

Aplikasi Pemrediksi Masa Studi Dan Predikat Kelulusan Mahasiswa Teknologi Informasi Menggunakan Metode Naive Bayes

1st Sudirman Sudirman
Teknologi Informasi
Universitas Bosowa
Makassar, Indonesia

<https://orcid.org/0000-0002-9892-8063>

2nd Yusrieh Mahendra
Teknologi Informasi
Universitas Bosowa
Makassar, Indonesia

yusriehmahendra9@gmail.com

3rd Rian Hadiwijaya Anwar
Teknologi Informasi
Universitas Bosowa
Makassar, Indonesia

rianbcc89@gmail.com

Abstract— teknologi informasi merupakan salah satu program studi di Universitas bosowa makassar yang memiliki data mahasiswa cukup besar baik data mahasiswa aktif maupun mahasiswa yang sudah lulus. Setiap tahun data tersebut semakin bertambah banyak. Di sisi lain, data yang dimiliki tersebut jika tidak dikelola dengan baik, maka hanya akan menjadi tumpukan data yang tidak bermanfaat, sehingga informasi yang dihasilkan juga tidak banyak. Salah satu cara untuk mengatasi masalah ini adalah data tersebut diubah menjadi sebuah informasi yang bersifat strategis. Oleh karena itu, tujuan penelitian ini adalah membuat suatu aplikasi untuk melakukan prediksi terhadap lama studi dan predikat kelulusan mahasiswa dengan menerapkan teknik data mining. Metode aplikasi yang digunakan dalam data mining untuk prediksi adalah algoritma naïve bayes. Hal ini digunakan untuk menganalisis data, terutama dalam proses pengenalan pola, memprediksi masa studi dan predikat kelulusan. Setelah pengolahan data, aplikasi akan menampilkan laporan serta laporan ringkasan. Hasil penelitian menunjukkan bahwa aplikasi dapat digunakan untuk membantu program studi informatika dalam rangka menemukan informasi strategis terkait dengan lama studi dan predikat kelulusan mahasiswa.

Kata Kunci : Data mining, Teknologi Informasi, metode naïve bayes, data base.

I. PENDAHULUAN

Dengan perkembangan teknologi informasi yang sangat pesat, informasi menjadi elemen penting dalam perkembangan kehidupan saat ini dan waktu mendatang. Data yang dihasilkan teknologi informasi sangat besar, termasuk bidang pendidikan, ekonomi, industry, dan berbagai bidang lainnya sehingga keperluan akan informasi yang tinggi tetapi tidak diimbangi dengan penyajian informasi yang memadai. Dari data yang jumlahnya sangat besar ini tentu belum digunakan secara optimal, sehingga perlu adanya penggalian ulang dari data yang jumlahnya sangat besar untuk menyajikan informasi yang dibutuhkan.

Dengan demikian data-data akademik mahasiswa yang ada juga cukup banyak. Jika data yang melimpah ini hanya dibiarkan menumpuk, maka hanya akan menjadi beban database yang dimiliki. Sementara itu, data-data yang melimpah ini sebenarnya bisa dimanfaatkan sebagai sumber informasi strategis bagi program studi untuk melakukan klasifikasi masa studi lulusan mahasiswa dengan menggunakan teknik data mining. Hal ini tentunya selain dapat memberikan informasi yang bersifat strategis bagi

fakultas dan program studi, juga dapat meningkatkan upaya untuk mendorong dan mempercepat kelulusannya. Sehingga selain dapat bermanfaat bagi mahasiswa sendiri, juga dapat meningkatkan nilai akreditasi bagi program studi.

Menurut buku panduan akademik Fakultas Teknik dan Teknologi Informasi Universitas Bosowa, dalam BAB II disebutkan bahwa “Beban studi mahasiswa program Strata satu (S-1) yang harus ditempuh minimal 144 satuan kredit mahasiswa (SKS) dan maksimal 146 satuan kredit mahasiswa (SKS), dengan batas waktu kurang dari 8 semester dan waktu paling lama 14 semester”. “Predikat Kelulusan untuk mahasiswa sarjana Strata satu (S-1) yaitu : Memuaskan, Sangat Memuaskan, dan Dengan Pujian (Cumlaude)”.

