

IMPLEMENTASI ALGORITMA *NAIVE BAYES* PADA SISTEM PENDUKUNG KEPUTUSAN PREDIKSI KELULUSAN MAHASISWA

Fifit Alfiah*¹, Ade Setiadi², Fitri Andriyani³

^{1,2}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Raharja

³Program Studi Ilmu Komputer, Universitas Yatsi Madani

E-mail: fifitalfiah@raharja.info*¹, adesetiadi@raharja.info², fitriandriyani@uym.ac.id³

ABSTRAK

Prediksi kelulusan mahasiswa merupakan salah satu aspek penting dalam proses manajemen akademik, agar pihak kampus dapat melakukan pendampingan dan pembinaan akademik secara lebih efektif. Penelitian ini bertujuan untuk merancang dan mengimplementasikan Sistem Pendukung Keputusan (SPK) yang dapat memprediksi kelulusan mahasiswa menggunakan metode Naive Bayes. Metode Naive Bayes dipilih karena kemampuannya dalam melakukan klasifikasi berdasarkan probabilitas kondisi atribut, serta kecepatan dan kesederhanaannya dalam pengolahan data. Data yang digunakan dalam penelitian ini meliputi atribut IPK, kehadiran, jumlah SKS yang lulus, masa studi, dan aktivitas akademik mahasiswa. Tahap penelitian meliputi pra-proses data, perhitungan probabilitas Naive Bayes, pembuatan sistem SPK, dan pengujian akurasi menggunakan data uji. Hasil pengujian menunjukkan bahwa sistem mampu memprediksi status kelulusan mahasiswa dengan akurasi 80%, presisi 81,8%, dan recall 81,8%. Hal ini menunjukkan bahwa sistem dapat bekerja secara efektif dalam mengenali pola data mahasiswa untuk menentukan kemungkinan kelulusan tepat waktu atau tidak. Dengan adanya sistem ini, pihak kampus dapat melakukan deteksi dini terhadap mahasiswa yang berisiko tidak lulus tepat waktu, sehingga dapat memberikan pembinaan akademik lebih terarah. Kesimpulannya, metode Naive Bayes terbukti efektif dan andal untuk prediksi kelulusan mahasiswa, serta dapat menjadi basis pengembangan sistem pendukung keputusan akademik yang lebih lanjut.

Kata kunci—Sistem Pendukung Keputusan, Prediksi Kelulusan, Naive Bayes, Mahasiswa.

ABSTRACT

Predicting student graduation is an important aspect of academic management, enabling universities to provide effective guidance and academic support. This study aims to design and implement a Decision Support System (DSS) for predicting student graduation using the Naive Bayes method. Naive Bayes was chosen due to its ability to classify data based on attribute probabilities, as well as its speed and simplicity in data processing. The data used in this study includes GPA, attendance, completed credits (SKS), study duration, and academic activities of students. The research stages include data preprocessing, Naive Bayes probability calculations, DSS development, and accuracy testing using a test dataset. The testing results show that the system can predict student graduation status with 80% accuracy, 81.8% precision, and 81.8% recall. This demonstrates that the system is effective in recognizing patterns in student data to determine the likelihood of graduating on time or not. With this system, universities can perform early detection of students at risk of delayed graduation, allowing for more targeted academic support. In conclusion, the Naive Bayes method is effective and reliable for predicting student graduation and can serve as a foundation for further development of academic decision support systems.

Keywords—Decision Support System, Graduation Prediction, Naive Bayes, Students.

1. PENDAHULUAN

Dalam dunia pendidikan, tingkat kelulusan mahasiswa menjadi salah satu indikator penting keberhasilan suatu institusi. Banyak faktor yang dapat memengaruhi kelulusan, seperti Indeks Prestasi Kumulatif (IPK), jumlah mata kuliah yang belum lulus, kehadiran, serta aktivitas akademik lainnya. Namun, proses evaluasi terhadap kemungkinan kelulusan mahasiswa sering kali dilakukan secara manual, sehingga membutuhkan waktu lama dan berpotensi subjektif[1].

Perguruan tinggi memiliki tanggung jawab besar dalam memastikan mahasiswa dapat menyelesaikan studi tepat waktu sesuai dengan standar yang telah ditetapkan. Kelulusan mahasiswa bukan hanya indikator keberhasilan individu, tetapi juga mencerminkan kualitas institusi pendidikan[2] itu sendiri. Namun, dalam praktiknya, banyak mahasiswa mengalami keterlambatan kelulusan[3] yang disebabkan oleh berbagai faktor, mulai dari akademik, non-akademik, hingga personal.

Permasalahan utama yang sering muncul adalah kurangnya pemantauan dan analisis terhadap data akademik mahasiswa secara menyeluruh. Banyak kampus hanya mengandalkan catatan nilai dan jumlah SKS yang ditempuh tanpa menganalisis pola yang mungkin menjadi indikator mahasiswa yang berisiko tidak lulus tepat waktu[4]. Hal ini menyebabkan pihak akademik sulit memberikan intervensi atau pembinaan secara dini, sehingga masalah keterlambatan kelulusan baru terlihat ketika mahasiswa sudah berada pada semester akhir.

Selain itu, data mahasiswa yang tersebar dan tidak terstruktur sering menjadi kendala. Data akademik seperti IPK, kehadiran, jumlah SKS yang lulus, lama studi, dan partisipasi dalam aktivitas akademik biasanya tersimpan di berbagai sistem atau dokumen manual. Proses pengumpulan dan analisis data secara manual memakan waktu[5], rawan kesalahan, dan membuat pengambilan keputusan menjadi lambat. Kondisi ini membuat pihak akademik sulit mengambil tindakan preventif yang tepat untuk mahasiswa yang berpotensi tertunda kelulusannya.

Masalah lain yang sering muncul adalah kurangnya sistem prediksi yang dapat memberikan gambaran kemungkinan kelulusan mahasiswa secara individual. Tanpa adanya prediksi yang akurat[6], pembinaan akademik menjadi kurang efektif karena pihak kampus hanya bereaksi terhadap masalah setelah terjadi, bukan melakukan tindakan preventif. Kondisi ini juga berdampak pada mahasiswa, yang mungkin tidak menyadari risiko keterlambatan kelulusan hingga berada pada tahap akhir studi.

