

## KLASIFIKASI TINGKAT PRESTASI MAHASISWA PADA MATA KULIAH PENAMBANGAN DATA MENGGUNAKAN NAÏVE BAYES

Basmallah Ramadhani Aisyah Putri<sup>1</sup>, Almurozy Mursidan<sup>2</sup>, Indra Sari Kusuma Wardhana<sup>3</sup>

<sup>1, 2, 3</sup>Universitas Indraprasta PGRI, Jl. Raya Tengah No.80, Jakarta Timur, DKI Jakarta

Email: [basmallahramadhani08@gmail.com](mailto:basmallahramadhani08@gmail.com)

---

### Article History

Received: 16-12-2026

Revision: 27-01-2026

Accepted: 29-01-2026

Published: 31-01-2026

**Abstract.** This study aims to classify the academic achievement levels of students in Data Mining courses using the Naïve Bayes algorithm and to evaluate the performance of the resulting classification model. This study uses a quantitative approach with academic data from 190 students, including assignment scores, mid-term exam scores, and final exam scores. The classification process was carried out by applying the Naïve Bayes algorithm, while model evaluation was performed using accuracy metrics, classification reports, and confusion matrices. The test results showed that the Naïve Bayes model produced an accuracy rate of 73.68%. Based on the classification report, classes B and B+ showed the best performance with recall values of 1.00 and f1-scores of 0.87 and 0.95, respectively. Confusion matrix analysis showed that most of the data in classes B and B+ were classified correctly. The results of this study indicate that the Naïve Bayes algorithm is quite effective in classifying students' academic achievement levels and has the potential to be used as an academic evaluation tool in learning decision-making.

**Keywords:** Classification, Naïve Bayes, Student Achievement, Data Mining

**Abstrak.** Penelitian ini bertujuan untuk mengklasifikasikan tingkat prestasi akademik mahasiswa pada mata kuliah Penambangan Data menggunakan algoritma *Naïve Bayes* serta mengevaluasi kinerja model klasifikasi yang dihasilkan. Penelitian ini menggunakan pendekatan kuantitatif dengan data akademik sebanyak 190 mahasiswa yang mencakup nilai tugas, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS). Proses klasifikasi dilakukan dengan menerapkan algoritma *Naïve Bayes*, sedangkan evaluasi model dilakukan menggunakan metrik akurasi, *classification report*, dan *confusion matrix*. Hasil pengujian menunjukkan bahwa model *Naïve Bayes* menghasilkan tingkat akurasi sebesar 73,68%. Berdasarkan *classification report*, kelas B dan B+ menunjukkan performa terbaik dengan nilai *recall* masing-masing sebesar 1,00 serta nilai *f1-score* sebesar 0,87 dan 0,95. Analisis *confusion matrix* menunjukkan bahwa sebagian besar data pada kelas B dan B+ berhasil diklasifikasikan dengan tepat. Hasil penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* cukup efektif dalam mengklasifikasikan tingkat prestasi akademik mahasiswa dan berpotensi digunakan sebagai alat bantu evaluasi akademik dalam pengambilan keputusan pembelajaran.

**Kata Kunci:** Klasifikasi, *Naïve Bayes*, Prestasi Mahasiswa, Penambangan Data

---

**How to Cite:** Putri, B. R. A., Mursidan, A., & Wardhana, I. S. K. (2026). Klasifikasi Tingkat Prestasi Mahasiswa Pada Mata Kuliah Penambangan Data Menggunakan Naïve Bayes. *Indo-MathEdu Intellectuals Journal*, 7 (1), 1207-1212. <http://doi.org/10.54373/imeij.v7i1.5050>

---

## PENDAHULUAN

Prestasi akademik mahasiswa menjadi salah satu tolak ukur penting dalam menilai efektivitas proses pembelajaran di lingkungan perguruan tinggi (Marliana et al., 2025). Capaian akademik mahasiswa mencerminkan tingkat penguasaan materi yang telah dipelajari serta keberhasilan strategi pembelajaran yang diterapkan (Biringkanae et al., 2023). Pada mata