II. PERNYATAAN MASALAH

Berdasarkan permasalahan tersebut, dalam penelitian ini akan dibangun sebuah aplikasi untuk membantu program studi dalam memprediksi lama studi dan predikat kelulusan mahasiswa dengan teknik data mining yaitu menggunakan metode Naïve Bayes. Aplikasi tersebut diharapkan dapat digunakan untuk menemukan informasi tentang lama studi dan predikat kelulusan sehingga dapat membantu program studi Teknologi Informasi dalam mencari solusi dan kebijakan untuk meningkatkan prestasi mahasiswa agar dapat menyelesaikan studinya tepat waktu

III. TINJAUAN PUSTAKA

Penelitian [1] dari Muh Amin Nurrohmat “Aplikasi Pemrediksi Masa Studi dan Predikat Kelulusan Mahasiswa Informatika Universitas Muhammadiyah Surakarta Menggunakan Metode Naive Bayes, Menggunakan metode Pengumpulan data dan penentuan atribut Data Mining ini dapat digunakan untuk menampilkan informasi hasil prediksi lama studi dan predikat kelulusan dalam bentuk tabel dengan menggunakan data mahasiswa yang telah lulus sebagai data training dan data mahasiswa yang masih aktif sebagai data testing. Variabel bebas (X) yang terdiri dari jurusan asal sekolah, gender, daerah asal, asal sekolah, dan asisten dapat digunakan sebagai variabel untuk memprediksi lama studi dan predikat kelulusan mahasiswa. Algoritma naïve bayes dapat digunakan untuk memprediksi suatu nilai class variabel

dalam data testing berdasarkan perbandingan nilai probabilitas masing-masing class variabel pada data training terhadap variabel X yang diajukan dalam data testing.

Penelitian [2] dari Hartatik “Optimasi model prediksi kelulusan mahasiswa menggunakan algoritma naïve bayes” Metode penelitian ini pendekatan yang dilakukan yaitu kuantitatif, penelitian dilaksanakan berdasarkan data pengamatan berdasarkan populasi dan sampel Teknik pengambilan sampel dilakukan dengan mengambil beberapa mahasiswa pada Angkatan tertentu dan dalam program studi tertentu yang mana selanjutnya diterapkan untuk menguji [4][5]. Adapun Langkah-langkah penelitian ini adalah, mengumpulkan hasil pengamatan, seleksi data, Teknik penelitian yang digunakan, penelitian dan pengujian algoritma, dan evaluasi.

Penelitian [3] dari Budanis Dwi Meilani “Aplikasi data mining untuk menghasilkan pola kelulusan siswa dengan metode naïve bayes” Metode naïve bayes classifier merupakan salah satu metode pengklasifikasi berpeluang sederhana yang berdasarkan pada penerapan teorema bayes dengan asumsi antar variable penjelas saling bebas (independen). Algoritma ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Dua kelompok peneliti, satu oleh pantel dan lin, dan yang lain oleh Microsoft research memperkenalkan metode statistik Bayesian ini pada teknologi anti spam filter. Tetapi yang membuat algoritma Bayesian filtering ini populer adalah pendekatan yang dilakukan oleh graham.

IV. METODOLOGI

PENGUMPULAN DATA DAN PENENTUAN ATRIBUT
 Dalam melakukan penelitian ini diperlukan data dari seluruh mahasiswa Teknologi Informasi baik yang sudah lulus maupun yang belum lulus. Semua data diperoleh dari Biro Administrasi Akademik dan Laboratorium Fakultas. Data yang digunakan dalam penelitian ini berhubungan dengan data kelulusan dan data akademik mahasiswa Teknologi Informasi Unibos.

Penentuan Nilai Class Data Nilai class data penelitian yang telah dikumpulkan kemudian diubah menyesuaikan kebutuhan data mining. Nilai class pada atribut Lama Studi dikategorikan berdasarkan semester yang ditempuh pada saat lulus, yaitu :

- 1). Tepat waktu, jika lama studi 8 semester atau kurang dari 8 semester.
- 2).Terlambat, jika lama studi lebih dari 8 semester. Sedangkan nilai class pada atribut Predikat Kelulusan ditentukan berdasarkan IPK yang dibagi menjadi tiga, yaitu: 1).Memuaskan, jika IPK antara 2,00 dan 2,75 2) Sangat Memuaskan, jika IPK antara 2,76 dan 3,50 3) Cumlaude, jika IPK antara 3,51 dan 4,00

Analisis Data Tahap analisis data dilakukan dengan perhitungan menggunakan algoritma *Naïve Bayes*. Algoritma naïve bayes merupakan pengklasifikasian dengan metode probabilitas dan statistic untuk memprediksi peluang di masa

depan. Adapun algoritma ini dapat dihitung menggunakan persamaan 1.

