

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi saat ini berkembang pesat di segala bidang kehidupan. Banyak sekali data yang dihasilkan oleh teknologi informasi, mulai dari bidang industri, ekonomi, serta berbagai ilmu lainnya. Penerapan teknologi informasi dalam dunia pendidikan perguruan tinggi juga dapat menghasilkan data yang berlimpah mengenai mahasiswa dan proses pembelajaran yang dihasilkan.

Jurusan Teknik Informatika UIN Sunan Gunung Djati Bandung memiliki rata-rata presentasi kelulusan mahasiswa tepat waktu yaitu 13.5%. Sedangkan tiap memasuki tahun ajaran baru kuota mahasiswa yang diterima semakin bertambah. Dari keterangan tersebut diketahui bahwa tidak semua mahasiswa dapat lulus tepat waktu sesuai dengan masa studi yang ditempuh sehingga mengakibatkan jumlah mahasiswa semakin banyak yang berdampak pula pada peningkatan jumlah data.

Salah satu teknik yang bisa digunakan untuk mengoptimalkan data agar memiliki nilai manfaat serta menjadi informasi dan pengetahuan adalah dengan *data mining*. *Data mining* adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan secara otomatis. Definisi lain dari *Data Mining* yaitu proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep yang akan dipelajari[2].

Pada penelitian sebelumnya [3] yang dilakukan dengan melakukan uji perbandingan algoritma klasifikasi yaitu *C4.5*, *Naive Bayes*, *KNN* dan *SVM* pada

prediksi penyakit kulit menunjukkan bahwa algoritma yang diuji mempunyai nilai akurasi, presisi, dan recall diatas 94%. Sedangkan dalam penelitian sekarang yaitu uji komparasi menggunakan *Naive Bayes Classifier* dan *C4.5*.

Algoritma *C4.5* merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma yang lainnya. Kelebihan algoritma *C4.5* dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani atribut bertipe diskret dan dapat menangani atribut bertipe diskret dan numerik[4]. sedangkan menurut Wahyu Eko Susanto, dan Candra Agustina memaparkan perbandingan algoritma *C4.5* dan *Naive Bayes*[5].

Algoritma	Accuracy	Precision	Recall
<i>C4.5</i>	67.27%	57.80%	88.73%
<i>Naive Bayes</i>	70.30%	67.19%	60.56%

Gambar 1.1 Perbandingan hasil pengujian *C4.5* dan *Naive Bayes*[5]

Dari penelitian yang sudah dilakukan terdapat aplikasi dalam menentukan *dress up matching* dengan menggunakan Metode *Naive Bayes*. Hasil penelitian tersebut menyatakan tingkat akurasi sebesar 84% dari 50 data latih.

Penelitian selanjutnya yaitu klasifikasi kelulusan mahasiswa menggunakan algoritma *Naive Bayes*. Hasil dari penelitian tersebut menyatakan tingkat akurasi algoritma *Naive Bayes* cukup tinggi yaitu sebesar 82%. Namun berdasarkan peneliti hasil tersebut dikarenakan kurangnya kompleksitas data. Saran untuk penelitian selanjutnya yaitu membandingkan dengan algoritma *C4.5* atau metode klasifikasi lainnya sehingga dapat mengetahui kelebihan masing-masing[6].

Berdasarkan hasil perbandingan penelitian tersebut bahwa algoritma *Naive Bayes Classifier* dan *C4.5* memiliki nilai akurasi yang cukup tinggi serta mudah

untuk diimplementasikan. Penulis akan mencoba membandingkan algoritma Naive Bayes Classifier dengan C4.5 berdasarkan data mahasiswa Teknik Informatika UIN Sunan Gunung Djati. Atribut yang akan digunakan dalam memprediksi kelulusan adalah NIM, nama, jenis kelamin, IPK, jalur masuk, tahfidz, asal sekolah, dan kegiatan ekstra selama masa perkuliahan. Kriteria mahasiswa yang dinyatakan lulus tepat waktu adalah mahasiswa yang menempuh masa pendidikan selama 4 tahun merujuk pada Ayat 2 (d) Pasal 17 Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) Nomor 49 Tahun 2014 tentang Standar Nasional Pendidikan Tinggi.

Dengan adanya penelitian ini diharapkan dapat menjadi informasi dan pengetahuan untuk bisa mengetahui algoritma mana yang cocok untuk menerapkan pada aplikasi prediksi masa studi mahasiswa Teknik Informatika Sunan Gunung Djati.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas maka diperoleh rumusan masalah sebagai berikut:

- a. Bagaimana akurasi algoritma *Naive Bayes Classifier* dan C4.5 dalam memprediksi masa kelulusan mahasiswa UIN Sunan Gunung Djati?
- b. Bagaimana kecepatan algoritma *Naive Bayes Classifier* dan C4.5 dalam memprediksi masa kelulusan mahasiswa UIN Sunan Gunung Djati?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini yaitu:

- a. mengetahui akurasi algoritma *Naive Bayes Classifier* dan C4.5 dalam memprediksi masa studi mahasiswa.
- b. mengetahui kecepatan algoritma *Naive Bayes Classifier* dan C4.5 dalam memprediksi masa studi mahasiswa.

