

# Penerapan Algoritma Naive Bayes untuk Klasifikasi Kelulusan Mahasiswa Pendidikan Teknik Informatika dan Komputer UIN SMDD Bukittinggi

Sri Atiqah Elvidamayan\*<sup>1</sup>, Liza Efriyanti<sup>2</sup>, Sarwo Derta<sup>3</sup>, Tasnim Rahmat<sup>4</sup>

<sup>1</sup>Pendidikan Teknik Informatika dan Komputer

<sup>2,3,4</sup>Fakultas Tarbiyah dan Ilmu Keguruan, UIN Sjech M.Djmail Djambek Bukittinggi

Email: <sup>1</sup>sri.atiqah.elvidamayanti@gmail.com, <sup>2</sup>lizaefriyanti@uinbukittinggi.ac.id,

<sup>3</sup>sarwoderta@iainbukittinggi.ac.id, <sup>4</sup>tasnimrahmat@uinbukittinggi.ac.id

\*Penulis Korespondensi

Diterima: 11 September 2025; Direvisi: 14 Mei 2026; Disetujui: 15 Mei 2026

## Abstrak

Ketepatan waktu kelulusan mahasiswa merupakan indikator penting dalam evaluasi kinerja perguruan tinggi, namun pemanfaatan data akademik untuk prediksi masih terbatas. Penelitian ini bertujuan untuk mengklasifikasikan ketepatan waktu kelulusan mahasiswa menggunakan pendekatan data mining dengan algoritma Naive Bayes Classifier. Metode yang digunakan adalah kuantitatif dengan tahapan preprocessing data, pemodelan klasifikasi, dan evaluasi menggunakan confusion matrix. Data yang dianalisis meliputi jenis kelamin, Indeks Prestasi Semester (IPS 1–8), Indeks Prestasi Kumulatif (IPK), serta sumber pendanaan mahasiswa. Hasil penelitian menunjukkan bahwa model memperoleh akurasi sebesar 87,5%, precision sebesar 57,2%, dan recall sebesar 80%. Variabel IPS dan IPK menjadi faktor dominan dalam menentukan ketepatan waktu kelulusan, sedangkan sumber pendanaan turut memberikan kontribusi terhadap hasil klasifikasi. Temuan ini menunjukkan bahwa algoritma Naive Bayes cukup efektif dalam mengidentifikasi pola kelulusan mahasiswa dan berpotensi mendukung pengambilan keputusan berbasis data di lingkungan pendidikan tinggi. Meskipun demikian, performa model masih dipengaruhi oleh ketidakseimbangan data dan keterbatasan jumlah dataset.

**Kata kunci:** data mining, klasifikasi, kelulusan mahasiswa, naive bayes, pendidikan tinggi

## Abstract

Timely graduation is an important indicator in evaluating higher education performance; however, the use of academic data for prediction remains limited. This study aims to classify student graduation timeliness using a data mining approach with the Naive Bayes Classifier algorithm. A quantitative method was applied, including data preprocessing, classification modeling, and evaluation using a confusion matrix. The dataset consists of gender, Semester Grade Point Average (GPA for semesters 1–8), Cumulative GPA, and funding source. The results show that the model achieved an accuracy of 87.5%, precision of 57.2%, and recall of 80%. Semester GPA and cumulative GPA were identified as dominant factors influencing graduation timeliness, while funding source also contributed to the classification results. These findings indicate that the Naive Bayes algorithm is effective in identifying graduation patterns and can support data-driven decision-making in higher education. However, the model performance is affected by data imbalance and limited dataset size.

**Keywords:** classification, data mining, higher education, naive bayes, student graduation

## I. PENDAHULUAN

Pendidikan tinggi memiliki peran penting dalam menghasilkan sumber daya manusia yang berkualitas, yang salah satunya dapat diukur melalui capaian akademik mahasiswa. Menurut standar Badan Akreditasi

Nasional Perguruan Tinggi (BAN-PT), indikator seperti ketepatan waktu kelulusan dan prestasi akademik menjadi faktor utama dalam menilai kualitas suatu perguruan tinggi [1] [2]. Namun, dalam praktiknya masih banyak institusi pendidikan yang mengalami kesulitan dalam memprediksi tingkat kelulusan mahasiswa secara akurat dan sistematis [3].