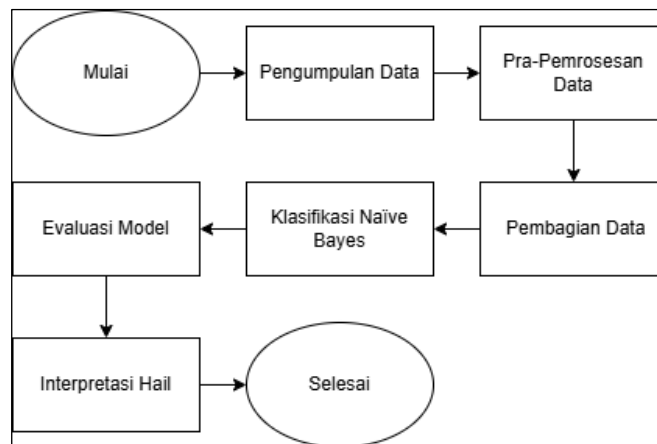
kuliah Penambangan Data, penilaian prestasi akademik memiliki peran yang lebih kompleks karena mata kuliah ini menuntut kemampuan berpikir analitis, pemahaman konsep, serta keterampilan logika dalam mengolah dan menganalisis data (Khasanah et al., 2025). Oleh sebab itu, diperlukan metode evaluasi yang mampu memberikan gambaran prestasi mahasiswa secara objektif dan akurat.

Pada praktiknya, penilaian prestasi mahasiswa masih banyak dilakukan menggunakan metode konvensional yang mengandalkan nilai akhir dari beberapa komponen penilaian, seperti tugas, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS). Pendekatan tersebut belum sepenuhnya memanfaatkan potensi data akademik yang tersedia dan berisiko menghasilkan evaluasi yang kurang komprehensif (Zulfiani et al., 2016). Seiring berkembangnya teknologi informasi, khususnya dalam bidang penambangan data, data akademik dapat diolah lebih lanjut untuk menghasilkan informasi yang bernilai dalam mendukung proses evaluasi dan pengambilan keputusan akademik (Sivi et al., 2023).

Penambangan data merupakan proses penggalian pengetahuan dari kumpulan data untuk menemukan pola atau informasi tersembunyi yang bermanfaat. Salah satu teknik yang umum digunakan dalam penambangan data adalah klasifikasi, yaitu metode pengelompokan data ke dalam kelas tertentu berdasarkan karakteristik yang dimiliki (Siregar et al., 2023). Algoritma Naïve Bayes merupakan salah satu algoritma klasifikasi berbasis probabilistik yang dikenal memiliki tingkat efisiensi dan akurasi yang cukup baik, terutama pada data dengan jumlah atribut yang terbatas (Saputra et al., 2024).

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan algoritma Naïve Bayes dalam mengklasifikasikan tingkat prestasi akademik mahasiswa pada mata kuliah Penambangan Data. Data penelitian terdiri dari 190 data mahasiswa dengan atribut nilai tugas, UTS, dan UAS. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, *classification report*, dan *confusion matrix*. Hasil penelitian ini diharapkan dapat menjadi alternatif solusi dalam pengembangan evaluasi akademik berbasis data serta membantu dosen dan pihak institusi dalam meningkatkan kualitas proses pembelajaran.

## METODE



**Gambar 1.** Tahapan penelitian

Gambar 1 menggambarkan tahapan penelitian dalam melakukan klasifikasi tingkat prestasi akademik mahasiswa menggunakan algoritma *Naïve Bayes*. Penelitian diawali dengan tahap pengumpulan data akademik mahasiswa yang meliputi nilai tugas, Ujian Tengah Semester (UTS), dan Ujian Akhir Semester (UAS). Data yang terkumpul kemudian melalui tahap pra-pemrosesan, yang mencakup perhitungan nilai akhir mahasiswa serta konversi nilai numerik ke dalam kategori *grade* sesuai dengan kriteria penilaian yang ditetapkan.

