompok keahlian dosen.

e. Bahasa dari dokumen yang digunakan dalam klasifikasi dokumen teks ini adalah Bahasa Indonesia.

f. *Output* yang dihasilkan berupa *class* dan rekomendasi dosen pembimbing dan penguji untuk judul yang di-*input*-kan.

1.4. Tujuan Penelitian

Penelitian ini bertujuan untuk:

a. Memberikan informasi tentang klasifikasi tugas akhir mahasiswa.

b. Membantu pihak jurusan dalam mengelompokan laporan tugas akhir baik dari sisi data digital maupun data fisik.
c. Membantu jurusan dalam menentukan dosen pembimbing dan penguji tugas akhir.

1.5. *State of The Art*

State of The Art menggambarkan posisi kajian penelitian yang sedang dilakukan terhadap penelitian-penelitian yang telah dilakukan sebelumnya dengan kajian yang sama. Berikut ini merupakan *review* dari penelitian-penelitian terdahulu yang kajiannya berkaitan dengan kajian peneliti yang sedang berlangsung yaitu tentang klasifikasi dokumen teks:

Penelitian yang dilakukan Silfia Andini (2013) tentang klasifikasi dokumen teks melalui proses *text mining* dengan menggunakan algoritma *Naïve Bayes Classifier* sebagai metode utamanya. Penelitian tersebut diimplementasikan ke dalam sistem dengan menggunakan bahasa pemrograman java. Penelitian tersebut dilatar belakangi oleh sulitnya menegetahuai dokumen mana yang dibutuhkan dari sebuah kumpulan dokumen yang tersedia [8].

Salah satu cara yang paling berhasil untuk mengorganisasikan informasi dalam jumlah banyak dan dapat dipahami oleh para pencari informasi adalah dengan melakukan klasifikasi dokumen berdasarkan topiknya. Oleh karena itu, diperlukan suatu strategi pengelompokan otomatis dokumen-dokumen berita tersebut. Penelitian yang dilakukan oleh Herny Februariyanti dan Eri Zuliarso (2012) ini menggunakan metode Ontologi. Obyek penelitian dari penelitian ini adalah artikel berita berbahasa Indonesia dari situs *http://www.google.com* [12].

Penelitian tentang klasifikasi teks dengan metode *Naïve Bayes* juga telah dilakukan oleh Amalia Indranandita, dkk. (2008). Data yang digunakan dalam penelitian tersebut adalah jurnal-jurnal dan artikel ilmiah. Penelitian ini dilatar belakangi oleh kebutuhan konsumen terhadap informasi dalam bentuk jurnal atau artikel ilmiah semakin yang meningkat, sehingga pengelompokan jurnal dibutuhkan untuk mempermudah pencarian informasi. Penelitian ini menggunkan Sistem klasifikasi dan pencarian jurnal dengan metode *Naive Bayes Classifier* dan *Vector Space Model* dengan pendekatan *Cosine* yang diharapkan dapat membantu pengguna dalam penentuan topik/ kategori dan menghasilkan daftar jurnal berdasarkan urutan tingkat kemiripan [13].

Buku referensi matakuliah harus ditentukan agar sesuai dengan silabus mata kuliah yang berkaitan, untuk mengatasi hal tersebut Atri Nurani, dkk. (2007) telah melakukan penelitian tentang klasifikasi buku referensi mata kuliah dengan menggunakan metode *Naive Bayes Classifier*. Data yang digunakan berupa data buku yang ada di perpustakaan. Penggunaan metode *Naïve Bayes Classifier* ini memiliki nilai presisi 63%. Dalam melaksanakan tugasnya untuk mengklasifikasikan daftar isi buku referensi, sistem ini dipengaruhi oleh berbagai faktor seperti pola data dan jumlah *data training* [14].