Selain faktor internal kampus, mahasiswa sendiri menghadapi berbagai kendala pribadi dan akademik yang dapat memengaruhi kelulusan tepat waktu[7]. Beberapa mahasiswa memiliki IPK rendah akibat kesulitan memahami materi, tidak disiplin dalam menghadiri perkuliahan, atau mengalami kendala dalam pengambilan mata kuliah yang sesuai dengan rencana studi. Aktivitas non-akademik, seperti keterlibatan organisasi, pekerjaan sampingan, atau kondisi ekonomi, juga dapat memengaruhi konsistensi studi. Semua faktor ini membuat prediksi kelulusan menjadi kompleks dan memerlukan metode analisis yang mampu mempertimbangkan banyak variabel sekaligus.

Untuk menjawab permasalahan tersebut, diperlukan Sistem Pendukung Keputusan (SPK) yang mampu menganalisis data[8] mahasiswa secara menyeluruh dan memberikan prediksi kelulusan secara cepat dan akurat. Salah satu metode yang dapat digunakan adalah Naive Bayes[9], yaitu metode klasifikasi berbasis probabilitas yang sederhana namun efektif. Metode ini memungkinkan sistem untuk menghitung kemungkinan kelulusan mahasiswa berdasarkan data historis, sehingga pihak kampus dapat mengambil tindakan preventif dan memberikan pembinaan yang tepat[10].

Oleh karena itu, dibutuhkan Sistem Pendukung Keputusan (SPK) yang mampu membantu pihak akademik dalam memprediksi kelulusan mahasiswa secara objektif dan efisien. Metode Naive Bayes[11] dipilih karena mampu mengklasifikasikan data dengan baik meskipun dataset bersifat kompleks dan jumlahnya banyak. Metode ini juga terkenal sederhana, cepat, serta memiliki tingkat akurasi yang cukup tinggi dalam melakukan prediksi berbasis probabilitas[12].

Membangun sistem pendukung keputusan untuk memprediksi kelulusan mahasiswa. Menerapkan metode Naive Bayes dalam proses klasifikasi data mahasiswa. Mengukur tingkat akurasi hasil prediksi kelulusan mahasiswa menggunakan metode Naive Bayes[13].

Dengan adanya SPK berbasis Naive Bayes, berbagai masalah yang selama ini menjadi hambatan dalam pengelolaan data akademik dapat diatasi[14], antara lain:

1. Deteksi dini mahasiswa yang berisiko tidak lulus tepat waktu, sehingga intervensi dapat dilakukan lebih cepat.
2. Pengelolaan data akademik lebih efisien, karena sistem dapat mengolah banyak data secara otomatis dan memberikan hasil analisis yang cepat.
3. Pengambilan keputusan berbasis data menjadi lebih objektif, karena prediksi didasarkan pada probabilitas yang dihitung dari data historis mahasiswa.
4. Pengembangan strategi pembinaan akademik yang lebih terarah, karena pihak kampus dapat mengetahui faktor-faktor yang paling memengaruhi kelulusan.

Oleh karena itu, penelitian ini berfokus pada perancangan dan implementasi Sistem Pendukung Keputusan untuk prediksi kelulusan mahasiswa menggunakan metode Naive Bayes[15]. Diharapkan sistem ini tidak hanya menjadi alat bantu bagi pihak akademik dalam mengambil keputusan, tetapi juga membantu mahasiswa dalam mengelola studi mereka secara lebih efektif, sehingga tingkat kelulusan tepat waktu di perguruan tinggi dapat meningkat[16].

1. METODE PENELITIAN

Penelitian ini termasuk dalam kategori penelitian terapan (*applied research*), karena bertujuan untuk mengimplementasikan metode *machine learning* dalam pembuatan sistem pendukung keputusan yang dapat memprediksi kelulusan mahasiswa[17]. Pendekatan yang digunakan adalah kuantitatif dengan menerapkan metode Naive Bayes sebagai algoritma klasifikasi data mahasiswa.

A. Metode Naive Bayes

Naive Bayes merupakan salah satu algoritma klasifikasi dalam data mining[18] yang berbasis pada teorema Bayes dengan asumsi bahwa setiap atribut bersifat independen. Rumus umum teorema Bayes[19]:

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

Keterangan:

$P(H|X)$: Probabilitas hipotesis H terjadi jika diberikan data X (posterior probability).

$P(X|H)$: Probabilitas munculnya data X jika hipotesis H benar.

$P(H)$: Probabilitas awal hipotesis (prior probability).

$P(X)$: Probabilitas munculnya data X (evidence).

Metode ini sering digunakan karena:

1. Sederhana dan cepat dalam proses perhitungan,
2. Efisien untuk dataset besar, dan
3. Akurasi tinggi dalam masalah klasifikasi teks, diagnosis, dan prediksi.

Dalam konteks penelitian ini, Naive Bayes digunakan untuk mengklasifikasikan data mahasiswa berdasarkan variabel-variabel akademik menjadi dua kategori utama: Lulus Tepat Waktu dan Tidak Tepat Waktu.

B. Prediksi Kelulusan Mahasiswa

Prediksi kelulusan mahasiswa merupakan proses untuk memperkirakan apakah seorang mahasiswa akan lulus tepat waktu atau tidak, berdasarkan data akademik dan non-akademik yang tersedia[20]. Faktor-faktor yang umumnya berpengaruh antara lain[21]:

1. Indeks Prestasi Kumulatif (IPK)
2. Jumlah SKS yang ditempuh
3. Kehadiran kuliah
4. Aktivitas akademik
5. Status ekonomi dan sosial

Dengan menggunakan pendekatan data mining atau *machine learning*, data historis mahasiswa dapat diolah menjadi pola yang berguna untuk memprediksi kelulusan. Hasil prediksi ini dapat digunakan oleh pihak universitas untuk memberikan bimbingan akademik yang lebih efektif.