Selanjutnya, data yang telah diproses dibagi menjadi dua bagian, yaitu data latih dan data uji, untuk membangun dan menguji model klasifikasi. Proses klasifikasi dilakukan dengan menerapkan algoritma *Naïve Bayes* pada data latih guna mempelajari pola hubungan antara atribut penilaian dan tingkat prestasi mahasiswa. Model yang dihasilkan kemudian diuji menggunakan data uji untuk mengetahui kemampuan klasifikasi terhadap data yang belum pernah dilihat sebelumnya. Tahap evaluasi model dilakukan dengan menggunakan beberapa metrik, yaitu akurasi, *precision*, *recall*, dan *f1-score*, guna menilai kinerja model secara menyeluruh. Hasil evaluasi selanjutnya dianalisis dan diinterpretasikan sebagai dasar dalam penarikan kesimpulan mengenai efektivitas algoritma *Naïve Bayes* dalam mengklasifikasikan tingkat prestasi akademik mahasiswa.

## HASIL

Berdasarkan hasil pengolahan data nilai mahasiswa pada mata kuliah Penambangan Data, diperoleh nilai akhir yang dihitung dari komponen tugas, UTS, dan UAS. Hasil klasifikasi tingkat prestasi mahasiswa menggunakan algoritma *Naïve Bayes*, sebagaimana ditunjukkan pada Gambar 3, memperlihatkan kesesuaian antara *grade* aktual dan *grade* prediksi pada

sebagian besar data mahasiswa. Beberapa perbedaan prediksi terjadi pada mahasiswa dengan nilai akhir yang berada pada batas antar kategori grade.

|   | No.  | NIM          | Nama                         | Tugas | UTS | UAS |
|---|------|--------------|------------------------------|-------|-----|-----|
| 0 | 1.0  | 202243500235 | FAHMI MAULANA YUSUF          | 100   | 66  | 75  |
| 1 | 2.0  | 202243500236 | DEA ALFITRI YUSUF            | 100   | 87  | 75  |
| 2 | 3.0  | 202243500242 | MUHAMMAD IQBAL PUTRA PRATAMA | 100   | 57  | 70  |
| 3 | 4.0  | 202243500246 | ZIDAN ARIB DENDIANSYAH       | 100   | 66  | 70  |
| 4 | 5.0  | 202243500248 | FIRZHA RIFAN ADILAH          | 100   | 82  | 75  |
| 5 | 6.0  | 202243500250 | FERINA PUTRI BR PANDIA       | 100   | 40  | 60  |
| 6 | 7.0  | 202243500252 | MUHAMMAD IQBAL               | 100   | 44  | 78  |
| 7 | 8.0  | 202243500255 | RIZKYSETIAWAN                | 100   | 41  | 48  |
| 8 | 9.0  | 202243500258 | MUHAMMAD SULTAN ARAFAH       | 100   | 64  | 75  |
| 9 | 10.0 | 202243500260 | MUCHAMAD SYAH RAMADHAN       | 100   | 54  | 75  |

|   | Nilai_Akhir | Bobot | Grade | Grade_Prediksi |
|---|-------------|-------|-------|----------------|
| 0 | 77.3        | 3.25  | B+    | B+             |
| 1 | 83.6        | 3.75  | A-    | B              |
| 2 | 72.1        | 3.00  | B     | B              |
| 3 | 74.8        | 3.00  | B     | B+             |
| 4 | 82.1        | 3.75  | A-    | B              |
| 5 | 62.0        | 2.25  | C+    | C+             |
| 6 | 72.2        | 3.00  | B     | B              |
| 7 | 56.3        | 2.00  | C     | C+             |
| 8 | 76.7        | 3.25  | B+    | B+             |
| 9 | 73.7        | 3.00  | B     | B              |