Permasalahan utama terletak pada pemanfaatan data akademik yang belum optimal. Data mahasiswa yang tersimpan setiap tahun cenderung hanya menjadi arsip tanpa dianalisis lebih lanjut untuk menghasilkan informasi yang bersifat prediktif [4]. Padahal, data akademik memiliki potensi besar untuk digunakan dalam mengidentifikasi pola kelulusan mahasiswa sehingga dapat membantu perguruan tinggi dalam mengambil keputusan akademik yang lebih tepat [5]. Kondisi tersebut menunjukkan adanya kebutuhan akan penerapan analisis berbasis data yang mampu mengolah informasi akademik menjadi pengetahuan yang bermanfaat bagi institusi Pendidikan [6].

Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah penerapan teknik Data Mining. Data mining merupakan proses analisis data untuk menemukan pola, hubungan, dan informasi yang sebelumnya tidak diketahui dari kumpulan data berukuran besar [7]. Dalam proses *Knowledge Discovery in Databases* (KDD), data mining melibatkan beberapa tahapan seperti seleksi data, pembersihan data (*data cleaning*), transformasi data, proses analisis, serta interpretasi hasil [8]. Teknik ini banyak dimanfaatkan dalam berbagai bidang, termasuk pendidikan, karena mampu membantu pengambilan keputusan berbasis data.

Dalam data mining terdapat beberapa metode utama, seperti klasifikasi, *clustering*, asosiasi, dan regresi [9]. Penelitian ini menggunakan teknik klasifikasi, yaitu metode yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan pola dari data sebelumnya [10]. Proses klasifikasi umumnya terdiri atas tahap pelatihan (*training*) dan tahap pengujian (*testing*). Pada tahap *training*, sistem membangun model berdasarkan data yang telah diketahui kelasnya, sedangkan pada tahap *testing* model diuji untuk mengetahui tingkat akurasi dalam memprediksi data baru. Dalam konteks pendidikan, klasifikasi dapat digunakan untuk memprediksi tingkat kelulusan mahasiswa berdasarkan data akademik seperti nilai mata kuliah, Indeks Prestasi Kumulatif (IPK), jumlah SKS, dan lama masa studi [11].

Metode klasifikasi yang digunakan dalam penelitian ini adalah Naïve Bayes Classifier. Naïve Bayes merupakan algoritma klasifikasi berbasis teori probabilitas Bayesian yang menghitung peluang suatu data termasuk ke dalam kelas tertentu berdasarkan atribut yang dimiliki [12]. Metode ini memiliki asumsi bahwa setiap atribut bersifat independen satu sama lain (*conditional independence*) [10]. Meskipun asumsi tersebut sederhana, Naïve Bayes tetap mampu memberikan hasil klasifikasi yang baik dalam berbagai kasus. Tahapan metode ini meliputi penentuan probabilitas awal (*prior probability*), perhitungan likelihood setiap atribut terhadap kelas tertentu, perhitungan probabilitas posterior, serta penentuan kelas dengan probabilitas tertinggi sebagai hasil klasifikasi [13].