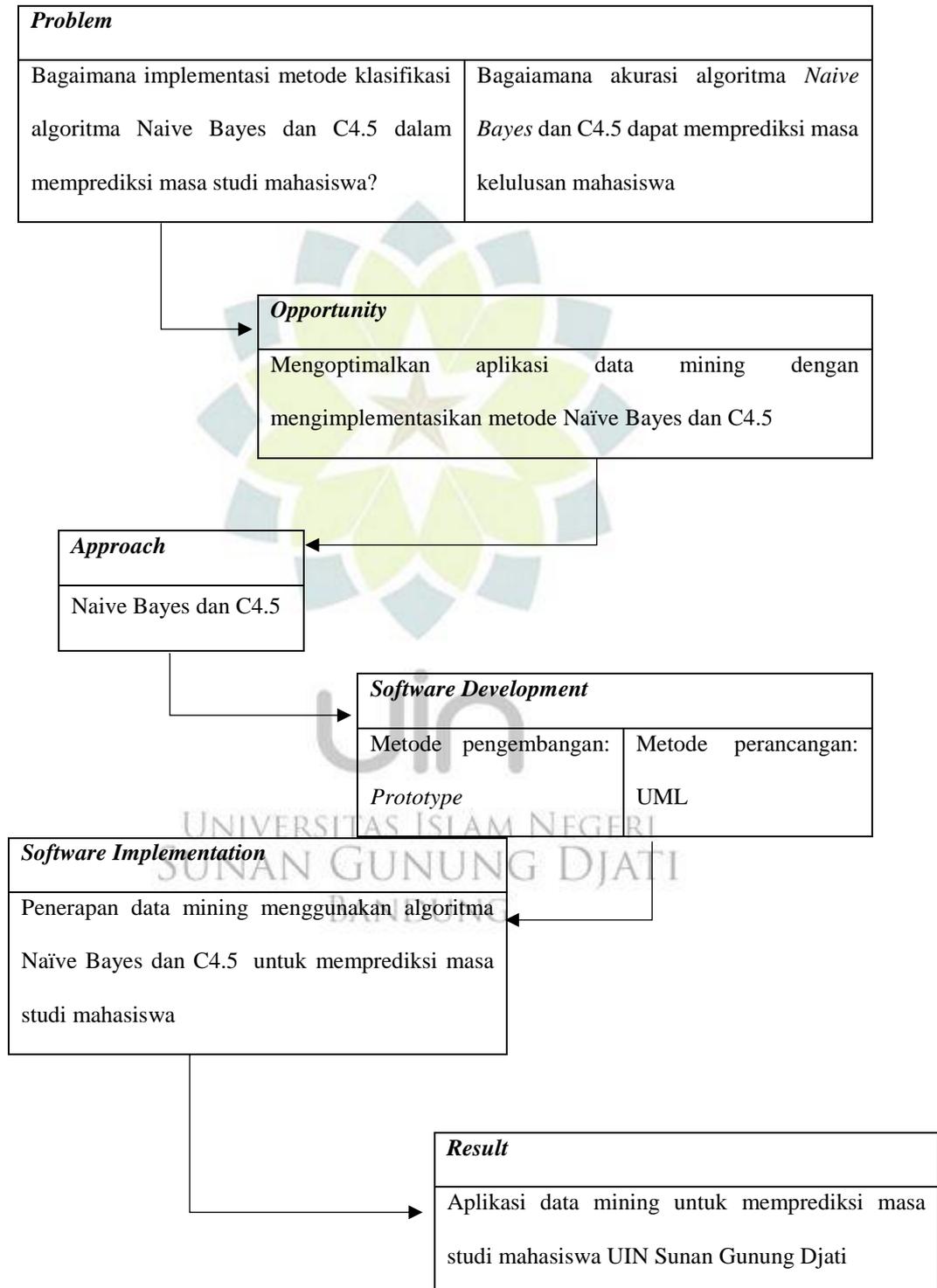
1.4 Batasan Masalah

Untuk membahas topik yang lebih terarah dan terfokus pada tujuan yang ingin dicapai, maka batasan masalahnya sebagai berikut:

- a. Ruang lingkup penelitian ini dibatasi pada data mahasiswa angkatan tahun 2011 dan 2012 Jurusan Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Gunung Djati Bandung.
- b. Kriteria lulus tepat waktu yaitu lama studi selama 4 tahun. Dihitung dari tahun lulus dikurangi tahun masuk mahasiswa.

1.5 Kerangka Penelitian

Berdasarkan tinjauan pustaka yang telah dijelaskan di atas, maka kerangka pemikiran dapat diuraikan Gambar 1.1:



Gambar 1.2 Kerangka Pemikiran Metodologi

1.6 Metodologi Penelitian

1.6.1 Gambaran Umum Penelitian

Metodologi yang akan digunakan dalam penelitian ini meliputi:

1. Studi Literatur

Studi literatur dilakukan dengan melakukan pembelajaran berbagai macam buku dan jurnal yang mendukung tugas akhir ini.