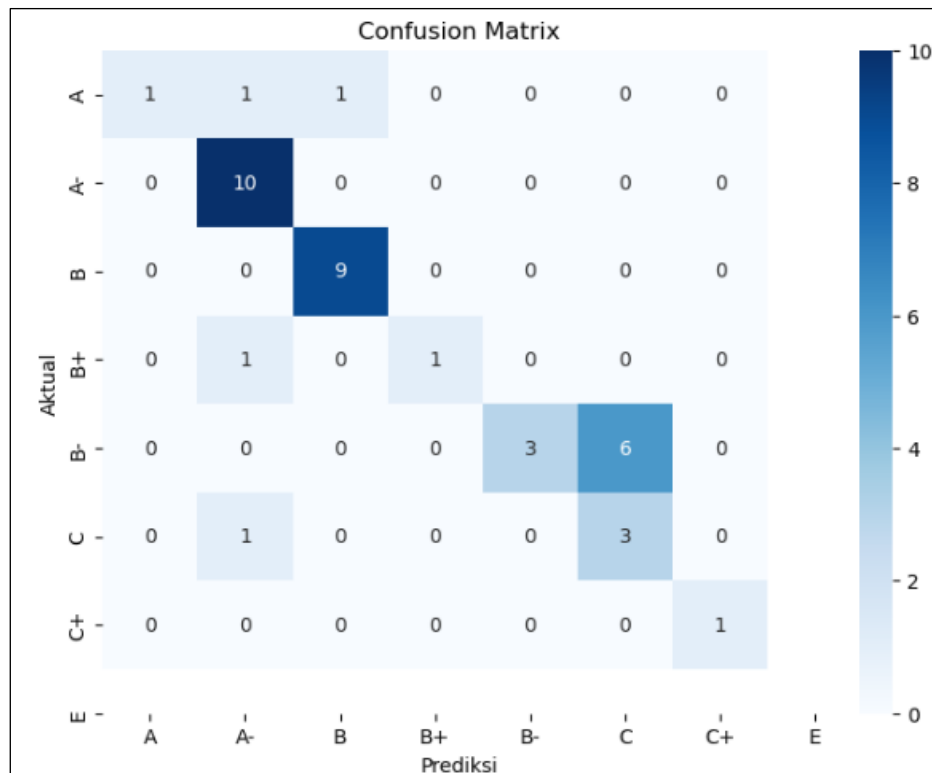
**Gambar 2.** Hasil klasifikasi prestasi mahasiswa

Hasil evaluasi model *Naïve Bayes* untuk klasifikasi tingkat prestasi mahasiswa menunjukkan akurasi sebesar 73,68% yang digambarkan pada Gambar 4. Berdasarkan *classification report*, model ini memiliki nilai *precision* tertinggi sebesar 1,00 pada kelas A-, B-, C, dan E, serta nilai *recall* sempurna pada kelas B dan E. Nilai *f1-score* tertinggi mencapai 1,00 pada kelas E, menunjukkan performa yang sangat baik dalam mengklasifikasikan kategori tersebut.

|                             |           |        |          |         |
|-----------------------------|-----------|--------|----------|---------|
| Akurasi: 0.7368421052631579 |           |        |          |         |
| Classification Report:      |           |        |          |         |
|                             | precision | recall | f1-score | support |
| A-                          | 1.00      | 0.33   | 0.50     | 3       |
| B                           | 0.77      | 1.00   | 0.87     | 10      |
| B+                          | 0.90      | 1.00   | 0.95     | 9       |
| B-                          | 1.00      | 0.50   | 0.67     | 2       |
| C                           | 1.00      | 0.33   | 0.50     | 9       |
| C+                          | 0.33      | 0.75   | 0.46     | 4       |
| E                           | 1.00      | 1.00   | 1.00     | 1       |
| accuracy                    |           |        | 0.74     | 38      |
| macro avg                   | 0.86      | 0.70   | 0.71     | 38      |
| weighted avg                | 0.85      | 0.74   | 0.72     | 38      |

**Gambar 3.** Hasil evaluasi kinerja algoritma

Pada Gambar 5 merupakan *confusion matrix* yang dihasilkan memperlihatkan distribusi kesalahan prediksi, dimana sebagian besar kesalahan terjadi pada kelas C+ yang sering diklasifikasikan sebagai kelas C, dan pada kelas B- yang beberapa kali salah prediksi sebagai kelas A-. Secara keseluruhan, model *Naïve Bayes* memberikan performa yang cukup baik dalam mengklasifikasikan tingkat prestasi mahasiswa dengan mempertimbangkan kompleksitas data dan variasi grade.