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

4.1. Perancangan Sistem

Berikut ini perancangan system dalam bentuk use case diagram, untuk melakukan analisis data mahasiswa terkait dengan kelulusannya.

V. HASIL

1. MENGHITUNG PROBABILITAS PRIOR MASING-MASING CLASS VARIABEL LAMA STUDI

Pada tahap ini dilakukan perhitungan, persamaan (2) dan (3), untuk mencari nilai probabilitas masing masing class yaitu “tepat” dan “terlambat”.

$$P(Y1 = \text{Tepat}) = \frac{\text{Jumlah class 'Tepat'}}{\text{jumlah data}} = \frac{4}{10}$$

$$P(Y1 = \text{Terlambat}) = \frac{\text{Jumlah class 'Terlambat'}}{\text{jumlah data}} = \frac{6}{10}$$

2. MENGHITUNG PROBABILITAS BERSYARAT SESUAI NILAI CLSS YANG DIAJUKAN DALAM DATA TESTING MENGGUNAKAN PERHITUNGAN NAÏVE BAYES.

Algoritma naïve bayes di gunakan untuk mencari nilai confidence atau peluang terjadinya masing-masing class variable Y1 berdasarkan variable X yang di ajkukan, persamaan (9) dan (10), nilai confidence masing-masing class tersebut kemudian dibandingkan untuk mengetahui peluang yang lebih besar untuk memprediksi nilai class yang muncul pada variable Y1

- a. Nilai confidence pada variable Y1 untuk class= ‘tepat’
 $P(X1=IPA.X2=Pria.X3= WIB.X5=y_a(Y1=Tepat))$

$$= \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{4}{4} \cdot \frac{4}{4} \cdot \frac{3}{4} \cdot \frac{4}{10} = 0,168$$

- b. Nilai confidence pada variable Y1 untuk class= ‘terlambat’
 $P(X1=IPA.X2=Pria.X3= WIB.X5=y_a(Y1=terlambat))$

$$= \frac{2}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{2}{6} \cdot \frac{3}{6} \cdot \frac{6}{10} = 0,00432$$

3. PROBALITAS BERSYARAT SESUAI NILAI CLASS YANG DIAJUKAN DALAM DATA TESTING MENGGUNAKAN PERHITUNGAN NAÏVE BAYES

Nilai confidence masing-masing class dalam variable Y2, persamaan (22)-(24), kemudian dibandingkan untuk mengetahui peluang yang lebih besar untuk memprediksi nilai class yang muncul pada variable Y2 tersebut.

- a. Nilai confidence pada variable Y2 untuk class= 'coumlaude'

$P(X1=IPA, X2=Pria, X3=WIB, X5=Ya | Y2=coumlaude)$

$$= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{3}{3} \cdot \frac{3}{3} \cdot \frac{2}{3} \cdot \frac{3}{10} = 0,097746$$

- b. Nilai confidence pada variable Y2 untuk class='sangat memuaskan'

$P(X1=IPA, X2=Pria, X3=WIB, X5=Ya | Y2=sangat memuaskan)$

$$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{2} \cdot \frac{3}{3} \cdot \frac{2}{3} \cdot \frac{3}{10} = 0,025$$