2. Analisis

Melakukan analisis algoritma Naive Bayes dalam penentuan masa studi mahasiswa

3. Perancangan Perangkat Lunak

Melakukan perancangan terhadap perangkat lunak yang akan dikembangkan berdasarkan hasil yang diperoleh dari analisis.

4. Implementasi

Melakukan implementasi terhadap perangkat lunak yang akan dikembangkan berdasarkan hasil yang diperoleh dari perancangan.

5. Pengujian

Melakukan pengujian perangkat lunak dengan menggunakan studi kasus yang telah ditentukan.

6. Evaluasi

Mengevaluasi hasil konstruksi perangkat lunak yang dikembangkan.

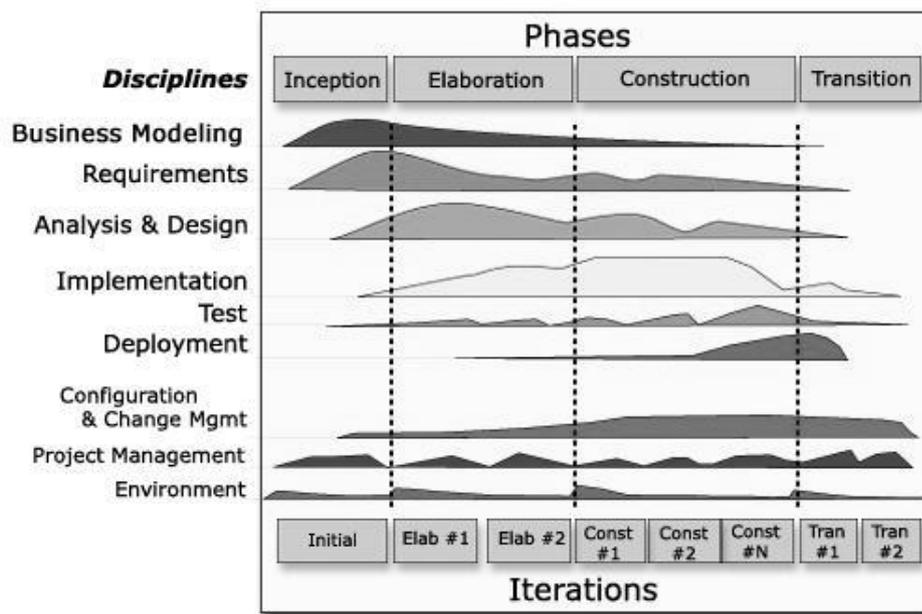
7. Bimbingan dan Diskusi

Mengadakan bimbingan dengan dosen pembimbing maupun dosen lain, serta berdiskusi dengan sesama mahasiswa.

1.6.2 Metode Pengembangan Perangkat Lunak

Metode pengembangan perangkat lunak yang digunakan untuk penelitian skripsi ini, yaitu *Rational Unified Process*, merupakan suatu metode rekayasa perangkat lunak yang dikembangkan dengan mengumpulkan berbagai *best practises* yang terdapat dalam industri pengembangan perangkat lunak. Ciri utama metode ini adalah menggunakan *use-case driven* dan pendekatan iteratif untuk siklus pengembangan perangkat lunak.

RUP menggunakan konsep *object oriented*, dengan aktifitas yang berfokus pada pengembangan model dengan menggunakan Unified Model Language (UML).



Gambar 1.3 Arsitektur Rational Unified Process [7]

Model pengembangan perangkat lunak *Rational Unified Process* sangat cocok diterapkan pada sistem atau perangkat lunak yang bersifat *customize* artinya perangkat lunak yang dikembangkan sesuai berdasarkan kebutuhan dan kondisi

tertentu dan sesuai untuk tugas akhir/skripsi yang memiliki tujuan untuk mengimplementasikan sebuah algoritma atau metode pada suatu kasus[3].

1.7 Sistematika Penulisan

Pembuatan skripsi ini dilakukan dengan pembagian bab sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini membahas mengenai latar belakang, rumusan masalah, batasan masalah, tujuan, metode penelitian, dan sistematika penulisan.

BAB II : TINJAUAN PUSTAKA

Bab ini menjelaskan tentang algoritma C4.5 dan Naive Bayes Classifier. Adapun literatur yang digunakan meliputi buku referensi dan dokumentasi internet.

BAB III : ANALISIS DAN PERANCANGAN

Bab ini dijelaskan tinjauan organisasi, tahap perencanaan, dan perancangan sistem informasi tentang perbandingan algoritma Naive Bayes Classifier dan C4.5 dalam memprediksi kelulusan mahasiswa.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini menjelaskan tahapan implementasi dan uji coba dari perancangan sistem serta hasil analisis.

BAB V : KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari bahasan dan saran yang bermanfaat untuk pengembangan skripsi ini.