



Prediksi Prestasi Akademik Mahasiswa Menggunakan Metode Naïve Bayes

Alif Fiandi Firdaus^{1*}, Anggita Syaidatul Soffa², Muhammad Haris³, Muhammad Nursalim Imron⁴, Aries Saiffudin⁵

Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

E-mail: ^{1*}aliffiandi1@gmail.com, ²luthfianggita3@gmail.com, ³mharis2290@gmail.com,
⁴mnsimron@gmail.com, ⁵aries.saiffudin@unpam.ac.id

(* : coresponding author)

Abstrak - Komponen penting dari pertumbuhan suatu negara adalah sistem pendidikannya. Memprediksi pencapaian akademis mahasiswa memungkinkan institusi pendidikan mengambil langkah proaktif untuk meningkatkan kinerja mahasiswa, yang merupakan salah satu cara pendidikan tinggi berupaya meningkatkan kualitas pengajaran. Di institusi pendidikan tinggi, mengelompokkan mahasiswa berdasarkan potensi keberhasilan akademisnya dapat menjadi strategi yang berguna untuk menurunkan tingkat kegagalan dalam bidang akademis, meningkatkan tingkat pencapaian, dan meningkatkan pengelolaan sumber daya. Penelitian ini bertujuan untuk menilai kemajuan akademik siswa menggunakan pendekatan Naïve Bayes, yang memperkirakan kinerja masa depan dengan menganalisis data masa lalu secara statistik dan melakukan perhitungan probabilitas.

Kata Kunci: Prediksi; Prestasi Akademik; Mahasiswa; Metode Naïve Bayes

Abstract - An important component of a country's growth is its education system. Predicting student academic achievement allows educational institutions to take proactive steps to improve student performance, which is one way higher education seeks to improve the quality of teaching. In higher education institutions, grouping students based on their potential for academic success can be a useful strategy for reducing academic failure rates, increasing achievement levels, and improving resource management. The aim of this research is to assess students' academic progress using the Naïve Bayes approach, which estimates future performance by statistically analyzing past data and performing probability calculations.

Keywords: Prediction; Academic Achievement; Student; Naïve Bayes Method

1. PENDAHULUAN

Kemampuan untuk memprediksi prestasi akademis mahasiswa sangat penting dalam kaitannya dengan pendidikan tinggi. Mahasiswa yang diprediksi berprestasi buruk dapat terhindar dari kegagalan akademik dengan memiliki akses terhadap prediksi awal pencapaian akademik. Selain itu, informasi ini sangat bermanfaat bagi lembaga pendidikan dalam memantau proses pembelajaran dan mengambil kebijakan yang dibutuhkan dalam meningkatkan kualitas pendidikan.

Dosen juga mendapat manfaat signifikan dari prediksi prestasi akademik, karena informasi ini memungkinkan mereka untuk mengevaluasi dan meningkatkan metode pengajaran mereka. Meskipun demikian, memprediksi prestasi akademik bukanlah tugas yang mudah karena melibatkan berbagai faktor seperti metode pengajaran, psikologi, status sosial ekonomi, dan lingkungan akademik.

Untuk menjamin mutu pendidikan, lembaga pendidikan di Indonesia biasanya menerapkan mekanisme evaluasi dan monitoring proses belajar mengajar. Evaluasi dosen umumnya dilakukan melalui kuesioner yang diisi oleh mahasiswa pada akhir semester. Namun, pendekatan ini memiliki beberapa kelemahan, kurangnya kemampuan mengidentifikasi secara akurat faktor-faktor yang dapat berdampak signifikan terhadap keberhasilan akademik mahasiswa dan tidak membantu mahasiswa dan dosen dalam melakukan evaluasi selama semester berlangsung.

Oleh karena itu, diperlukan strategi dan teknik baru untuk mengatasi permasalahan tersebut. Penelitian ini bertujuan untuk mengukur prestasi akademik mahasiswa menggunakan metode Naïve Bayes. Metode ini menggunakan perhitungan probabilitas dan analisis statistik dari data historis untuk memprediksi kinerja akademik di masa depan.