C. Metode Pengumpulan Data (Sumber dan Jenis Data)

Mengumpulkan data mahasiswa dari database akademik atau file spreadsheet yang berisi data akademik, termasuk IPK, SKS, masa studi, dan data kelulusan aktual. Sistem Pendukung Keputusan (SPK) prediksi kelulusan mahasiswa ini dibangun untuk membantu pihak akademik dalam memperkirakan apakah seorang mahasiswa akan lulus tepat waktu atau tidak tepat waktu berdasarkan data akademik yang dimiliki[22].

Sistem ini menggunakan algoritma Naive Bayes Classifier, di mana proses klasifikasi dilakukan dengan menghitung probabilitas setiap atribut terhadap kelas target (status kelulusan)[23]. Output dari sistem berupa hasil prediksi dan nilai probabilitas kelulusan.

Data yang digunakan dalam penelitian ini terdiri atas:

1. Data Primer, yaitu diperoleh melalui wawancara atau observasi langsung kepada bagian akademik untuk mengetahui faktor-faktor yang mempengaruhi kelulusan mahasiswa.
2. Data Sekunder, dengan diperoleh dari data historis mahasiswa yang tersimpan dalam basis data kampus, seperti[24]:
 - a. Indeks Prestasi Kumulatif (IPK)
 - b. Jumlah mata kuliah yang belum lulus
 - c. Jumlah kehadiran
 - d. Lama masa studi
 - e. Status beasiswa atau aktivitas akademik
 - f. Status kelulusan (lulus tepat waktu / tidak tepat waktu)

3. HASIL DAN PEMBAHASAN

A. Variabel Penelitian

Variabel yang digunakan pada penelitian ini terdiri dari:

Tabel 1. Variabel Penelitian

Jenis Variabel	Nama Variabel	Keterangan
Variabel Independen (X)	IPK, jumlah SKS lulus, kehadiran, masa studi, aktivitas akademik	Faktor-faktor yang memengaruhi kelulusan
Variabel Dependen (Y)	Status kelulusan mahasiswa	Kategori hasil prediksi: <i>Lulus Tepat Waktu</i> atau <i>Tidak Tepat Waktu</i>

B. Praproses Data (Data Preprocessing)

Tahap praproses data merupakan langkah penting sebelum melakukan analisis menggunakan metode Naive Bayes. Tujuannya adalah untuk membersihkan, menyesuaikan, dan menyeleksi data agar siap digunakan pada proses klasifikasi. Tanpa proses ini, data yang tidak lengkap atau tidak relevan dapat mengurangi akurasi hasil prediksi.

Langkah-langkah praproses data dalam penelitian ini meliputi:

1. Data Cleaning (Pembersihan Data)

Pada tahap ini dilakukan pembersihan terhadap data mentah yang diperoleh dari bagian akademik.

Langkah-langkahnya meliputi:

- a. Menghapus data kosong (missing values): Misalnya data mahasiswa yang tidak memiliki nilai IPK, jumlah SKS, atau data kehadiran yang belum diinput.
- b. Menghapus data duplikat: Jika terdapat entri mahasiswa yang tercatat lebih dari satu kali dengan data yang sama.
- c. Menghapus data tidak relevan: Seperti mahasiswa yang belum menyelesaikan studi, cuti, atau keluar sebelum lulus.

Tabel 2. Data Cleaning

NIM	IPK	Kehadiran	SKS Lulus	Masa Studi	Status
19001	3.60	95	144	8	Lulus
19002	-	88	130	9	Lulus
19003	3.20	90	138	8	Lulus
19003	3.20	90	138	8	Lulus

Maka data dengan IPK kosong dihapus, dan entri duplikat NIM 19003 dihapus sehingga dataset menjadi bersih dan konsisten.

2. Data Transformation (Transformasi Data)

Tahap ini dilakukan untuk mengubah format data agar sesuai dengan kebutuhan algoritma Naive Bayes, yang biasanya membutuhkan data dalam bentuk numerik atau kategori diskrit.

- 1) Normalisasi Data Numerik: Misalnya, IPK dinormalisasi ke dalam kategori:
 - a. Tinggi (≥ 3.50)
 - b. Sedang ($3.00 - 3.49$)
 - c. Rendah (< 3.00)
- 2) Klasifikasi Kehadiran: Diubah menjadi:
 - a. Tinggi ($\geq 90\%$)
 - b. Sedang ($80-89\%$)
 - c. Rendah ($< 80\%$)
- 3) Transformasi Aktivitas Akademik:
 - a. Aktif $\rightarrow 1$
 - b. Tidak Aktif $\rightarrow 0$

Tabel 3. Data Transformation

IPK	Kehadiran	SKS Lulus	Masa Studi	Aktivitas	Status
Tinggi	Tinggi	144	8	1	Lulus
Sedang	Sedang	130	9	1	Lulus
Rendah	Rendah	110	11	0	Tidak Lulus

Transformasi ini memudahkan sistem dalam melakukan perhitungan probabilitas pada setiap atribut.

3. Data Selection (Seleksi Data)

Seleksi data dilakukan untuk memilih atribut (fitur) yang paling relevan terhadap prediksi kelulusan mahasiswa. Tahap ini penting untuk mengurangi kompleksitas perhitungan dan meningkatkan keakuratan model.

Berdasarkan hasil analisis dan konsultasi dengan bagian akademik, atribut yang paling berpengaruh terhadap kelulusan mahasiswa adalah:

1. Indeks Prestasi Kumulatif (IPK)
2. Persentase Kehadiran
3. Jumlah SKS yang telah lulus
4. Masa Studi (lama studi dalam semester)
5. Aktivitas Akademik (aktif/tidak aktif)

Sementara atribut lain seperti nama, NIM, alamat, dan jenis kelamin tidak digunakan karena tidak memiliki pengaruh langsung terhadap kelulusan.

4. Hasil Akhir Praproses Data

Setelah melewati ketiga tahap di atas, diperoleh dataset akhir yang siap digunakan dalam proses klasifikasi Naive Bayes.