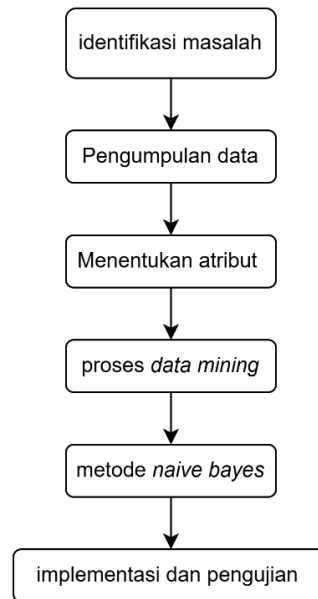
Naïve Bayes memiliki beberapa keunggulan, antara lain efisiensi komputasi, kemudahan implementasi, serta kemampuan dalam menangani dataset berukuran besar [14]. Oleh karena itu, metode ini banyak digunakan dalam berbagai penelitian, termasuk pada bidang Pendidikan [15]. Melalui penerapan metode klasifikasi menggunakan Naïve Bayes, pola kelulusan mahasiswa dapat diidentifikasi secara lebih sistematis sehingga perguruan tinggi dapat memahami faktor-faktor yang mempengaruhi keberhasilan studi mahasiswa [16].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan algoritma Naïve Bayes dalam mengklasifikasikan tingkat kelulusan mahasiswa berdasarkan data akademik. Penelitian ini diharapkan dapat menghasilkan model prediksi yang akurat serta memberikan kontribusi dalam pemanfaatan data akademik sebagai dasar pengambilan keputusan di lingkungan perguruan tinggi.

## II. METODE PENELITIAN

Pada bagian ini dapat diuraikan mengenai hasil dari penelitian beserta pengujian yang telah dilakukan. Selain itu, disampaikan juga mengenai pembahasan dari penelitian maupun pengujian yang telah dilakukan. Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif, yaitu data dalam bentuk numerik yang merepresentasikan nilai dari variabel-variabel yang diteliti. Oleh karena itu, penelitian ini menggunakan pendekatan kuantitatif yang berfokus pada analisis data numerik untuk menghasilkan kesimpulan yang objektif dan terukur [17].

Tahapan penelitian meliputi identifikasi masalah, pengumpulan data, penentuan atribut, proses data mining, serta penerapan model Naive Bayes Classifier. Data yang digunakan mencakup informasi akademik mahasiswa, seperti identitas dasar dan kinerja akademik selama masa studi di perguruan tinggi.



**Gambar 1.** Alur penelitian

Penelitian ini dilakukan melalui beberapa tahapan sistematis sebagai berikut:

#### **A. Identifikasi Masalah**

Tahap awal penelitian bertujuan untuk mengkaji permasalahan terkait prediksi ketepatan waktu kelulusan mahasiswa. Fokus penelitian diarahkan pada penerapan algoritma Naive Bayes Classifier untuk mengklasifikasikan mahasiswa ke dalam kategori lulus tepat waktu atau terlambat.

#### **B. Pengumpulan Data**

Data diperoleh dari Program Studi Pendidikan Teknik Informatika dan Komputer (PTIK), yang mencakup data identitas mahasiswa serta data akademik selama masa perkuliahan.

#### **C. Penentuan Atribut**

Atribut yang digunakan dalam proses klasifikasi meliputi:

- 1) Jenis kelamin (laki-laki dan perempuan)
- 2) Indeks Prestasi Semester (IPS 1–8), yang dikategorikan menjadi: Tinggi ( $>3,50$ ), Cukup ( $2,75-3,50$ ), Rendah ( $<2,75$ )
- 3) Indeks Prestasi Kumulatif (IPK), yang di kategorikan menjadi: Dengan Pujian ( $3.5 - 4.00$ ), Sangat Memuaskan ( $2.76 - 3.5$ ), Memuaskan ( $2.00 - 2.75$ )
- 4) Variabel keputusan: Kategori kelulusan (tepat waktu dan tidak tepat waktu)

#### **D. Proses Data Mining**

Tahapan pengolahan data dilakukan melalui beberapa langkah, yaitu:

- 1) Data *Cleaning*: Untuk menghilangkan data yang tidak lengkap, tidak konsisten, atau redundan.
- 2) Data Selection: Untuk memilih atribut yang relevan dengan tujuan penelitian.
- 3) Data Transformation: Untuk mengubah data ke dalam format yang sesuai untuk analisis.
- 4) Proses Klasifikasi: Menerapkan algoritma Naïve Bayes untuk mengidentifikasi pola dalam data [18].

#### **E. Penerapan Metode Naïve Bayes**

Algoritma Naïve Bayes merupakan metode klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dengan asumsi independensi antar atribut [19].