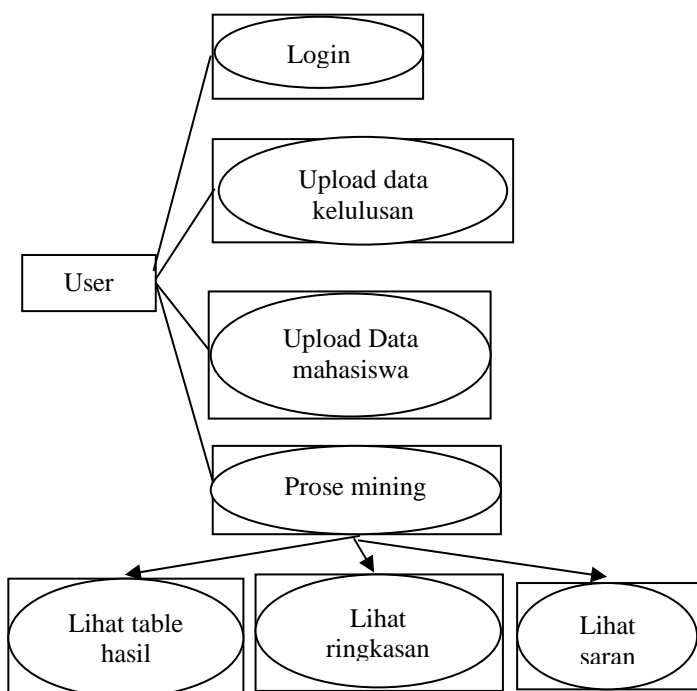
- c. Nilai confidence pada variabel Y2 untuk class 'memuaskan'

$P(X1=IPA, X2=Pria, X3=WIB, X4=WIB, X5=Ya | Y2=memuaskan)$

$$= \frac{2}{5} \cdot \frac{5}{5} \cdot \frac{0}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{5}{10} = 0$$

Berdasarkan nilai confidence masing-masing class pada variabel Y2 dapat dilihat bahwa nilai confidence untuk class= 'coumlaude' lebih besar daripada nilai confidence confidence pada class= 'sangat memuaskan' dan class= 'memuaskan' dengan demikian, prediksi class variabel Y2 pada data testing adalah "caumlaude".

Implementasi dimulai dari tahap pembuatan use case diagram seperti di bawah ini.



Gambar 1. Use Case Diagram dengan User sebagai Aktor

Tabel 1. Keterangan use case diagram

Nama Case	Use	Aktor	Deskripsi
Loging		user	Menggambarkan kegiatan input username dan password
Ambil data training		User	Menggambarkan kegiatan mengambil data kelulusan
Ambil data tasting		User	Menggambarkan kegiatan mengambil data mahasiswa aktif
Proses Mining		User	Menggambarkan kegiatan proses mining atribut lama studi dan IPK
Laporan hasil maining		user	Menggambarkan kegiatan menampilkan hasil mining
Laporan Rangkuman		User	Menggambarkan kegiatan menampilkan rangkuman dari proses mining
Laporan saran		user	Menggambarkan kegiatan melihat saran yang diberikan untuk prodi

Hasil dari penelitian ini adalah sebuah aplikasi data maining untuk memprediksi lama studi dan predikat kelulusan mahasiswa program studi Teknologi Informasi menggunakan metode *Naïve Bayes* hasil nya seperti di bawah ini'

VI. KESIMPULAN

Berdasarkan hasil penelitian dan uji coba yang dilakukan maka dapat disimpulkan bahwa aplikasi DATA MAINING ini dapat di gunakan untuk menampilkan informasi hasil prediksi lama studi dan predikat kelulusan dalam bentuk table dengan menggunakan data mahasiswa yang telah lulus. Algoritma *Naïve Bayes* dapat digunakan untuk memprediksi suatu nilai class variable dalam data testing berdasarkan perbandingan nilai probabilitas masing-masing class variable pada data training terhadap variable X yang diajukan dalam data testing.

DAFTAR PUSTAKA

- [1]. Nurrohmat, M. A. (2015). Aplikasi Pemrediksi Masa Studi dan Predikat Kelulusan Mahasiswa Informatika Universitas Muhammadiyah Surakarta Menggunakan Metode Naive Bayes. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, 1(1), 29-34.
- [2]. Hartatik, H. Optimasi Model Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Naive Bayes. *IJAI (Indonesian Journal of Applied Informatics)*, 5(1), 32-38.
- [3]. Sudirman, S. (2021). Machine Learning Deteksi Jatuh Menggunakan Algoritma Human Posture Recognition. *Proceeding KONIK (Konferensi Nasional Ilmu Komputer)*, 5, 462-466.
- [4]. Sudirman, S., Zainuddin, Z., & Suyuti, A. (2021). Fall Detection in the Elderly With Android Mobile IoT Devices Using Nodemcu And Accelerometer Sensors.
- [5]. Meilani, B. D., & Susanti, N. (2016). Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naive Bayes. *Network Engineering Research Operation*, 1(3).