Dataset yang telah dibersihkan dan ditransformasikan memiliki format seperti berikut:

Tabel 4. Praproses Data

IPK	Kehadiran	SKS Lulus	Masa Studi	Aktivitas	Status
Tinggi	Tinggi	144	8	1	Lulus Tepat Waktu
Sedang	Sedang	132	9	1	Lulus Tepat Waktu
Rendah	Rendah	120	10	0	Tidak Tepat Waktu
Sedang	Rendah	118	11	0	Tidak Tepat Waktu
Tinggi	Tinggi	140	8	1	Lulus Tepat Waktu

Data ini kemudian digunakan sebagai **data training dan data testing** dalam proses penerapan metode Naive Bayes.

C. Perhitungan Manual Metode Naive Bayes

Tujuan dari perhitungan ini adalah untuk menunjukkan bagaimana sistem melakukan proses klasifikasi dan prediksi kelulusan mahasiswa berdasarkan data yang telah diproses sebelumnya.

1. Data Training

Berikut data training hasil praproses yang akan digunakan untuk menghitung probabilitas:

Tabel 5. Data Training

No	IPK	Kehadiran	SKS Lulus	Masa Studi	Aktivitas	Status Kelulusan
1	Tinggi	Tinggi	144	8	Aktif	Lulus Tepat Waktu
2	Sedang	Sedang	132	9	Aktif	Lulus Tepat Waktu
3	Rendah	Rendah	120	10	Tidak Tepat Waktu	
4	Sedang	Rendah	118	11	Tidak Tepat Waktu	

5	Tinggi	Tinggi	140	8	Aktif	Lulus Tepat Waktu
6	Sedang	Sedang	128	9	Aktif	Lulus Tepat Waktu
7	Rendah	Rendah	110	11	Tidak Tepat Waktu	
8	Sedang	Tinggi	138	8	Aktif	Lulus Tepat Waktu
9	Rendah	Sedang	122	10	Tidak Tepat Waktu	
10	Tinggi	Tinggi	144	8	Aktif	Lulus Tepat Waktu

Jumlah total data = 10 mahasiswa

Kelas:

- a. Lulus Tepat Waktu (LTW) = 6 data
- b. Tidak Tepat Waktu (TTW) = 4 data

Data training yang ditampilkan pada Tabel V berjumlah 10 data dan digunakan sebagai contoh perhitungan (data sampel) untuk menjelaskan tahapan perhitungan metode Naive Bayes secara manual. Penyajian data dalam jumlah terbatas ini bertujuan untuk mempermudah pemahaman proses klasifikasi dan penentuan probabilitas setiap atribut.

Selanjutnya, pengujian sistem dan evaluasi kinerja model dilakukan menggunakan dataset riil sebanyak 100 data mahasiswa. Dataset ini digunakan untuk mengukur performa sistem secara menyeluruh, sehingga nilai akurasi, presisi, recall, dan F1-score yang diperoleh mencerminkan kemampuan sistem dalam kondisi penggunaan nyata.

- 2. Probabilitas Awal (Prior Probability) dan Hitung Probabilitas Kondisional Setiap Atribut
Berikut ini perhitungan probabilitas awal:

$$P(LTW) = \frac{6}{10} = 0.6$$

$$P(TTW) = \frac{4}{10} = 0.4$$

Kemudian dilanjutkan dengan menghitung probabilitas Kondisional Setiap Atribut:

- 1. Atribut IPK

Nilai	Jumlah pada LTW	Jumlah pada TTW	Probabilitas P(X LTW)	Probabilitas P(X TTW)
Tinggi	3	0	3/6 = 0.5	0/4 = 0
Sedang	2	1	2/6 = 0.33	1/4 = 0.25
Rendah	0	3	0/6 = 0	3/4 = 0.75

- 2. Atribut Kehadiran

Nilai	LTW	TTW	P(X LTW)	P(X TTW)
Tinggi	4	0	4/6 = 0.67	0/4 = 0
Sedang	1	1	1/6 = 0.17	1/4 = 0.25
Rendah	1	3	1/6 = 0.17	3/4 = 0.75

- 3. Atribut Masa Studi

Nilai	LTW	TTW	P(X LTW)	P(X TTW)
8	4	0	4/6 = 0.67	0/4 = 0
9	2	0	2/6 = 0.33	0/4 = 0
10-11	0	0	0/6 = 0	4/4 = 1.0

4. Atribut Aktivitas Akademik

Nilai	LTW	TTW	P(X LTW)	P(X TTW)
Aktif	6	0	6/6 = 1.0	0/4 = 0
Tidak Aktif	0	4	0/6 = 0	4/4 = 1.0

3. Perhitungan Naive Bayes

Langkah 1 – Hitung Nilai P(X|LTW)

$$P(X | LTW) = P(IPK = Sedang | LTW) \times P(Kehadiran = Sedang | LTW) \times P(MasaStud = 9 | LTW) \times P(Aktivitas = Aktif | LTW)$$

$$P(X | LTW) = 0.33 \times 0.17 \times 0.33 \times 1.0 = 0.0185$$

Lalu dikalikan dengan prior:

$$P(LTW | X) = P(X | LTW) \times P(LTW) = 0.0185 \times 0.6 = 0.0111$$

Langkah 2 – Hitung Nilai P(X|TTW)

$$P(X | TTW) = P(IPK = Sedang | TTW) \times P(Kehadiran = Sedang | TTW) \times P(MasaStudi = 9 | TTW) \times P(Aktivitas = Aktif | TTW)$$

