

BAB I

PENDAHULUAN

I.1. Latar Belakang

Tingginya tingkat keberhasilan mahasiswa dan rendahnya tingkat kegagalan mahasiswa merupakan cermin kualitas dari suatu perguruan tinggi. Kegagalan mahasiswa dan faktor-faktor penyebabnya merupakan topik yang menarik untuk diteliti.

Saat ini sekolah tinggi berada dalam lingkungan yang sangat kompetitif. Setiap perguruan tinggi berusaha untuk terus memperbaiki menejemennya untuk meningkatkan mutu pendidikan dan meningkatkan akreditasi. Salah satu elemen penilaian akreditasi perguruan tinggi adalah lulus tepat waktu. Selain itu wisuda tepat waktu merupakan isu yang penting karena tingkat kelulusan sebagai dasar efektifnya suatu perguruan tinggi. Salah satu persoalan yang masih menjadi bahan pembicaraan tentang kegagalan mahasiswa adalah mengenai mahasiswa berhenti studi (drop out/non-aktif) dan kelulusan. Mahasiswa non-aktif (mangkir) adalah mahasiswa yang tidak melakukan registrasi pada awal semester atau tidak mengikuti perkuliahan sekurang-kurangnya satu semester (Diana Laily Fithri ; 2014 :1).

Data mining adalah teknik bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang disimpan. Salah satu teknik data mining adalah teknik klasifikasi. Teknik klasifikasi adalah

teknik pembelajaran untuk prediksi suatu nilai dari target variabel kategori. Karena dalam prediksi mahasiswa non-aktif ini memiliki dua variable target yaitu aktif dan non-aktif maka penelitian ini akan menggunakan data mining teknik klasifikasi (Relita Buaton ; 2012 : 90-91).

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 (Yosoa Putra Raharja, 2012 : 2).

Oleh karena itulah, penulis mengangkat sebuah judul “**Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa Pada Politeknik Ganesha**”.

I.2. Ruang lingkup Permasalahan

Adapun beberapa tahap yang dilakukan dalam membuat ruang lingkup permasalahan adalah :

I.2.1 Identifikasi Masalah

Peneliti mencoba untuk mengidentifikasi masalah sebagai berikut :

1. mahasiswa yang tidak lulus sesuai dengan waktu studi.
2. Minimnya transformasi pengetahuan dari prediksi ketepatan jumlah data mahasiswa baru yang setiap tahun terus meningkat ke dalam bentuk representasi pengetahuan yang akan dirancang.

3. penerapan dan perbandingan data mining dan algoritma C4.5 prediksi ketepatan waktu kelulusan.

I.2.2. Rumusan Masalah

Berdasarkan identifikasi masalah yang ditemukan oleh penulis dalam melakukan penelitian ini, maka perumusan masalah dapat dirumuskan sebagai berikut :

1. Bagaimana merancang penerapan data mining dan C4.5 dalam prediksi kelulusan mahasiswa yang dapat lulus sesuai dengan waktu studi ?
2. Bagaimana melakukan perbandingan hasil dan akurasi data mining dan algoritma C4.5?

I.2.3. Batasan Masalah

Agar pembahasan terarah dan tidak menyimpang dari pokok permasalahan yang dibahas maka sistem ini dibatasi permasalahannya sebagai berikut :

1. Penerapan data mining algoritma C4.5 dalam prediksi kelulusan mahasiswa yang dapat lulus sesuai dengan waktu studi.
2. Analisa perbandingan hasil dan akurasi data mining dan algoritma C4.5.
3. Bahasa pemrograman yang digunakan untuk membuat aplikasi yaitu *Java Script* dengan IDE yang digunakan untuk mengembangkan aplikasi yaitu Netbeans 8.0.
4. *Database* yang digunakan yaitu *SQLite*.
5. Pemodelan sistem dilakukan dengan UML 2.0.

I.3. Tujuan Dan Manfaat

Tujuan dan manfaat yang penulis peroleh dari penelitian skripsi ini adalah sebagai berikut :

I.3.1. Tujuan

Tujuan penelitian ini yaitu:

1. mencari dan menemukan pola yang terdapat pada data mahasiswa berdasarkan data.
2. memprediksi mahasiswa yang lulus atau tidak lulus sesuai dengan waktu studi dengan menggunakan algoritma C4.5, kemudian membandingkan hasil dan akurasi algoritma tersebut.