**Gambar 5.** *Confusion Matrix*

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* mampu memberikan performa yang cukup baik dalam mengklasifikasikan tingkat prestasi mahasiswa pada mata kuliah Penambangan Data. Meskipun masih terdapat kesalahan prediksi pada beberapa kelas tertentu, temuan ini memperkuat hasil penelitian sebelumnya yang menyatakan bahwa *Naïve Bayes* layak digunakan sebagai alat bantu evaluasi akademik. Dengan pengembangan lebih lanjut, seperti penambahan atribut penilaian atau pengaturan interval grade yang lebih adaptif, akurasi klasifikasi diharapkan dapat ditingkatkan.

## KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* dapat digunakan untuk mengklasifikasikan tingkat prestasi mahasiswa pada mata kuliah Penambangan Data dengan akurasi sebesar 73,68%. Hasil evaluasi menunjukkan performa model yang cukup baik,

meskipun masih terdapat kesalahan klasifikasi pada beberapa kelas yang memiliki batas nilai berdekatan. Secara keseluruhan, metode *Naïve Bayes* dapat dimanfaatkan sebagai alat bantu evaluasi akademik, dan pengembangan lebih lanjut dengan jumlah data yang lebih besar serta perbandingan dengan metode lain disarankan untuk meningkatkan kinerja model.

## REKOMENDASI

Penelitian selanjutnya disarankan untuk menggunakan jumlah data yang lebih besar dan lebih beragam agar hasil klasifikasi menjadi lebih akurat. Selain itu, perlu dilakukan perbandingan dengan algoritma klasifikasi lain serta penambahan atribut penilaian yang relevan guna meningkatkan performa dan keandalan model dalam mengklasifikasikan prestasi mahasiswa.

## REFERENSI

- Biringkanae, P., Bunahri, R. R., Kona, M., Supardam, D., & Prayitno, H. (2023). Pengaruh Kualitas Pembelajaran terhadap Capaian Akademik Mahasiswa Perguruan Tinggi Kedinasan. *Buletin Pengembangan Perangkat Pembelajaran*, 5(2), 1-11.
- Khasanah, U., Alanur, S. N., Trisnawati, S. N. I., Sulistyowati, R., Isma, A., Agustina, E., & Hamsar, I. (2025). Deep Learning Dalam Pendidikan: Pendekatan Pembelajaran Bermakna, Sadar, dan Menyenangkan. *Penerbit Tahta Media*.
- Marliana, Gracea, A., Anatasia, C., & Khoesasi, W. (2025). Analisis Faktor yang Mempengaruhi Prestasi Akademik Mahasiswa yang Kuliah sambil Bekerja di Universitas Universal. *Fortunate Business Review*, 5(2), 50-69.
- Saputra, A. W., Purnamasari, A. I., & Ali, I. (2024). Implementasi Algoritma Naïve Bayes Untuk Memprediksi Kualitas Air Yang Dapat Di Konsumsi. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 133-140.
- Siregar, Y. S., Handoko, D., Khairani, M., Syahputri, N. I., & Harahap, H. (2023). Implementasi Data Mining Klasifikasi Algoritma Chaid Dalam Menentukan Pola Penerima Mahasiswa Baru. *Digital Transformation Technology (Digitech)| E*, 3(2), 978-989.
- Sivi, N. A., Hartono, R., & Hanafi, P. (2023). Penerapan Algoritma C4. 5 untuk Prediksi Kelulusan Mahasiswa berdasarkan Data Akademik. *Polygon: Jurnal Ilmu Komputer dan Ilmu Pengetahuan Alam*, 1(5), 01-17.
- Zulfiani, Herlanti, Y., & Sofyan, A. (2016). *Kajian Penerapan Pendampingan Penelitian Tindakan Kelas Kolaboratif Antara Perguruan Tinggi Dan Sekolah*. Yogyakarta State University.