

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Universitas Bina Darma merupakan salah satu universitas swasta yang ada di provinsi Sumatera Selatan. Universitas ini memiliki beberapa fakultas, fakultas teknik ilmu komputer merupakan salah satu dari beberapa fakultas yang ada. Setiap tahunnya mahasiswa baru fakultas ilmu komputer selalu bertambah namun jumlah yang lulus semakin berkurang.

Pada universitas Bina Darma untuk jenjang pendidikan Program Sarjana (S1) di fakultas Bina Darma mempunyai waktu delapan semester atau 4 tahun dengan jumlah minimal 124 sks yang harus diselesaikan, dihitung 4 tahun dengan perhitungan tahun lulus mahasiswa dikurangi tahun masuk. Namun mahasiswa yang mampu menyelesaikan studinya paling cepat 4 tahun dan selama ini mahasiswa Fakultas Ilmu Komputer yang lulus tepat waktu mengalami penurunan atau masih banyak mahasiswa Fakultas Ilmu Komputer yang tidak lulus tepat waktu yaitu sekitar 4,5 tahun bahkan lebih. Sks merupakan sistem kredit semester suatu program untuk jenjang pendidikan yang menggunakan semester sebagai waktu terkecil dan tidak mempengaruhi kenaikan tingkat[1]

Masing-masing mahasiswa setiap semester memiliki pengambilan jenis mata kuliah yang berbeda-beda, hal itu juga dapat membuat nilai indeks prestasi sementara setiap mahasiswa berbeda-beda. Nilai indeks prestasi semester ialah

nilai yang berupa dalam menentukan tingkat prestasi mahasiswa dalam setiap semester. Nilai ini digunakan untuk mengambil beban studi mahasiswa dalam semester berikutnya. Tujuan penelitian dari penjelasan diatas ialah untuk memprediksi tingkat kelulusan mahasiswa menggunakan Absensi, Pola Belajar, Organisasi, Aktif Elearning dan Semester Pendek. Pengolahan data tersebut di implementasikan dengan teknik data mining menggunakan metode *naïve bayes classifier*.

Data mining gabungan sejumlah disiplin ilmu kompuetr (ACM 2006), (Clifton 2010)[2], yang didefinisikan sebagai proses penemuan pola-pola baru dari kumpullan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence*, *machine learning*, *statistik* dan *database system* (ACM 2006). Terdapat banyak teknik dalam pengerjaan penerapan data mining, diantaranya *Decision tree*, *Algoritma naïve bayes*, *K-Nearest neighbour rule*, *Algoritma apriori*, dan masih banyak lagi. Dalam penelitian ini yang digunakan ialah *naïve bayes classifier* karena memiliki keuntungan proses yang cepat efisien ruang serta dapat diaplikasikan dengan masalah apa saja.

Naïve bayes merupakan teknik prediksi berbasis probablistik sederhana yang berdasar pada penerapan teorema bayes (atau aturan bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat (naif). Sistem ini dibuat bertujuan untuk membantu fakultas ilmu komputer dalam memantau mahasiswanya sejak dini dapat lulus tepat waktu atau tidak tepat waktu. Mahasiswa yang lulus tepat waktu dapat meningkatkan jumlah mahasiswa yang lulus. Dan hal tersebut mampu meningkatkan akreditasi fakultas komputer. Berdasarkan uraian diatas,

maka perlu dirancang suatu sistem yang diharapkan dapat membantu dan mempermudah dalam pemantauan setiap mahasiswanya. Hal ini juga yang menjadi latar belakang peneliti dalam melakukan penelitian yang berjudul “Implementasi Data Mining Menggunakan Metode Naïve Bayes untuk Memprediksi Ketepatan Waktu Tingkat Kelulusan Mahasiswa Fakultas Ilmu Komputer Universitas Bina Darma”.

Kamagi, D.H dan Hangsun, S (2014) dengan penelitian berjudul “Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Mahasiswa”. Penelitian ini untuk memprediksi kelulusan mahasiswa menggunakan Algoritma C4.5 sebagai sebuah referensi untuk membuat kebijakan dan tindakan akademik bidang(BAAK) dalam mengurangi mahasiswa yang tidak lulus dan lulus terlambat. Dalam penelitian ini kriteria yang diambil ialah IP semester 1-6, asal usul SMA, jenis kelamin dan jumlah kredit semester, dapat memprediksi mahasiswa dengan empat kategori lulus tepat, lulus cepat, terlambat dan drop out, menggunakan data mining Algoritma C4.5 dengan akurasi kelulusan prediksi yang diperoleh adalah 87,5%.