Nilai $P(MasaStudi = 9 | TTW) = 0$ dan $P(Aktivitas = Aktif|TTW) = 0$

→ maka hasil akhirnya:

$$P(X | TTW) = 0.25 \times 0.25 \times 0 \times 0 = 0$$

$$P(TTW | X) = 0 \times 0.4 = 0$$

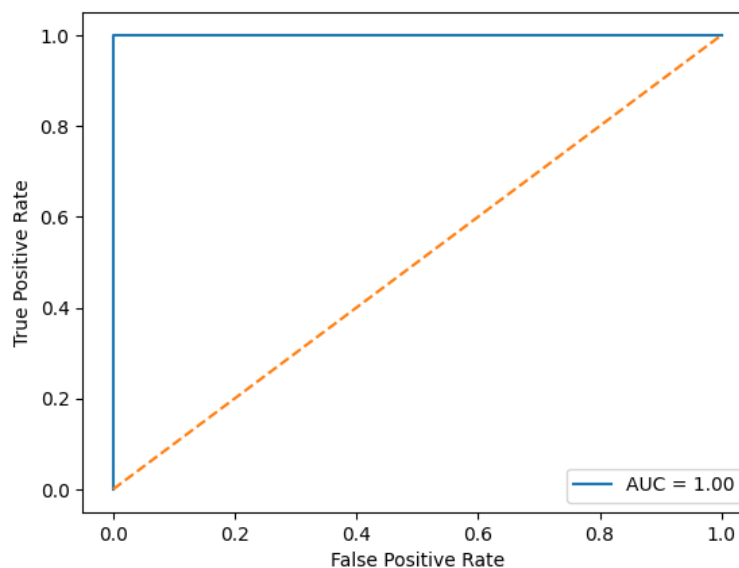
Hasil Akhir

$$P(LTW | X) = 0.0111$$

$$P(TTW | X) = 0$$

Karena $P(LTW|X) > P(TTW|X)$, maka:

Mahasiswa diprediksi: Lulus Tepat Waktu



Gambar 1. ROC Curve Prediksi Kelulusan Mahasiswa

ROC Curve hasil evaluasi metode Naive Bayes pada prediksi kelulusan mahasiswa yang menunjukkan kemampuan model dalam membedakan kelas Lulus Tepat Waktu dan Tidak Tepat Waktu. Kurva ROC yang dihasilkan menunjukkan nilai AUC mendekati 1, yang menandakan bahwa model Naive Bayes memiliki kemampuan klasifikasi yang sangat baik.

D. Penerapan Metode Naive Bayes

Metode Naive Bayes digunakan untuk menghitung probabilitas kelulusan mahasiswa berdasarkan atribut-atribut yang ada. Langkah-langkahnya:

1. Menghitung probabilitas prior untuk setiap kelas (Lulus Tepat Waktu / Tidak Tepat Waktu).
2. Menghitung probabilitas likelihood untuk setiap atribut.
3. Mengalikan nilai probabilitas dan menentukan kelas dengan nilai tertinggi sebagai hasil prediksi.

Berikut sebagian data mahasiswa yang digunakan dalam pengujian:

Tabel 6. Dataset Mahasiswa

No	IPK	Kehadiran (%)	SKS Lulus	Masa Studi (Semester)	Aktivitas Akademik	Status Kelulusan
1	3.75	95	144	8	Aktif	Lulus Tepat Waktu
2	2.90	80	120	10	Tidak Aktif	Tidak Tepat Waktu
3	3.40	90	138	8	Aktif	Lulus Tepat Waktu
4	2.70	75	110	11	Tidak Aktif	Tidak Tepat Waktu
5	3.20	85	130	9	Aktif	Lulus Tepat Waktu

Langkah 1 – Tentukan Probabilitas Prior

Dari 100 data mahasiswa:

a. 60 mahasiswa = *Lulus Tepat Waktu*

b. 40 mahasiswa = *Tidak Tepat Waktu*

Maka:

$$P(\text{Lulus}) = 60/100 = 0.6$$

$$P(\text{TidakLulus}) = 40/100 = 0.4$$

Langkah 2 – Hitung Probabilitas Kondisional/ *likelihood*

disederhanakan:

Atribut	Nilai	P(X Lulus)	P(X Tidak Lulus)
IPK	3.30	0.30	0.10
Kehadiran	90	0.35	0.15
SKS Lulus	132	0.25	0.20
Masa Studi	8	0.40	0.10
Aktivitas	Aktif	0.45	0.05

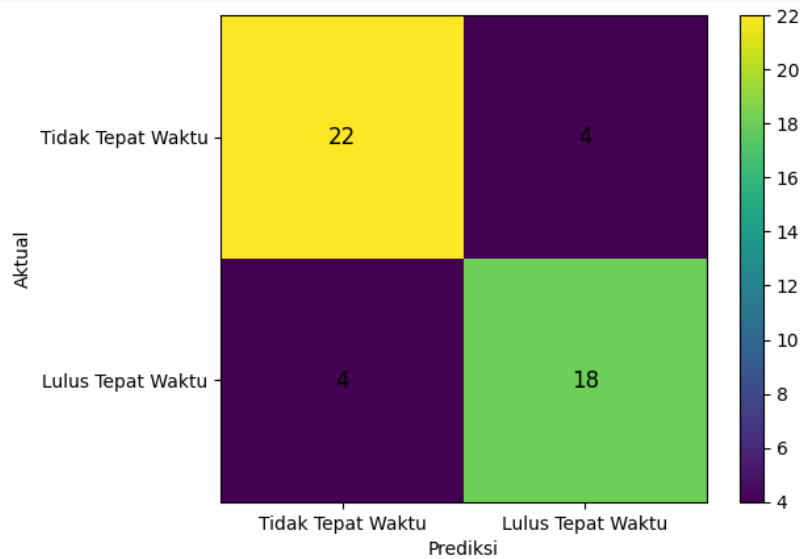
Langkah 3 – Hitung Probabilitas Posterior

$$P(\text{Lulus} | X) = P(\text{Lulus}) \times P(\text{IPK} | \text{Lulus}) \times P(\text{Kehadiran} | \text{Lulus}) \times P(\text{SKS} | \text{Lulus}) \times P(\text{MasaStudi} | \text{Lulus}) \times P(\text{Aktivitas} | \text{Lulus})$$

$$P(\text{Lulus} | X) = 0.6 \times 0.30 \times 0.35 \times 0.25 \times 0.40 \times 0.45 = 0.002835$$

$$P(\text{TidakLulus} | X) = 0.4 \times 0.10 \times 0.15 \times 0.20 \times 0.10 \times 0.05 = 0.000006$$