Persamaan (1) berikut adalah dasar Naïve Bayes

$$P(C | X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (1)$$

Keterangan:

- $P(C | X)$ = probabilitas posterior (kelas C terhadap data X)
- $P(X | C)$ = probabilitas likelihood
- $P(C)$ = probabilitas prior
- $P(X)$ = probabilitas evidence

Karena  $P(X)$  bersifat konstan, maka perhitungan dapat disederhanakan menjadi persamaan (2) berikut:

$$P(C | X) \propto P(C) \cdot \prod_{i=1}^n P(x_i | C) \quad (2)$$

Dimana:

- $x_i$ = atribut ke-i
- $n$ = jumlah atribut

Langkah Perhitungan:

Menghitung probabilitas prior setiap kelas dalam persamaan (3) berikut:

$$P(C) = \frac{\text{jumlah data pada kelas } C}{\text{total data}} \quad (3)$$

Menghitung likelihood dalam persamaan (4) berikut:

$$P(x_i | C) = \frac{\text{jumlah atribut } x_i \text{ pada kelas } C}{\text{jumlah data pada kelas } C} \quad (4)$$

Evaluasi dilakukan menggunakan *confusion matrix*, dengan metrik sebagai berikut [20]:

Akurasi dalam persamaan (5) berikut:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

*Precision* dalam persamaan (6) berikut:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

*Recall* dalam persamaan (7) berikut:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

Keterangan:

- $TP$  (*True Positive*)
- $TN$  (*True Negative*)
- $FP$  (*False Positive*)
- $FN$  (*False Negative*)

### III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data akademik mahasiswa sebanyak 75 data latih dan 8 data uji dari total 83 data. Variabel yang digunakan meliputi jenis kelamin, Indeks Prestasi Semester (IPS 1–8), Indeks Prestasi Kumulatif (IPK), serta status kelulusan sebagai variabel target, yaitu *tepat waktu* dan *terlambat*.

#### A. Perhitungan Probabilitas Awal (Prior)

Berdasarkan data latih, diperoleh distribusi kelas yaitu kelas tepat waktu sebanyak 66 data, kelas terlambat sebanyak 9 data. Sehingga probabilitas awal dihitung sebagai:

$$P(\text{Tepat}) = 66/75 = 0,88$$

$$P(\text{Terlambat}) = 9/75 = 0,12$$

Nilai ini menunjukkan bahwa sebagian besar mahasiswa dalam dataset cenderung lulus tepat waktu.

## B. Menghitung Probabilitas Atribut

Pada proses perhitungan pada atribut dari masing-masing kategori yang digunakan untuk penelitian. dilihat pada masing-masing atribut kelas kemudian diterapkan proses menghitung jumlah label data yang Tepat dan juga jumlah label data Terlambat.

- 1) Atribut Jenis Kelamin:
  - “Tepat”
    - a) LK =  $14 / 18 = 0,777777$
    - b) PR =  $52 / 57 = 0,912280$
  - “Terlambat”
    - a) LK =  $4 / 18 = 0,222222$
    - b) PR =  $5 / 57 = 0,087719$
- 2) IPS 1:
  - IPS 1
    - a) T - Tepat =  $34 / 66 = 0,515151$
    - b) T - Terlambat =  $1 / 9 = 0,111111$
  - IPS 1 Cukup
    - a) C - Tepat =  $32 / 66 = 0,484848$
    - b) C - Terlambat =  $7 / 9 = 0,777777$
  - IPS 1 Rendah
    - a) R - Tepat =  $0 / 66 = 0$
    - b) R - Terlambat =  $1 / 9 = 0,111111$
- 3) IPS 2:
  - IPS 2
    - a) T - Tepat =  $60 / 66 = 0,909090$
    - b) T - Terlambat =  $3 / 9 = 0,333333$
  - IPS 2 Cukup
    - a) C - Tepat =  $4 / 66 = 0,060606$
    - b) C - Terlambat =  $8 / 9 = 0,888888$
  - IPS 2 Rendah
    - a) R - Tepat =  $0 / 66 = 0$
    - b) R - Terlambat =  $0 / 9 = 0$
- 4) IPS 3:
  - IPS 3
    - a) T - Tepat =  $53 / 66 = 0,803030$
    - b) T - Terlambat =  $1 / 9 = 0,111111$
  - IPS 3 Cukup
    - a) C - Tepat =  $12 / 66 = 0,181818$
    - b) C - Terlambat =  $7 / 9 = 0,777777$
  - IPS 3 Rendah
    - a) R - Tepat =  $1 / 66 = 0,0151515$
    - b) R - Terlambat =  $1 / 9 = 0,111111$
- 5) IPS 4:
  - IPS 4
    - a) T - Tepat =  $56 / 66 = 0,848484$
    - b) T - Terlambat =  $1 / 9 = 0,111111$
  - IPS 4 Cukup
    - a) C - Tepat =  $11 / 66 = 0,166667$
    - b) C - Terlambat =  $5 / 9 = 0,555556$