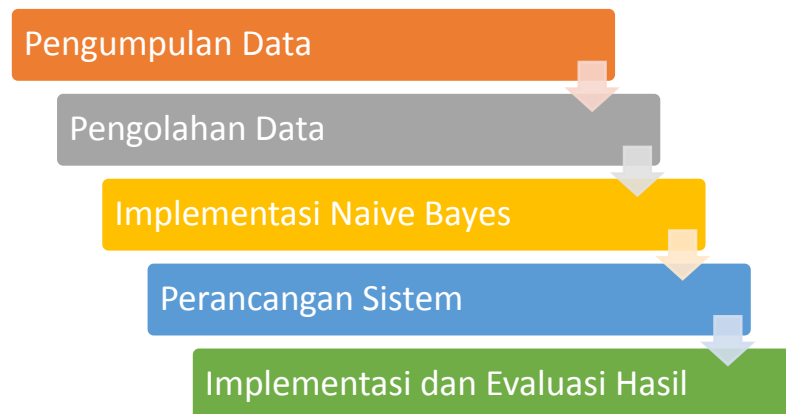


Sistem prediksi yang dirancang diharapkan dapat membantu universitas dalam meningkatkan kualitas prestasi akademik mahasiswa, dengan menyediakan informasi yang akurat dari semester pertama sampai semester terakhir yang ditempuh saat ini. Dengan demikian, bagian akademik atau admin universitas dapat dengan mudah mengidentifikasi mahasiswa yang berprestasi menggunakan metode Naïve Bayes.

2. METODOLOGI

Metode Naïve Bayes, sebuah metode pembelajaran mesin berdasarkan perhitungan probabilitas, digunakan dalam penelitian ini untuk menggunakan data historis guna memprediksi potensi kejadian di masa depan. Tujuan dari metode Naive Bayes adalah untuk mengklasifikasikan siswa dan menentukan apakah mereka termasuk dalam kategori berprestasi atau tidak. Proses klasifikasi ini melibatkan pembuatan model berdasarkan dataset pelatihan, yang kemudian digunakan untuk memprediksi kelas data uji. Salah satu keunggulan utama dari metode ini adalah kebutuhan data pelatihan yang relatif sedikit, sehingga memungkinkan untuk menghasilkan model yang efisien.

Langkah-langkah dalam penelitian ini meliputi pengumpulan, pengolahan data, implementasi metode Naive Bayes, perancangan sistem, serta implementasi dan evaluasi hasil.



Gambar 1. Metode Naive Bayes

a. Pengumpulan Data

Proses pengumpulan data meliputi studi literatur, observasi, wawancara dan dokumentasi terkait data akademik mahasiswa. Data yang dikumpulkan berupa biodata mahasiswa seperti nomor induk mahasiswa, usia, jenis kelamin, jurusan, semester, indeks prestasi semester (IPS), indeks prestasi kumulatif (IPK), keikutsertaan dalam organisasi, dan kepemilikan piagam penghargaan. Data ini akan digunakan sebagai atribut dalam prediksi prestasi akademik mahasiswa.

b. Pengolahan Data

Pengolahan data melibatkan berbagai tahapan data mining, termasuk pembersihan data, pemilihan data, validasi data, dan identifikasi dataset untuk tujuan pelatihan dan pengujian. Pembersihan data adalah tentang menghilangkan informasi yang tidak penting. Seleksi data dilakukan untuk memilih atribut yang diperlukan dan menghilangkan yang tidak diperlukan dalam dataset mahasiswa, seperti nomor induk mahasiswa, usia, jenis kelamin, jurusan, semester, IPK dari semester 1 hingga semester terakhir, keikutsertaan dalam organisasi, dan kepemilikan piagam penghargaan. Data kemudian diformat agar sesuai untuk proses data mining.