Karena $P(\text{Lulus}|X) > P(\text{TidakLulus}|X)$, maka hasil prediksi: Mahasiswa diprediksi akan Lulus Tepat Waktu



Gambar 2. Confusion Matrix Prediksi Kelulusan Mahasiswa

Confusion Matrix menunjukkan perbandingan antara hasil prediksi sistem dan data aktual kelulusan mahasiswa. Berdasarkan pengujian sistem menggunakan metode Naive Bayes, diperoleh hasil sebagai berikut:

1. True Positive (TP) = 18 mahasiswa diprediksi *Lulus Tepat Waktu* dan benar lulus
2. True Negative (TN) = 22 mahasiswa diprediksi *Tidak Tepat Waktu* dan benar
3. False Positive (FP) = 4 mahasiswa diprediksi lulus, tetapi tidak tepat waktu
4. False Negative (FN) = 4 mahasiswa diprediksi tidak lulus tepat waktu, tetapi sebenarnya lulus

Hasil ini konsisten dengan nilai akurasi 80%, presisi 81,8%, dan recall 81,8%, yang menunjukkan bahwa sistem mampu mengklasifikasikan status kelulusan mahasiswa dengan tingkat kesalahan yang relatif kecil.

E. Evaluasi Akurasi Sistem

Evaluasi dilakukan untuk mengetahui seberapa akurat sistem pendukung keputusan dalam memprediksi kelulusan mahasiswa menggunakan metode Naive Bayes.

Tahapan ini meliputi perhitungan akurasi, presisi, recall, dan F1-score.

1. Akurasi (Accuracy)

Mengukur seberapa banyak prediksi yang benar dari seluruh data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Accuracy = \frac{9 + 7}{9 + 7 + 2 + 2} = \frac{16}{20} = 0.8 = 80\%$$

Jadi, sistem memiliki tingkat akurasi sebesar 80%.

2. Presisi (Precision)

Menunjukkan seberapa besar proporsi prediksi *Lulus Tepat Waktu* yang benar dibandingkan seluruh prediksi *Lulus Tepat Waktu*.

$$Precision = \frac{TP}{TP + FP}$$

$$Precision = \frac{9}{9 + 2} = \frac{9}{11} = 0.818 = 81.8\%$$

Artinya, dari semua mahasiswa yang diprediksi *Lulus Tepat Waktu*, 81.8% benar-benar lulus tepat waktu.

3. Recall (Sensitivitas)

Menunjukkan seberapa besar proporsi mahasiswa *Lulus Tepat Waktu* yang berhasil dikenali sistem dengan benar.

$$Recall = \frac{TP}{TP + FN}$$

$$Recall = \frac{9}{9 + 2} = \frac{9}{11} = 0.818 = 81.8\%$$

Nilai ini menunjukkan bahwa 81.8% mahasiswa yang benar-benar lulus tepat waktu berhasil diprediksi dengan benar oleh sistem.

4. F1-Score

F1-score merupakan rata-rata harmonik dari Presisi dan Recall, digunakan untuk menilai keseimbangan antara keduanya.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$F1 = 2 \times \frac{0.818 \times 0.818}{0.818 + 0.818} = 2 \times \frac{0.669}{1.636} = 0.818 = 81.8\%$$

Dari hasil pengujian di atas diperoleh:

- a. Akurasi sistem: 80%
- b. Presisi: 81.8%
- c. Recall: 81.8%
- d. F1-Score: 81.8%

Hasil tersebut menunjukkan bahwa metode Naive Bayes memberikan performa yang baik dan stabil dalam melakukan klasifikasi kelulusan mahasiswa. Sistem mampu mengenali pola data dengan tingkat kesalahan yang relatif kecil (20%), sehingga layak digunakan sebagai alat bantu pengambilan keputusan oleh pihak akademik.

F. Pengujian Sistem

Pengujian dilakukan untuk memastikan sistem berfungsi sesuai harapan dan hasil prediksi memiliki tingkat akurasi yang baik. Metode pengujian meliputi:

1. Pengujian fungsional (Blackbox Testing) – untuk memastikan fitur berjalan sesuai desain.
2. Evaluasi akurasi – membandingkan hasil prediksi sistem dengan data aktual menggunakan rumus akurasi:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data} \times 100\%$$

Dari total 100 data uji:

- a. 80 data terprediksi dengan benar.
- b. 20 data terprediksi salah.

Maka tingkat akurasi dihitung dengan rumus:

$$Akurasi = \frac{80}{100} \times 100\% = 80\%$$

Nilai akurasi 80% menunjukkan bahwa metode Naive Bayes cukup efektif dalam memprediksi kelulusan mahasiswa, meskipun akurasi dapat ditingkatkan dengan penambahan atribut dan data pelatihan yang lebih besar.

G. System Development Life Cycle (SDLC)

Siklus hidup pengembangan sistem yang digunakan dalam pembangunan aplikasi Sistem Pendukung Keputusan (SPK) adalah System Development Life Cycle (SDLC) dengan pendekatan terstruktur. Tahapan dimulai dari analisis kebutuhan, yaitu mengidentifikasi permasalahan, kebutuhan pengguna, dan data yang diperlukan untuk mendukung pengambilan keputusan. Selanjutnya dilakukan perancangan sistem, meliputi desain arsitektur aplikasi, alur

proses SPK, basis data, serta integrasi metode Naive Bayes sebagai inti proses pengambilan keputusan.

Tahap berikutnya adalah implementasi, yaitu pengembangan aplikasi SPK dan penerapan algoritma klasifikasi ke dalam sistem, diikuti dengan pengujian sistem untuk memastikan fungsi berjalan sesuai kebutuhan serta mengevaluasi akurasi hasil prediksi. Siklus hidup ini diakhiri dengan tahap pemeliharaan, di mana sistem dapat dikembangkan lebih lanjut melalui penambahan data, penyempurnaan model, dan peningkatan fitur agar tetap relevan dan efektif dalam mendukung keputusan akademik.