- IPS 4 Rendah
- R -Tepat =  $0 / 66 = 0$
  - R-Terlambat =  $2 / 9 = 0,222222$
- 6) IPS 5:
- IPS 5 Tinggi
- T -Tepat =  $58/ 66 = 0,878787$
  - T -Terlambat =  $3 / 9 = 0,333333$
- IPS 5 Cukup
- C-Tepat =  $6/ 66 = 0,090909$
  - C -Terlambat =  $5/ 9 = 0,555555$
- IPS 5 Rendah
- R -Tepat =  $2 / 66 = 0,030303$
  - R -Terlambat =  $1 / 9 = 0,111111$
- 7) IPS 6:
- IPS 6 Tinggi
- T -Tepat =  $61/ 66 = 0,924242$
  - T -Terlambat =  $5 / 9 = 0,555555$
- IPS 6 Cukup
- C -Tepat =  $4/ 66 = 0,060606$
  - C -Terlambat =  $4/ 9 = 0,444444$
- IPS 6 Rendah
- R -Tepat =  $1 / 66 = 0,0151515$
  - R -Terlambat =  $0 / 9 = 0$
- 8) IPS 7:
- IPS 7 Tinggi
- T -Tepat =  $65/ 66 = 0,984848$
  - T -Terlambat =  $8 / 9 = 0,888888$
- IPS 7 Cukup
- C-Tepat =  $0/ 66 = 0$
  - C-Terlambat =  $1/ 9 = 0,111111$
- IPS 7 Rendah
- R- Tepat =  $1 / 66 = 0,0151515$
  - R-Terlambat =  $0/ 9 = 0$
- 9) IPS 8:
- IPS 8 Tepat
- T-Tepat =  $62/ 66 = 0,939393$
  - T -Terlambat =  $6 / 9 = 0,666666$
- IPS 8 Cukup
- C-Tepat =  $1/ 66 = 0,151515$
  - C-Terlambat =  $3/9 = 0,3333333$
- IPS 8 Rendah
- R-Tepat =  $3/66 = 0,045454545$
  - R-Terlambat =  $0/9 = 0$
- 10) Atribut IPK:
- Dengan Pujian – Tepat =  $65/66= 0,98484848$
  - Dengan Pujian – Terlambat =  $1/9 = 0,11111111$
  - Sangat Memuaskan – Tepat =  $1/66 = 0,015151515$
  - Sangat Memuaskan – Terlambat =  $8/9 = 0,88888888$
  - Memuaskan – Tepat =  $0/66 = 0$
  - Memuaskan – Terlambat =  $0/9 = 0$

Tahap selanjutnya adalah menghitung data testing yaitu dengan 8 data testing dari 83 total data. Hasil Data testing yang dilakukan perhitungan dapat dilihat pada Tabel 1 berikut :