Nikmatun, I.A dan Waspada, I (2019) dengan penelitian berjudul “Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa menggunakan Algoritma *K-Nearest Neighbor*”. Penelitian ini untuk memprediksi lama studi mahasiswa <5 tahun atau >5 tahun, berdasarkan kriteria yang digunakan dan metode *Algoritma K-Nearest Neighbor*. Pada penelitian menggunakan atribut mata kuliah yang dibagi menjadi 2 mata kuliah pilihan dan mata kuliah wajib. Dengan menggunakan mata kuliah pilihan didapat nilai akurasi kelulusan prediksi diperoleh adalah 75,95%.

Berdasarkan uraian penelitian diatas peneliti akan melakukan penelitian yang berjudul “Implementasi Data Mining Menggunakan Metode Naïve Bayes untuk Memprediksi Ketepatan Waktu Tingkat Kelulusan Mahasiswa Fakultas Ilmu Komputer Universitas Bina Darma”. Penelitian ini mengimplementasikan metode Algoritma *Naïve Bayes Classifier* untuk memprediksi kelulusan mahasiswa fakultas ilmu komputer dengan ketentuan lulus tepat waktu atau terlambat. Mahasiswa yang lulus tepat waktu dapat membantu dalam menaikkan nilai akreditasi fakultas.

1.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini yaitu sebagai berikut :

1. Bagaimana teknik data mining dengan metode Naïve Bayes dapat menghitung dan menampilkan hasil dalam memprediksi kelulusan Mahasisa Bina Darma?
2. Berapa tingkat nilai akurasi dalam prediksi kelulusan menggunakan metode Naïve Bayes?

1.3 Ruang Lingkup dan Batasan Masalah

1. Semua proses dalam perhitungan yang digunakan melalui data mining metode Naïve Bayes.
2. Data mahasiswa yang akan testing menggunakan data angkatan 2018.
3. Jurusan yang menjadi objek penelitian ialah mahasiswa Fakultas Ilmu Komputer Universitas Bina Darma.
4. Prediksi yang dihasilkan ialah tepat atau tidak tepat waktunya dalam menyelesaikan masa studi.

5. Atribut yang digunakan Absensi, Pola Belajar, Organisasi, Aktif Elearning dan Semester Pendek.
6. Implementasi dan pengujian menggunakan program aplikasi.

1.4 Tujuan dan Manfaat Penelitian

1.4.2 Tujuan Penelitian

1. Untuk mengetahui dan menerapkan kegunaan data mining metode *naïve bayes classifier* dalam menyajikan informasi untuk memprediksi kelulusan berdasarkan data training menggunakan data alumni mahasiswa fakultas ilmu komputer.
2. Mendapatkan akurasi ketepatan kelulusan mahasiswa fakultas ilmu komputer menggunakan metode *naïve bayes classifier*.

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Bagi Penulis
 - a. Memberikan informasi prediksi kelulusan yang lulus tepat waktu ataupun yang tidak tepat waktu.
 - b. Menjadikan sebagai perbandingan bagi peneliti lain dalam menerapkan metode-metode data mining di area pendidikan.

2. Bagi Perusahaan

Adapun bagi perusahaan ialah untuk mempermudah dalam melihat perbandingan kelulusan tahun sebelumnya.

1.5 Metodologi Penelitian

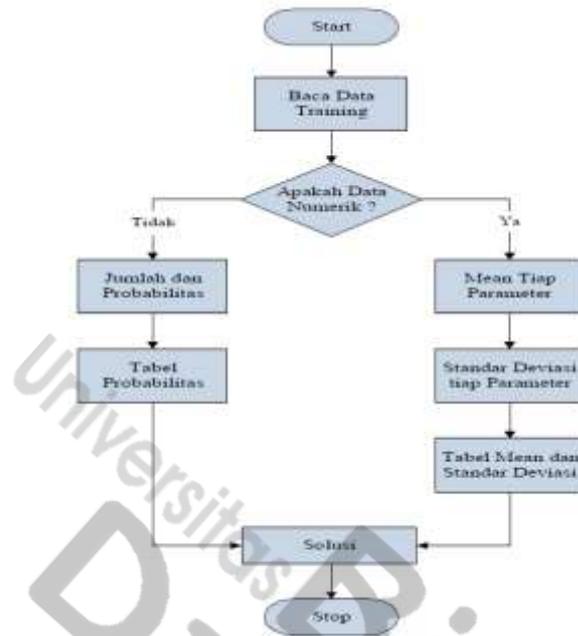
1.5.1 Penentuan Atribut

Dalam memprediksi kelulusan mahasiswa maka diperlukan data-data seluruh mahasiswa baik yang sudah lulus maupun yang belum lulus. Semua data-data diperoleh dari Direktorat Sistem Teknologi Informasi (DSTI) yang terletak di Universitas Bina Darma. Data yang digunakan sebagai data training ialah Absensi, Pola Belajar, Organisasi, Aktif Elearning dan Semester Pendek.