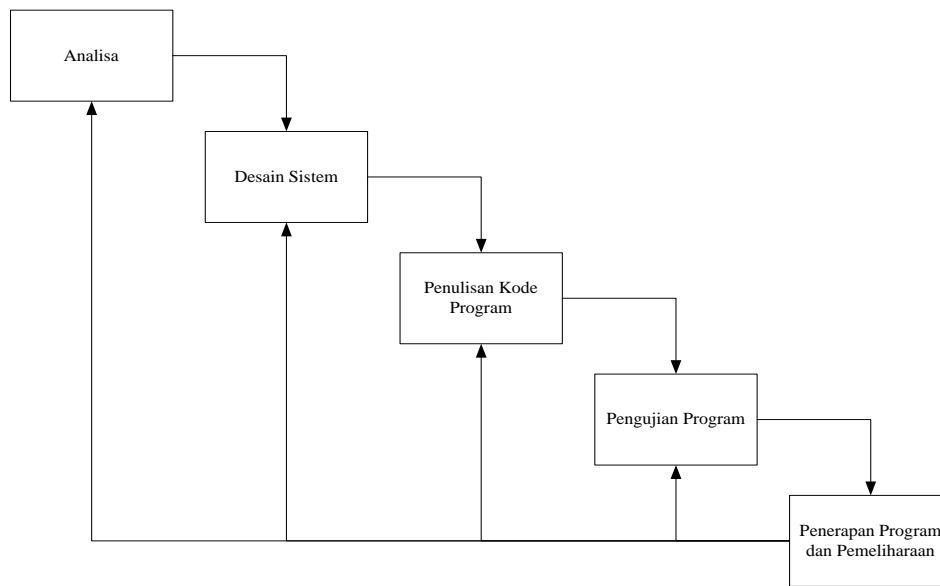
I.3.2. Manfaat

Manfaat penelitian ini yaitu:

1. Hasil penelitian dapat dijadikan bahan referensi oleh peneliti berikutnya yang akan membahas masalah mengenai sistem data mining.
2. Memberikan kemudahan bagi mahasiswa untuk mengetahui kelulusan.

I.4. Metodologi Penelitian

Didalam melakukan pengembangan sistem penulis menggunakan paradigma waterfall. Adapun metode waterfall mempunyai tahapan-tahapan sebagai berikut :



Gambar I.1. Waterfall

1.4.1. Analisa Kebutuhan

Adapun tahap yang dilakukan pada analisa kebutuhan yaitu mengumpulkan data data dari politeknik Ganesha. Pada tahapan ini untuk mengetahui data mining menggunakan algoritma C4.5 untuk prediksi ketepatan waktu kelulusan mahasiswa pada Politeknik Ganesha. Dengan menggunakan algoritma C4.5 akan diketahui waktu kelulusan mahasiswa. Adapun pengumpulan data dengan metode-metode sebagai berikut:

1. Studi lapangan

Merupakan teknik yang dilakukan dengan cara mengumpulkan data dengan mengadakan penelitian langsung terhadap objek penelitian dan pengumpulan data melalui :

2. Wawancara

Adalah salah satu metode atau suatu cara yang digunakan untuk memperoleh data dengan mengajukan pertanyaan-pertanyaan dengan pihak yang terkait yang dapat memberikan informasi khususnya pada penulis.

3. Observasi/pengamatan

Adalah pengumpulan data dengan mengadakan peninjauan langsung, mengkaji dan menganalisa terhadap prosedur-prosedur pada sistem waktu kelulusan mahasiswa pada Politeknik Ganesha.

Tabel I.1. Kebutuhan Sistem Fungsional

No	Kebutuhan	Rincian Kebutuhan
1.	Fungsi Sistem	– Sebagai <i>interface</i> penyampaian informasi
2.	Perangkat Lunak	– JavaScript – SQLite
3.	Pelaksana Sistem	– <i>User</i>
4.	Pengolah Sistem	– <i>Programmer</i>

Tabel I.2. Kebutuhan Sistem Nonfungsional

No	Kebutuhan	Rincian Kebutuhan
1.	Sistem Operasi	– Minimal Windows XP SP 2
2.	Prosesor	– Minimal Intel DualCore
3.	RAM	– Minimal 2GB
4.	Hardisk	– Minimal 120GB
5.	Monitor/LCD	– Minimal Resolusi 1024x768

1.4.2. Desain Sistem

Desain sistem ini dirancang dengan permodelan *UML* menggunakan *Microsoft Visio* 2010 yang digunakan untuk membuat desain sistem laporan nilai kelulusan mahasiswa.