Tabel 1.1 Penelitian yang telah dilakukan

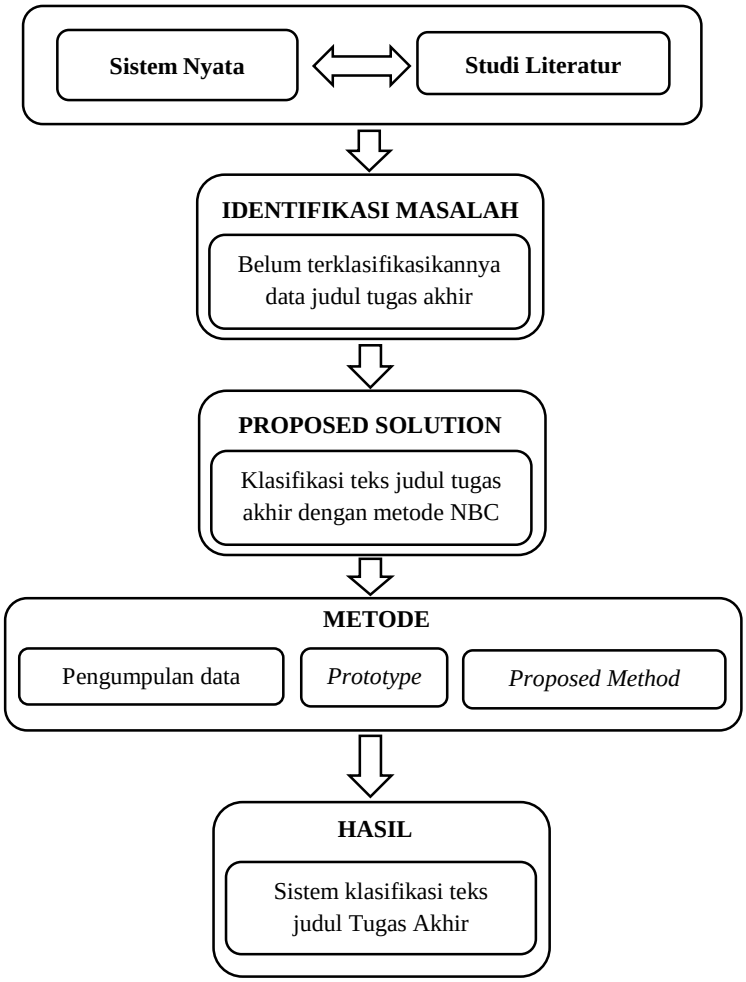
No.	Peneliti	Metode	Data	Teknologi	Hasil
1	(Silfia Andini. 2013)	*Text mining* dan *Naïve Bayes Classifier*	Dokumen teks	Java	Klasifikasi dokumen teks ke dalam kategori tertentu

Tabel 1.1 Penelitian yang telah dilakukan (Lanjutan)

No.	Peneliti	Metode	Data	Teknologi	Hasil
2	(Herny Februariyanti, Eri Zuliarso. 2012)	Ontologi	Artikel berita berbahasa Indonesia	Protégé 4.0	Klasifikasi dokumen berita teks bahasa Indonesia ke dalam dataset yang telah ditentukan
3	(Amalia Indranandita, Budi Susanto, Antonius Rahmat C. 2008)	*Naïve Bayes Classifier* dan *Vector Space Model*	Jurnal dan artikel ilmiah	VSM	Klasifikasi dan pencarian jurnal yang diinginkan
4	(Atri Nurani, Budi Susanto, Umi Proboyekti. 2007)	*Naïve Bayes Classifier*	Buku diperpustakaan	*Naïve Bayes Classifier*	Klasifikasi buku referensi mata kuliah

Berdasarkan hasil *review* dari keempat sumber di atas, terdapat beberapa perbedaan antara penelitian yang sedang dilakukan peneliti dengan penelitian yang telah dilakukan. Kajian penelitian yang sedang dilakukan adalah pengklasifikasian dokumen teks yang berupa judul-judul tugas akhir teknik informatika ke dalam kelas-kelas yang telah ditentukan. Metode yang digunakan adalah *Naïve Bayes Classifier* dan diimplementasikan dengan bahasa pemrograman php. Perbedaan dengan penelitian sebelumnya terletak pada data serta bahasa pemrograman yang digunakan.