1.5.2 Penggunaan metode

Metode yang digunakan dalam penelitian ini ialah metode *Algoritma naïve bayes Classifier* untuk memprediksi masa studi mahasiswa fakultas ilmu komputer. Pemilihan metode ini dikarenakan relatif mudah digunakan, lebih efisien bila digunakan untuk memprediksi dalam jumlah yang sangat besar dan memiliki tingkat keakurasian yang relatif tinggi dalam hasil prediksi.

Berikut alur perhitungan metode *Naïve Bayes*:



Gambar 1. Alur Metode Naïve Bayes

1. Baca data training
2. Hitung jumlah dan probabilitas, namun apabila data numerik maka:
 - a. Cari nilai mean dan standar deviasi dari masing-masing parameter yang merupakan data numerik. Adapun persamaan yang digunakan untuk menghitung nilai rata-rata hitung (mean) dapat dilihat sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

atau

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

- b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.

3. Mendapatkan nilai dalam table mean, standart deviasi dan probabilistik
4. Solusi kemudian dihasilkan

1.6 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah:

a. Studi Pustaka (*Literature Review*)

Studi pustaka adalah metode mengumpulkan data dengan cara mencari informasi lewat buku, jurnal dan *literature* lainnya yang bertujuan membentuk sebuah landasan teori. Serta peneliti melakukan pencarian dan pembacaan buku-buku, jurnal atau referensi yang berhubungan dengan data mining, *naïve bayes classifier*, prediksi, weka dan lain-lain.

b. Observasi

Observasi adalah kegiatan yang dilakukan dengan sebuah pengamatan pada objek yang sedang diteliti, pengamatan ini dilakukan dengan mengamati IPK, asal sekolah, Nim, program studi, jenis kelamin, SKS, dan yang sudah lulus pada mahasiswa fakultas Ilmu Komputer.

c. Kuisisioner

[Kuesioner adalah](#) salah satu metode survei dalam melakukan penelitian yang dipakai untuk mengumpulkan data dari responden. Kuesioner berisi pertanyaan-pertanyaan untuk dijawab responden. Jawaban tersebut kemudian dikumpulkan, diolah dan dijadikan sebuah teori atau kesimpulan.

1.7 Metode Pengembangan

Dalam penelitian ini, penulis menggunakan metode pengembangan program aplikasi. Program aplikasi yang akan dibuat ialah suatu program yang menampilkan sebuah input hasil dari perhitungan metode naïve bayes dengan atribut yang sudah ada.

1.8 Lokasi dan Waktu Penelitian

1.8.1 Lokasi Penelitian

Penelitian ini dilakukan dikampus Utama Bina Darma yang beralamat di Jl. Jendral Ahmad Yani No.3, 9/10 Ulu, Kecamatan Seberang Ulu 1, Kota Palembang, Sumatera Selatan 30111.

1.8.2 Waktu Penelitian

Penelitian ini dilaksanakan mulai dari bulan Desember 2021.

1.9 Alat dan Bahan

Alat yang digunakan:

1. Perangkat keras

Perangkat keras yang peneliti gunakan dalam pembuatan program dan penulisan skripsi adalah:

- a. Processor Intel(R) Core(TM) i3-7020U

- b. RAM 4GB

2. Perangkat Lunak

Perangkat lunak yang digunakan adalah sebagai berikut:

- a. Windows 10 sebagai Operating Sistem
- b. Excel 2013
- c. WEKA

1.10 Sistematis Penulisan

BAB I PENDAHULUAN

Pada bab ini berisi latar belakang, identifikasi masalah, rumusan masalah, tujuan dan manfaat penelitian, metode penelitian, lokasi dan waktu penelitian, metode pengumpulan data, alat dan bahan penelitian serta sistematis penelitian skripsi.

BAB II LANDASAN TEORI

Pada bab ini akan menjelaskan mengenai teori-teori yang berhubungan dengan judul dan permasalahan yang akan diteliti.

BAB III METODOLOGI PENELITIAN

Pada bab ini penulis akan menjelaskan mengenai metodologi yang digunakan dalam penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini penulis akan membahas hasil-hasil yang diperoleh dalam penelitian dan pembahasan terhadap hasil yang telah dicapai maupun masalah-masalah yang telah diuraikan pada penelitian ini.

BAB V PENUTUP

Pada bab ini penulis akan menguraikan kesimpulan dari penelitian dan juga saran yang diberikan oleh penulis kepada pihak Universitas Bina Darma.