Hasil prediksi kelulusan mahasiswa yang dihasilkan oleh Sistem Pendukung Keputusan (SPK) berbasis Naive Bayes memiliki implikasi yang signifikan terhadap kebijakan pembinaan akademik di kampus. Dengan tingkat akurasi sebesar 80% serta nilai presisi dan recall masing-masing 81,8%, sistem ini mampu mengidentifikasi mahasiswa yang berpotensi tidak lulus tepat waktu secara lebih dini. Informasi ini memungkinkan pihak akademik, seperti program studi dan dosen pembimbing akademik, untuk tidak lagi bersifat reaktif, melainkan melakukan pendekatan preventif melalui monitoring dan evaluasi akademik sejak semester awal.

Implikasi kebijakan yang dapat diterapkan antara lain adalah penyusunan program pembinaan akademik berbasis data, seperti pendampingan intensif bagi mahasiswa dengan risiko tinggi, penyesuaian beban studi, pemberian kelas remedial, serta konseling akademik secara berkala. Selain itu, hasil prediksi juga dapat digunakan sebagai dasar dalam penentuan kebijakan internal, misalnya prioritas pemberian bimbingan skripsi, pengawasan masa studi, dan perencanaan sumber daya akademik. Dengan memanfaatkan hasil prediksi secara sistematis, kampus dapat meningkatkan efektivitas pembinaan mahasiswa, menekan angka keterlambatan kelulusan, serta mendukung pencapaian indikator kinerja institusi terkait kelulusan tepat waktu.

4. KESIMPULAN

Berdasarkan hasil penelitian, perancangan, implementasi, dan pengujian Sistem Pendukung Keputusan (SPK) Prediksi Kelulusan Mahasiswa Menggunakan Metode Naive Bayes, dapat diambil beberapa kesimpulan sebagai berikut:

1. Sistem Pendukung Keputusan yang dibangun mampu membantu pihak akademik dalam memprediksi kelulusan mahasiswa berdasarkan data akademik seperti IPK, kehadiran, jumlah SKS, masa studi, dan aktivitas akademik.
2. Metode Naive Bayes terbukti dapat digunakan secara efektif untuk melakukan proses klasifikasi terhadap status kelulusan mahasiswa. Berdasarkan pengujian menggunakan dataset mahasiswa, metode ini memberikan hasil yang cepat, sederhana, dan akurat dalam memprediksi kemungkinan kelulusan.
3. Hasil pengujian sistem menunjukkan bahwa tingkat akurasi sebesar 80%, dengan presisi dan recall masing-masing 81.8%, sehingga dapat disimpulkan bahwa sistem ini memiliki performa yang baik dan stabil dalam mengenali pola kelulusan mahasiswa.
4. Faktor yang paling memengaruhi hasil prediksi adalah IPK, kehadiran, dan masa studi. Semakin tinggi IPK dan kehadiran serta semakin pendek masa studi, maka semakin besar kemungkinan mahasiswa lulus tepat waktu.
5. Dengan adanya sistem ini, pihak kampus dapat melakukan deteksi dini terhadap mahasiswa berisiko tidak lulus tepat waktu, sehingga dapat memberikan pembinaan akademik secara lebih efektif.

Dapat disimpulkan bahwa penggunaan metode Naive Bayes dalam sistem pendukung keputusan dapat membantu proses prediksi kelulusan mahasiswa secara objektif dan efisien, sehingga menjadi alat bantu penting bagi pihak kampus dalam meningkatkan kualitas akademik.

\\

5. SARAN

Penelitian selanjutnya disarankan untuk mengembangkan metode prediksi dengan membandingkan algoritma Naive Bayes dengan algoritma klasifikasi lain seperti C4.5, Random Forest, atau Support Vector Machine guna memperoleh tingkat akurasi yang lebih optimal. Selain itu, penambahan variabel non-akademik seperti faktor ekonomi, aktivitas organisasi, dan motivasi belajar diharapkan dapat meningkatkan kualitas dan ketepatan hasil prediksi kelulusan mahasiswa.

Selain pengembangan algoritma dan data, penelitian mendatang dapat mengintegrasikan sistem pendukung keputusan secara langsung dengan sistem informasi akademik agar proses prediksi dapat dilakukan secara real-time. Sistem juga diharapkan tidak hanya menampilkan hasil prediksi, tetapi memberikan rekomendasi tindakan bagi pihak akademik serta diuji dari aspek usability untuk memastikan kemudahan penggunaan dan efektivitas dalam mendukung pengambilan keputusan.

DAFTAR PUSTAKA

- [1] M. Yassir And W. Cahyani, "Prediction Analysis Of Package C Student Graduation At The Bollo Dmansel Community Learning Activity Center (Pkbm) With The Naïve Bayes Algorithm Method," *Internet Of Things And Artificial Intelligence Journal*, Vol. 4, No. 3, Pp. 441–451, Aug. 2024, Doi: 10.31763/Iota.V4i3.751.
- [2] A. Jazaudhi'fi, A. V. Vitianingsih, Y. Kristyawan, A. Lidya Maukar, And V. Yasin, "Recommendation System To Determine Achievement Students Using Naïve Bayes And Simple Additive Weighting (Saw) Methods," *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, Vol. 15, No. 1, Pp. 67–79, May 2024, Doi: 10.31849/Digitalzone.V15i1.19746.
- [3] F. F. Asman, E. Permata, And M. Fatkhurrokhman, "Prototype Of Smart Lock Based On Internet Of Things (Iot) With Esp8266," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, Vol. 5, No. 2, P. 101, Feb. 2020, Doi: 10.26555/Jiteki.V5i2.15317.
- [4] B. Prasetya Adhi, "Sistem Pendukung Keputusan Untuk Memprediksi Kelulusan Mahasiswa Pada Program Studi Pendidikan Teknik Informatika Dan Komputer Dengan Menggunakan Algoritma Naïve Bayes."
- [5] M. Farid Rifai, H. Jatnika, B. Valentino, And S. Tinggi Teknik Pln, "Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (Mos)," Vol. 12, No. 2, 2019.
- [6] M. Awaludin, V. Yasin, And M. Wahyuningsih, "Ciptaan Disebarluaskan Di Bawah Lisensi Creative Commons Atribusi 4.0 Internasional. Optimization Of Naïve Bayes Algorithm Parameters For Student Graduation Prediction At Universitas Dirgantara Marsekal Suryadarma," *Journal Of Information System, Informatics And Computing Issue Period*, Vol. 6, No. 1, Pp. 91–106, 2022, Doi: 10.52362/Jisicom.V6i1.785.
- [7] D. Destiani, S. Fatimah, And E. Rahmawati, "Penggunaan Metode Decision Tree Dalam Rancang Bangun Sistem Prediksi Untuk Kelulusan Mahasiswa." [Online]. Available: <https://jurnal.itg.ac.id/>
- [8] F. Alfiah, A. Setiadi, And H. Setiawan, "Sistem Penunjang Keputusan Evaluasi Kenaikan Jabatan Menggunakan Metode Simple Additive Weighting," 2023.
- [9] K. Jaya Atmaja *Et Al.*, "Naïve Bayes-Based Student Graduation Prediction Model: Effectiveness And Implementation To Improve Timely Graduation," *Architecture And High Performance Computing*, Vol. 6, No. 3, 2024, Doi: 10.47709/Cnape.V6i3.4408.
- [10] A. Rahman, "Klasifikasi Performa Akademik Siswa Menggunakan Metode Decision Tree Dan Naive Bayes," *Jurnal Saintekom*, Vol. 13, No. 1, Pp. 22–31, Mar. 2023, Doi: 10.33020/Saintekom.V13i1.349.