**Tabel 1.** Hasil *data testing*

No	Nama	Tepat	Terlambat	Prediksi	Fakta	Masa Studi
1	Ely Nurhalizah Nst	0,0074287612	0,0000001630	Tepat	Tepat	4 Tahun 0 Bulan
2	Yustimudiansyah	0,0000000037	0,0005063627	Terlambat	Terlambat	5 Tahun 0 Bulan
3	Armita Sari	0,0000009603	0,0042602594	Terlambat	Terlambat	4 Tahun 3 Bulan
4	M. Ali Akbar	0,0063334937	0,0000016954	Tepat	Tepat	4 Tahun 0 Bulan
5	Rozatul Nasihin	0,0000008023	0,0132920205	Terlambat	Terlambat	4 Tahun 3 Bulan
6	Devi Yuriza	0,0074287612	0,0000001630	Tepat	Tepat	4 Tahun 0 Bulan
7	Rahman Saleh Pulungan	0,0000044182	0,0007911917	Tepat	Terlambat	4 Tahun 3 Bulan
8	Ely Nurhalizah Nst	0,2027547925	0,0000034234	Tepat	Tepat	4 Tahun 0 Bulan

Selanjutnya dilakukan pengujian dari data yang digunakan. Dilanjutkan dengan melakukan evaluasi memanfaatkan confusion matrix dengan rumus dan tabel 2 sebagai berikut :

**Tabel 2.** *Confussion matrix*

Aktual / Prediksi	Tepat	Terlambat
Tepat	4	1
Terlambat	0	3

Berdasarkan hasil pengujian yang direpresentasikan dalam *confusion matrix*, diperoleh nilai akurasi sebesar 87,5%, yang menunjukkan bahwa model mampu mengklasifikasikan 7 dari 8 data uji secara benar. Nilai precision sebesar 57,2% mengindikasikan bahwa sebagian prediksi pada kelas tertentu masih mengandung kesalahan, sedangkan nilai recall yang relatif tinggi (80%) mengindikasikan bahwa model cukup baik dalam mengidentifikasi mahasiswa yang benar-benar termasuk dalam kategori tertentu [21]. Namun, nilai precision yang lebih rendah menunjukkan bahwa masih terdapat kesalahan dalam memprediksi kelas tertentu, khususnya pada kategori kelulusan tepat waktu.

Hasil ini menunjukkan bahwa algoritma Naïve Bayes cukup efektif dalam memodelkan hubungan antara variabel akademik dan ketepatan waktu kelulusan mahasiswa. Variabel seperti Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK) terbukti memiliki kontribusi signifikan terhadap hasil klasifikasi. Selain itu, atribut sumber pendanaan juga menunjukkan pengaruh terhadap ketepatan waktu kelulusan, yang mengindikasikan bahwa faktor finansial dapat berperan dalam mendukung keberhasilan studi mahasiswa.

Selain itu, atribut IPS dari semester awal hingga akhir menunjukkan pengaruh yang signifikan terhadap hasil klasifikasi. Mahasiswa dengan nilai IPS yang konsisten tinggi cenderung memiliki probabilitas lebih besar untuk lulus tepat waktu. Hal ini sejalan dengan konsep bahwa performa akademik yang stabil merupakan indikator utama keberhasilan studi.

Namun demikian, terdapat beberapa faktor yang memengaruhi performa model. Distribusi data yang tidak seimbang antara kelas *tepat waktu* dan *terlambat* berpotensi menyebabkan bias pada model, yang tercermin pada nilai precision yang relatif lebih rendah. Selain itu, jumlah data yang terbatas juga dapat membatasi kemampuan model dalam melakukan generalisasi.

Jika dibandingkan dengan penelitian terdahulu, nilai akurasi sebesar 87,5% tergolong cukup baik dan masih berada dalam rentang performa metode Naïve Bayes pada kasus klasifikasi akademik. Meskipun demikian, peningkatan performa masih dimungkinkan melalui pemilihan atribut yang lebih optimal serta penambahan jumlah data yang lebih representatif.

Dengan demikian, penerapan metode Naïve Bayes dapat memberikan kontribusi dalam analisis data akademik, khususnya dalam membantu institusi pendidikan untuk memprediksi dan mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan kelulusan. Model ini dapat dijadikan sebagai dasar dalam pengambilan keputusan akademik, seperti pemberian bimbingan atau intervensi lebih awal kepada mahasiswa yang berisiko.

