

## Literatur Review : Efektivitas Metode Naïve Bayes dalam Memprediksi Kelulusan Mahasiswa Tepat Waktu

Agung Sapto Hermawan<sup>1</sup>, Firdaus Hajjiyansah<sup>2</sup>, Sulistia Herti Riana<sup>3</sup>, Ines Heidiani Ikasari<sup>4</sup>

<sup>1</sup> Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspitek, Buaran, Kec. Pamulang, Kota Tangerang Selatan, Banten 15310, Indonesia.

Email: <sup>1</sup>[Agungsaptohermawan@gmail.com](mailto:Agungsaptohermawan@gmail.com), <sup>2</sup>[Firdaushjy@gmail.com](mailto:Firdaushjy@gmail.com), <sup>3</sup>[sulistiaheriana@gmail.com](mailto:sulistiaheriana@gmail.com), <sup>4</sup>[dosen01374@unpam.ac.id](mailto:dosen01374@unpam.ac.id).

(\* : coressponding author)

**Abstrak**– Penelitian ini bertujuan untuk melakukan tinjauan literatur terhadap penerapan metode Naïve Bayes dalam memprediksi kelulusan mahasiswa tepat waktu. Dari berbagai penelitian yang diulas, metode ini menunjukkan tingkat akurasi yang tinggi dalam klasifikasi data pendidikan. Faktor-faktor yang memengaruhi hasil prediksi meliputi nilai akademik, kehadiran, dan atribut sosial ekonomi mahasiswa. Hasil kajian menunjukkan bahwa metode Naïve Bayes dapat memberikan kontribusi signifikan dalam mendukung pengambilan keputusan institusi pendidikan.

**Kata Kunci:** Naïve Bayes, prediksi kelulusan, data mining, pendidikan, klasifikasi.

*Abstract*– This study aims to review the application of the Naïve Bayes method in predicting on-time student graduation. From various studies reviewed, this method demonstrates high accuracy levels in educational data classification. Factors influencing prediction outcomes include academic performance, attendance, and students' socioeconomic attributes. The findings indicate that the Naïve Bayes method significantly contributes to supporting educational institutions' decision-making processes

**Keywords:** Naïve Bayes, graduation prediction, data mining, education, classification.

### 1. PENDAHULUAN

Kelulusan mahasiswa tepat waktu merupakan salah satu indikator kunci dalam menilai efektivitas sistem pendidikan tinggi. Perguruan tinggi menghadapi tantangan besar untuk memastikan bahwa mahasiswa dapat menyelesaikan studi sesuai waktu yang direncanakan. Hal ini penting karena kelulusan tepat waktu tidak hanya mencerminkan keberhasilan institusi dalam menyediakan pendidikan berkualitas tetapi juga berkontribusi pada efisiensi alokasi sumber daya.

Beberapa faktor diketahui memengaruhi kelulusan tepat waktu mahasiswa, di antaranya adalah kualitas akademik mahasiswa, dukungan sosial, kemampuan finansial, serta proses pembelajaran di perguruan tinggi. Dalam konteks ini, metode prediksi berbasis *data mining*, khususnya algoritma *Naïve Bayes*, telah menjadi alat yang penting dalam membantu institusi pendidikan memahami pola dan faktor yang berkontribusi pada keberhasilan akademik mahasiswa.

*Naïve Bayes* dikenal karena kemampuannya dalam mengolah data dengan skala besar dan kompleks, serta memberikan hasil prediksi yang akurat dalam waktu yang relatif singkat. Dalam berbagai penelitian, algoritma ini telah digunakan untuk mengidentifikasi mahasiswa yang berisiko tidak lulus tepat waktu, sehingga institusi dapat mengambil langkah proaktif untuk memberikan intervensi yang sesuai.

Tujuan utama dari penelitian ini adalah untuk merangkum hasil-hasil penelitian sebelumnya terkait implementasi *Naïve Bayes* dalam memprediksi kelulusan mahasiswa tepat waktu. Penelitian ini tidak hanya bertujuan untuk mengidentifikasi tingkat akurasi algoritma tersebut tetapi juga untuk memahami kekuatan dan kelemahan pendekatan ini serta implikasinya dalam konteks pendidikan tinggi di Indonesia.

### 2. METODOLOGI PENELITIAN

#### 2.1 Pendekatan Penelitian

Pendekatan yang digunakan dalam penelitian ini adalah tinjauan literatur sistematis (*systematic literature review*). Metode ini bertujuan untuk mengumpulkan, mengevaluasi, dan mensintesis hasil penelitian sebelumnya terkait implementasi metode *Naïve Bayes* dalam prediksi

kelulusan mahasiswa. Langkah-langkah utama dalam pendekatan ini meliputi perencanaan, seleksi literatur, ekstraksi data, analisis, dan interpretasi.

Perencanaan melibatkan penentuan tujuan penelitian, kriteria inklusi dan eksklusi, serta strategi pencarian literatur. Seleksi literatur dilakukan berdasarkan relevansi artikel dengan topik penelitian. Ekstraksi data bertujuan untuk mengidentifikasi variabel kunci, metode analisis, dan hasil penelitian yang relevan. Analisis dilakukan untuk mengidentifikasi pola, tren, dan temuan utama dari literatur yang dikaji.