1.6. Kerangka Pemikiran



Gambar 1.1 Kerangka Pemikiran

Pada Gambar 1.1 menjelaskan alur penelitian yang akan dilakukan dalam tugas akhir, dari analisa terhadap permasalahan sampai menemukan hasil dari penelitian. Adapun langkah-langkahnya sebagai berikut:

1. Analisa sistem nyata yaitu proses analisa terhadap sistem yang sedang berjalan dan dijadikan sebagai acuan dalam pembuatan sistem.

2. Studi literatur menerangkan tentang teori-teori yang relefan sebagai perbandingan antara jurnal yang telah membahas kajian yang sama dengan tema penelitian yang akan dilakukan.
3. Identifikasi masalah yaitu menentukan permasalahan yang diambil dari analisa sistem nyata dan studi literatur.
4. *Proposed solution* merupan solusi yang ditawarkan dalam memecahkan masalah dalam penelitian.
5. Metode menerangkan proses pembuatan perangkat lunak dari metode perangkat lunak yang digunakan, pengumpulan data yang didapat sampai proses *logic* yang ada pada sistem.
6. Hasil akhir yaitu *output* berupa sistem yang dihasilkan dari penelitian yaitu klasifikasi teks judul tugas akhir.

1.7. Metodologi Penelitian

a. Teknik Pengumpulan Data

Untuk mengumpulkan berbagai data yang diperlukan dalam penelitian ini ada beberapa cara yang akan dilakukan sehingga penelitian yang dilakukan dapat berjalan secara maksimal. Berikut ini merupakan teknik-teknik yang akan digunakan peneliti dalam mengumpulkan data yang diperlukan:

1. Observasi: Melakukan observasi dan analisa data yang diperlukan di Jurusan Teknik Informatika Universitas Islam Negeri Sunan Gunung Djati Bandung.
2. Wawancara: Melakukan wawancara kepada pihak jurusan

b. Metode Pengembangan Perangkat Lunak

Metode ini merupakan metode yang akan digunakan peneliti dalam proses pengembangan perangkat lunak sebagai *end product* dari penelitian ini. Metode yang akan digunakan adalah metode *prototype.*

1.8. Sistematika Penulisan

Sistematika penulisan karya ilmiah ini dibagi menjadi lima bab. Setiap bab memiliki memiliki tujuan masing-masing. Berikut ini merupakan penjelasan tentang peran masing-masing bab:

BAB I PENDAHULUAN

Bab ini mendeskripsikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, *state of the art*, kerangka pemikiran, metodologi penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini membahas teori-teori yang berkaitan dengan penelitian yang dilakukan oleh penulis. Teori ini diambil dari berbagai sumber, buku-buku dan jurnal-jurnal yang berkaitan dengan topik penelitian yang digarap penulis. Dalam bab ini juga dijelaskan teori-teori yang menunjang pembuatan produk akhir dari penelitian ini, yaitu perangkat lunak.

BAB III ANALISIS DAN PERANCANGAN

Bab ini menjelaskan tentang hasil analisis *requirements system*, yaitu analisa kebutuhan sistem yang akan dibangun. Bab ini juga menjelaskan tentang faktor-faktor apa saja yang harus diperhatikan dalam perangcangan sistem yang akan

dibangun secara rinci yang meliputi perancangan *global*, perancangan kode, serta perancangan basis data.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Pada bab ini akan dijelaskan modul-modul yang dibentuk yaitu tabel-tabel basis data, struktur *menu*, spesifikasi *hardware* dan bahasa pemrograman yang digunakan serta tabel pengujian dengan menggunakan metode *Blackbox*.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran-saran.