#### IV. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Naive Bayes Classifier untuk mengklasifikasikan ketepatan waktu kelulusan mahasiswa berdasarkan data akademik. Model yang dibangun menunjukkan kinerja yang cukup baik dengan nilai akurasi sebesar 87,5%, yang menandakan bahwa sebagian besar data uji dapat

diprediksi dengan tepat. Nilai recall yang tinggi 80% menunjukkan kemampuan model dalam mengidentifikasi mahasiswa pada kelas tertentu dengan baik, meskipun nilai precision 57,2% masih menunjukkan adanya kesalahan klasifikasi pada beberapa kasus.

Hasil analisis juga mengungkap bahwa variabel akademik, terutama Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK), merupakan faktor dominan dalam menentukan ketepatan waktu kelulusan. Selain itu, atribut lain turut memberikan kontribusi terhadap hasil klasifikasi, yang mengindikasikan adanya pengaruh faktor non-akademik terhadap keberhasilan studi mahasiswa.

Meskipun demikian, performa model masih dipengaruhi oleh beberapa keterbatasan, seperti ketidakseimbangan distribusi data antar kelas dan jumlah dataset yang relatif terbatas. Secara keseluruhan, metode Naïve Bayes terbukti dapat digunakan sebagai pendekatan yang efektif dan efisien dalam analisis data akademik, serta berpotensi mendukung pengambilan keputusan berbasis data di lingkungan pendidikan tinggi.