## 2.2 Pengumpulan Data

Pengumpulan data dilakukan melalui beberapa tahapan berikut:

### 1. Penelusuran Literatur:

- Menggunakan basis data akademik seperti Scopus, IEEE Xplore, PubMed, Google Scholar, dan SpringerLink.
- Kata kunci yang digunakan antara lain "Naïve Bayes", "prediction", "graduation", "education", dan "data mining".
- Menggunakan operator boolean seperti AND, OR, dan NOT untuk mempersempit atau memperluas pencarian.

### 2. Kriteria Inklusi dan Eksklusi:

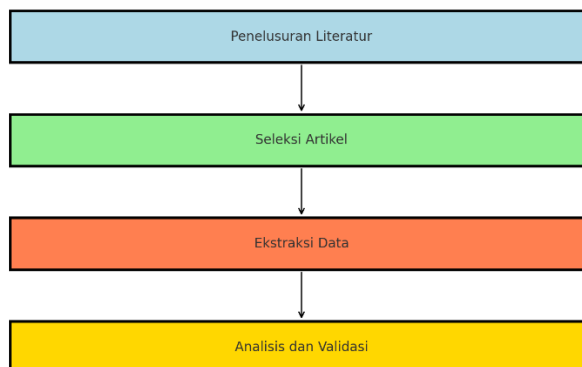
- Artikel yang diterbitkan dalam 10 tahun terakhir (2013-2023).
- Artikel yang menggunakan metode *Naïve Bayes* untuk prediksi kelulusan mahasiswa.
- Artikel yang tersedia dalam teks penuh dan berbahasa Inggris atau Indonesia.
- Artikel yang tidak relevan atau tidak mencantumkan hasil implementasi metode ini dikecualikan.

### 3. Ekstraksi Data:

- Setiap artikel yang dipilih dievaluasi untuk mengidentifikasi informasi berikut:
  - Jenis dataset yang digunakan (contoh: dataset akademik, dataset simulasi).
  - Metode preprocessing data (contoh: normalisasi, imputasi data hilang).
  - Parameter utama yang diukur (contoh: akurasi, presisi, sensitivitas).

### 4. Pengorganisasian Data:

- Data yang dikumpulkan dikelompokkan berdasarkan kategori, seperti jenis dataset, teknik analisis, dan hasil yang dicapai. Informasi ini kemudian disusun dalam tabel untuk memudahkan analisis perbandingan.



Gambar 1. Flowchart Pengumpulan data

## 2.3 Pengumpulan Data

Data yang telah dikumpulkan dianalisis menggunakan pendekatan deskriptif dan komparatif. Analisis deskriptif dilakukan untuk menggambarkan karakteristik penelitian yang diulas, seperti jenis dataset, jumlah sampel, dan variabel yang digunakan. Analisis komparatif dilakukan untuk membandingkan tingkat akurasi, presisi, dan keandalan metode *Naïve Bayes* dalam berbagai studi.

Selain itu, evaluasi terhadap metode preprocessing data, parameter tuning, dan teknik validasi yang digunakan dalam penelitian juga dilakukan untuk mengidentifikasi faktor-faktor yang memengaruhi hasil prediksi. Temuan dari analisis ini diinterpretasikan untuk memberikan wawasan yang lebih mendalam tentang keunggulan dan keterbatasan metode *Naïve Bayes* dalam konteks pendidikan.

## 2.4 Validasi Hasil

Hasil analisis divalidasi dengan membandingkan temuan penelitian ini dengan hasil dari literatur lainnya yang relevan. Validasi dilakukan untuk memastikan konsistensi temuan dan memberikan keyakinan terhadap kesimpulan yang dihasilkan. Langkah ini melibatkan evaluasi kritis terhadap metodologi penelitian, kualitas dataset, dan keandalan hasil prediksi.

Dengan pendekatan ini, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam memahami potensi metode *Naïve Bayes* untuk meningkatkan efisiensi dan efektivitas sistem pendidikan tinggi.

## 3. HASIL PENELITIAN

Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* memiliki akurasi yang tinggi, berkisar antara 75% hingga 95% dalam berbagai konteks. Faktor utama yang memengaruhi akurasi meliputi kualitas dataset, jumlah data training, dan preprocessing data. Bisa dilihat dalam tabel berikut :