1.4.3. Penulisan Kode Program

Penulisan kode program menggunakan *javascript*. Hal ini sangat memudahkan proses pasca perancangan kode program. Setelah pengkodean selesai maka akan dilakukan *testing* terhadap sistem yang telah dibuat tadi. Tujuan *testing* adalah menemukan kesalahan kesalahan terhadap sistem tersebut dan kemudian bisa diperbaiki.

1.4.4. Pengujian Program

Berisi langkah-langkah yang dilakukan dalam pembuatan alat serta tahapan-tahapan pengujian yang dilakukan untuk masing masing blok peralatan yang dirancang.

- a. Menganalisis beberapa kesalahan yang ada pada sistem yang lama.
- b. Melakukan pengujian aplikasi yang baru untuk meminimalisir kesalahan yang ada.
- c. Melakukan perawatan sistem yang baru apabila terjadi kesalahan..

1.4.5. Penerapan Program dan Pemeliharaan

Perangkat lunak yang merupakan suatu kegiatan untuk memelihara perangkat lunak yang sudah dibuat, pemeliharaan tersebut dilakukan agar keutuhan program dapat terjaga seperti validasi data, update data, dan integrasi data.

I.5. Keaslian Penelitian

Sebagai bukti penelitian yang akan dibuat, maka penelitian akan dibandingkan terhadap penelitian sejenis yang pernah dilakukan.

Tabel I.3. Keaslian Penelitian

No	Nama	Tahun	Judul	Hasil Penelitian
1	David Hartanto Kamagi, ULTIMATICS, Vol. VI, No. 1	2014	Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa	Data mining dengan algoritma C4.5 dapat diimplementasikan untuk memprediksi tingkat kelulusan mahasiswa dengan empat kategori yaitu lulus cepat, lulus tepat, lulus terlambat dan drop out
2	Dwi Untari	2010	Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5	Ratio data training yang digunakan mempengaruhi nilai akurasi pada setiap percobaan. Pada percobaan ke- 1 nilai akurasi terbaik adalah 92.79% dengan ratio data training 90%

3	Yosoa Putra Raharja	2014	RANCANG BANGUN SISTEM REKOMENDASI BEASISWA MENGUNAKAN ALGORITMA KLASIFIKASI C4.5 PADA UNIVERSITAS DIAN NUSWANTORO	Tingkat akurasi dari implementasi algoritma klasifikasi C4.5 dipengaruhi oleh beberapa hal seperti jenis, jumlah, isi dataset dan jumlah partisi data set. Dari hasil uji coba partisi data set ditemukan tingkat akurasi tertinggi pada jumlah partisi data sebanyak 90% dan menghasilkan akurasi sebanyak 92,31%. Dan semakin besar jumlah partisi data maka akan menghasilkan jumlah akurasi yang semakin tinggi pula.
---	------------------------	------	---	---

I.6. Lokasi Penelitian

Melakukan riset di Politeknik Ganesha Medan, Sumatera Utara. Telp : (061) 661-0910.

I.7. Sistematika Penulisan

Adapun sistematika penulisan yang diajukan dalam tugas akhir ini adalah sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini menerangkan tentang latar belakang, ruang lingkup permasalahan, tujuan dan manfaat, metode penelitian dan sistematika penulisan.

BAB II : TINJAUAN PUSTAKA

Pada bab ini menerangkan tentang teori-teori dan metode yang berhubungan dengan topik yang dibahas atau permasalahan yang sedang dihadapi.

BAB III : ANALISA DAN DESAIN SISTEM

Pada bab ini mengemukakan tentang analisa sistem yang sedang berjalan, evaluasi sistem yang berjalan dan desain sistem secara detail.

BAB IV : HASIL DAN UJI COBA

Pada bab ini menerangkan hasil dan pembahasan program yang dirancang serta kelebihan dan kekurangan sistem yang dirancang.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan penulisan dan saran dari penulis sebagai perbaikan di masa yang akan datang untuk sistem.