## DAFTAR PUSTAKA

- [1] BAN-PT, “Peraturan Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) No. 1 tahun 2020,” *Banpt.or.Id. P.* 1, 2020.
- [2] M. Pendidikan, D. A. N. Kebudayaan, and R. Indonesia, “Peraturan Menteri Pendidikan Dan Kebudayaan Nomor 03 Tahun 2020 Tentang Standar Nasional Perguruan Tinggi,” no. 47, 2020.
- [3] M. Hasan, Milawati, Darodjat, H. Khairani, and T. Tahrir, *Media Pembelajaran*. 2021.
- [4] R. Sepriansyah, S. D. Purnamasari, K. R. N. Wardani, and N. Halim, “Prediksi Kelulusan Mahasiswa Fakultas Teknik Universitas Bina Darma Menggunakan Algoritma Naïve Bayes,” *JIPI (Jurnal Ilm. Penelit. Dan Pembelajaran Inform.*, vol. 8, no. 1, pp. 313–322, 2023, doi: 10.29100/jipi.v8i1.3459.
- [5] A. Yandi Saputra and Y. Primadasa, “Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbour Implementation of Classification Method to Predict Student Graduation Using K-Nearest Neighbor Algorithm,” *Techno.Com*, vol. 17, no. 4, p. 9, 2018.
- [6] B. K. Paju, D. Y. A. Fallo, and S. E. Mowata, “Analisis Algoritma Klasifikasi dalam Pembelajaran Sistem Informasi Geografis di Pendidikan Informatika,” *J. Kridatama Sains dan Teknol.*, vol. 7, no. 01, pp. 480–488, 2025, doi: 10.53863/kst.v7i01.1676.
- [7] T. Destiana, Y. Umaidah, and U. Enri, “Penerapan Algoritma Gaussian Naive Bayes Dalam Penentuan Prioritas Rehabilitasi Daerah Aliran Sungai Berdasarkan Parameter Lahan Kritis,” *INFOTECH J.*, vol. 9, no. 2, pp. 507–513, 2023, doi: 10.31949/infotech.v9i2.6501.
- [8] V. Wulandari, W. J. Sari, and Z. Alfian, “Implementation of Naïve Bayes Classifier and K-Nearest Neighbor Algorithms for Chronic Kidney Disease Classification Implementasi Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor untuk Klasifikasi Penyakit Ginjal Kronik,” vol. 4, no. April, pp. 710–718, 2024.
- [9] D. L. Safitry, A. Al Harani, E. Divayaning, F. H. Hanifa, D. F. Chema, and I. Paryudi, “Perbandingan Metode Decision Tree , Naive Bayes , dan Induction Rule untuk Klasifikasi Penyakit Diabetes,” *J. Informatics Adv. Comput.*, vol. 4, no. 1, pp. 22–30, 2023, [Online]. Available: <https://journal.univpancasila.ac.id/index.php/jiac/article/view/5488>
- [10] Syarli and A. A. Muin, “Metode Naive Bayes Untuk Prediksi Kelulusan,” *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2020, [Online]. Available: <https://media.neliti.com/media/publications/283828-metode-naive-bayes-untuk-prediksi-kelulu-139fcfea.pdf>
- [11] Y. E. Yuspita, R. Okra, and M. Rezeki, “Penerapan Algoritma Klasifikasi Untuk Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Rappidminer,” *Djtechno J. Teknol. Inf.*, vol. 6, no. 1, pp. 376–388, 2025, doi: 10.46576/djtechno.v6i1.6169.
- [12] A. Nata and S. Suparmadi, “Analisis Sistem Pendukung Keputusan Dengan Model Klasifikasi Berbasis Machine Learning Dalam Penentuan Penerima Program Indonesia Pintar,” *J. Sci. Soc. Res.*, vol. 5, no. 3, p. 697, 2022, doi: 10.54314/jssr.v5i3.1041.
- [13] F. Marudut, T. Pane, and D. Hindarto, “Comparative Analysis of Machine Learning Models for Stunting Prediction in Jakarta,” vol. 9, no. December, pp. 1365–1375, 2025.
- [14] S. Suriana, E. D. Bintari, and D. Kurniawan, “Desain Aplikasi Klasifikasi Kelulusan Mahasiswa Menggunakan Metode Naive Bayes dan Algoritma C4.5,” *J. Big Data Anal. Artif. Intell.*, vol. 3, no. 1, pp. 31–41, 2017, [Online]. Available: <https://www.neliti.com/publications/308842/>
- [15] Z. Amri, Kusri, and Kusnawi, “Edumatic : Jurnal Pendidikan Informatika Prediksi Tingkat Kelulusan Mahasiswa menggunakan Algoritma Naïve Bayes, Decision Tree, ANN, KNN, dan SVM,” vol. 7, no. 2, pp. 187–196, 2023, doi: 10.29408/edumatic.v7i2.18620.
- [16] T. M. Rahayu, B. A. Ningsi, I. Isnurani, and I. Arofah, “Klasifikasi Ketepatan Waktu Kelulusan Mahasiswa dengan Metode Naïve Bayes,” *Media Bina Ilm.*, vol. 15, no. 8, pp. 4993–5000, 2021.
- [17] D. H. Z. Abdussamad, *Metode Penelitian Kuantitatif*. Dr. Patta Rapanna,.

- [18] A. David Imanuel, N. Nawaningtyas Pusparini, and A. Sani, "Klasifikasi Untuk Memprediksi Tingkat Kelulusan Mahasiswa Stmik Widuri Menggunakan Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 12, no. 01, pp. 1–7, 2024, doi: 10.33884/jif.v12i01.8201.
- [19] N. Purwati and A. Dwi Januanti, "Aplikasi Data Mining Dengan Algoritma Naive Bayes Untuk Memprediksi Tingkat Kelulusan Mahasiswa," *J. Pepadun*, vol. 2, no. 1, pp. 123–137, 2021, doi: 10.23960/pepadun.v2i1.38.
- [20] A. Indriani, "Analisa Perbandingan Metode Naïve Bayes Classifier Dan K-Nearest Neighbor Terhadap Klasifikasi Data," *Sebatik*, vol. 24, no. 1, pp. 1–7, 2020, doi: 10.46984/sebatik.v24i1.909.
- [21] M. Tafsir, "Penerapan Algoritma Naïve Bayes untuk Prediksi Waktu Kelulusan Mahasiswa," *Nucl. Phys.*, no. 1. Universitas Putra Indonesia "YPTK" PADANG, Padang, 2023.