Tabel 1. Daftar Literatur Review

No	Nama Penulis	Asal	Tujuan Penulisan	Dataset	Variabel Digunakan	Prosedur Penelitian	Hasil	Kelebihan Penelitian	Kekurangan Penelitian
1	Siti Nurulia et al.	Universitas Nahdlatul Ulama	Mengukur akurasi <i>Naïve Bayes</i> dalam prediksi	Data Kaggle dengan 500 observasi	Jenis kelamin, beasiswa, IPS, IPK	Preprocessing data, evaluasi confusion matrix	Akurasi mencapai 90%	Metode sederhana, cepat diterapkan	Tidak menguji variabel non-akademik
2	Armansyah et al.	UIN Sumatera Utara	Prediksi kelulusan tepat waktu	Dataset lokal lulusan dengan 44 data	IP semester 1-4, jalur masuk, total SKS	RapidMiner untuk klasifikasi dan validasi	Akurasi mencapai 100%	Relevan dengan konteks lokal	Ukuran dataset kecil
3	Lila Setiyani et al.	STMIK ROSMA	Sistematik review <i>Naïve Bayes</i> untuk prediksi	10 penelitian dengan fokus kelulusan tepat waktu	Tidak disebutkan secara eksplisit	Analisis komparatif literatur	Akurasi hingga 90%	Kajian menyeluruh dari berbagai studi	Tidak ada uji coba langsung
4	Pinkan Tri.A et al.	Universitas Muhammadiyah Jember	Analisis faktor kelulusan siswa	Data siswa lokal dengan atribut nilai ujian	Absensi siswa, nilai rata-rata, nilai ujian tengah semester	Data preprocessing, klasifikasi <i>Naïve Bayes</i>	Akurasi mencapai 95%	Fokus pada faktor spesifik	Tidak mencakup faktor sosial ekonomi
5	Bustami Yusuf et al.	UIN Ar-Raniry	Implementasi algoritma untuk prediksi akademik	Kombinasi dataset lokal dan simulasi	Usia, nilai rata-rata, tingkat dropout	Kombinasi <i>Naïve Bayes</i> dan algoritma lain	Akurasi hingga 78%	Inovasi dalam kombinasi algoritma	Akurasi relatif rendah dibandingkan metode lain

#### 4. PEMBAHASAN

Hasil kajian menunjukkan bahwa metode *Naïve Bayes* menawarkan solusi yang efektif dan efisien untuk memprediksi kelulusan mahasiswa tepat waktu. Salah satu keunggulan utama dari metode ini adalah kemampuannya dalam mengolah data dengan berbagai atribut dan memberikan hasil prediksi yang cukup akurat meskipun dalam waktu yang relatif singkat. Hal ini menjadikan metode ini cocok untuk diterapkan di institusi pendidikan dengan sumber daya komputasi terbatas.

Namun, ada beberapa tantangan yang harus diatasi, seperti sensitivitas algoritma terhadap dataset yang tidak seimbang dan keterbatasan dalam menangani hubungan kompleks antarvariabel. Beberapa penelitian menunjukkan bahwa penggunaan preprocessing data yang lebih mendalam, seperti normalisasi atau teknik sampling untuk mengatasi ketidakseimbangan data, dapat meningkatkan akurasi hasil prediksi.

Dari sisi penerapan praktis, integrasi *Naïve Bayes* dengan alat pengolah data modern seperti RapidMiner atau Python memungkinkan analisis yang lebih cepat dan efisien. Institusi pendidikan dapat memanfaatkan metode ini untuk mengidentifikasi mahasiswa yang memerlukan dukungan tambahan dalam proses akademiknya. Langkah-langkah seperti ini dapat berkontribusi pada peningkatan kelulusan tepat waktu dan efisiensi operasional perguruan tinggi.

#### 5. KESIMPULAN

Tinjauan literatur ini menunjukkan bahwa metode *Naïve Bayes* memiliki potensi besar untuk diterapkan dalam prediksi kelulusan mahasiswa. Dengan akurasi yang konsisten tinggi, metode ini dapat menjadi alat pendukung keputusan yang handal di institusi pendidikan. Penelitian lanjutan disarankan untuk mengeksplorasi kombinasi *Naïve Bayes* dengan algoritma lain guna meningkatkan kinerja prediksi.

#### REFERENCES

- Siti N., Harliana, Tito P. (2023). Implementasi Naïve Bayes Classifier Dalam Memprediksi Kelulusan Mahasiswa. JACIS : Journal Automation Computer Information System. Vol.3, No.01, Mei 2023, pp. 63~72.
- Armansyah, Rakhmat, K. R. (2022). Model Prediksi Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes Edumatic: Jurnal Pendidikan Informatika Vol. 6 No. 1, Juni, 2022, Hal.
- Lila S., Mokhammad W., Dudi A., Sri P. (2020). ANALISIS PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN METODE DATA MINING NAÏVE BAYES : SYSTEMATIC REVIEW. Faktor Exacta 13 (1): 35-43, 2020.
- Pinkan T. A., Bagus S. R., Deni A. (2014). MEMPREDIKSI HASIL KELULUSAN SISWA MENGGUNAKAN METODE NAIVE BAYES . Universitas Muhammadiyah Jember.
- Bustami Y, Muthmainna Q, Basrul, Ima D, Malahayati, & Mega E. (2020). IMPLEMENTASI ALGORITMA NAIVE BAYES DAN RANDOM FOREST DALAM MEMPREDIKSI PRESTASI AKADEMIK MAHASISWA UNIVERSITAS ISLAM NEGERI AR-RANIRY BANDA ACEH. Cyberspace: Jurnal Pendidikan Teknologi Informasi. Volume 4, Nomor 1, Maret 2020, 50-58