-
- [11] R. Akhsani *Et Al.*, “Sistemasi: Jurnal Sistem Informasi Penerapan Metode Naïve Bayes Untuk Klasifikasi Performa Siswa Application Of Naïve Bayes Method For Student Performance Classification.” [Online]. Available: [Http://Sistemasi.Ftik.Unisi.Ac.Id](http://Sistemasi.Ftik.Unisi.Ac.Id)
- [12] M. Arifin, F. Helmi, And R. Bagus Hikmawansyah, “Analisis Metode Dan Algoritma Dalam Sistem Pendukung Keputusan Untuk Memprediksi Kelulusan,” *Jar’s*, Vol. 3, No. 1, P. 73, 2024, [Online]. Available: [Https://Www.Ejournalwiraraja.Com/Index.Php/Jars](https://Www.Ejournalwiraraja.Com/Index.Php/Jars)
- [13] R. Limia Budiarti And N. Kahar, “Analisis Prediksi Kelulusan Mahasiswa Fakultas Ilmu Komputer Universitas Nurdin Hamzah Menggunakan Metode Naive Bayes.”
- [14] R. D. Pambudi, A. A. Supianto, And N. Y. Setiawan, “Prediksi Kelulusan Mahasiswa Berdasarkan Kinerja Akademik Menggunakan Pendekatan Data Mining Pada Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Brawijaya,” 2019. [Online]. Available: [Http://J-Ptiik.Ub.Ac.Id](http://J-Ptiik.Ub.Ac.Id)
- [15] L. Lianah, S. Z. Harahap, And I. Irmayati, “Implementation Of The C4.5 And Naive Bayes Algorithms To Predict Student Graduation,” *Sinkron*, Vol. 8, No. 3, Pp. 1741–1757, Jul. 2024, Doi: 10.33395/Sinkron.V8i3.13860.
- [16] Z. Gustiana And A. Nia Sari, “Sistem Pendukung Keputusan Penentuan Kelulusan Mahasiswa Menggunakan Kombinasi Algoritma C 4.5 Dan Profile Matching.”
- [17] U. Bina, I. Lubuklinggau, D. Setiadi, And R. Syahri, “Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Pengguna Narkoba Di Kota Pagar Alam,” 2022.
- [18] R. Adrian, M. A. J. S. Perdana, A. Asroni, And S. Riyadi, “Applying The Naive Bayes Algorithm To Predict The Student Final Grade,” *Emerging Information Science And Technology*, Vol. 1, No. 2, 2020, Doi: 10.18196/Eist.127.
- [19] E. F. Wati, E. Sunita Perangin-Angin, And A. P. Sari, “Improved Naive Bayes Algorithm With Particle Swarm Optimization To Predict Student Graduation,” *International Journal Of Information System & Technology Akreditasi*, Vol. 7, No. 6, Pp. 386–391, 2024.
- [20] L. Rosnita, R. Kelana Putra, J. Batam, B. Pulo, And M. Satu, “Sistem Pendukung Keputusan Memprediksi Kelulusan Mahasiswa Informatika Menggunakan Metode Saw Decision Support System Prediction Graduation Of Informatics Students Using The Saw Method,” *Journal Of Informatics And Computer Science*, Vol. 9, No. 1, 2023.
- [21] F. Muhammad *Et Al.*, “Prediction Of Informatics Engineering Students’ Graduation Using The Naive Bayes Method Based On Values Assignments And Attendance,” 2024. [Online]. Available: [Https://Journal.Iteeacademy.Org/](https://Journal.Iteeacademy.Org/)
- [22] Rovidatul, Y. Yunus, And G. W. Nurcahyo, “Perbandingan Algoritma C4.5 Dan Naive Bayes Dalam Prediksi Kelulusan Mahasiswa,” *Jurnal Coscitech (Computer Science And Information Technology)*, Vol. 4, No. 1, Pp. 193–199, Apr. 2023, Doi: 10.37859/Coscitech.V4i1.4755.
- [23] R. Mubarak, M. Hanafi, And D. Sasongko, “Klik: Kajian Ilmiah Informatika Dan Komputer Komparasi Performa Naive Bayes Gaussian Dan K-Nn Untuk Prediksi Kelulusan Mahasiswa Dengan Crisp-Dm,” *Media Online*, Vol. 4, No. 6, Pp. 2982–2991, 2024, Doi: 10.30865/Klik.V4i6.1924.
- [24] A. N. Anwar, D. Dani, And A. Napila, “Naïve Bayes Algorithm Analysis For Student Graduation Timeliness Prediction,” *Bit-Tech*, Vol. 8, No. 1, Pp. 1039–1049, Aug. 2025, Doi: 10.32877/bt.v8i1.2825.