c. Implementasi Naïve Bayes



Pada tahap ini data mining dilakukan dengan membaca kumpulan data pelatihan dan data pengujian. Saat menggunakan algoritma Naive Bayes, probabilitas ditambahkan dan dihitung untuk menentukan klasifikasi.

d. Perancangan Sistem

Sebelum mengimplementasikan sistem, kebutuhan perangkat lunak harus ditentukan dengan menggunakan bahasa pemrograman. Proses ini mencakup pemodelan arsitektur perangkat lunak, struktur data, dan representasi antarmuka pengguna. Proses ini mencakup tahapan seperti struktur data, representasi antarmuka pengguna, dan pemodelan arsitektur perangkat lunak.

e. Implementasi dan Evaluasi Hasil

Menerjemahkan desain sistem ke dalam bahasa pemrograman diperlukan pada tahap ini. Hasilnya dievaluasi dengan melakukan pengujian fungsionalitas sistem serta pengujian data untuk memverifikasi bahwa prediksi sesuai dengan tujuan penelitian.

Dengan metodologi yang terstruktur ini, penelitian diharapkan dapat menghasilkan sistem prediksi yang akurat dan efisien untuk membantu universitas dalam meningkatkan prestasi akademik mahasiswa.

3. HASIL DAN PEMBAHASAN

3.1. Implementasi Sistem

Proses pengambilan data dilakukan dengan memilih data mahasiswa secara acak. Langkah pertama adalah melakukan pembersihan data, seleksi, dan penentuan data yang akan dijadikan sebagai data pelatihan (training). Proses pengujian data dibagi menjadi dua bagian utama: pertama, mengukur akurasi dan prediksi data menggunakan metode Naive Bayes. Data pelatihan digunakan untuk membangun tabel probabilitas, sedangkan data pengujian digunakan untuk mengevaluasi probabilitas yang dilatih. Dalam penelitian ini menggunakan 50 data yang memiliki 14 atribut, seperti: NIM, nama, jenis kelamin, umur, semester, peminatan, IPK semester 1 hingga IPK semester 5, IPK, keikutsertaan dalam organisasi dan kepemilikan sertifikat penghargaan.

NIM	Nama	Jenis Kelamin	Usia	Semester	Nama Jurusan	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	IPK	Organisasi	Piagam Penghargaan
111	Adi	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.35	3.6	3.65	3.6	3.4	3.6	Ada	Tidak Ada
112	Aji	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.2	3.25	3.3	3.2	3.25	3.24	Tidak Ada	Tidak Ada
113	Ali	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.85	3.9	3.95	3.9	3.9	3.9	Ada	Ada
114	Ani	Perempuan	20 Tahun	5 (lima)	Teknik Informatika	3.3	3.35	3.4	3.35	3.4	3.36	Tidak Ada	Tidak Ada
115	Ari	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.15	3.2	3.25	3.2	3.2	3.2	Tidak Ada	Tidak Ada
116	Aziz	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.75	3.8	3.85	3.8	3.7	3.78	Ada	Ada
117	Badri	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.8	3.85	3.75	3.85	3.8	3.81	Ada	Ada
118	Beni	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.6	3.65	3.6	3.65	3.65	3.63	Ada	Tidak Ada
119	Budi	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.3	3.35	3.3	3.35	3.3	3.32	Tidak Ada	Tidak Ada
120	Caca	Perempuan	20 Tahun	5 (lima)	Teknik Informatika	3.8	3.85	3.9	3.85	3.85	3.85	Ada	Ada
121	Ceri	Perempuan	20 Tahun	5 (lima)	Teknik Informatika	3.35	3.4	3.45	3.4	3.4	3.4	Tidak Ada	Tidak Ada
122	Dika	Perempuan	20 Tahun	5 (lima)	Teknik Informatika	3.6	3.65	3.7	3.6	3.65	3.64	Ada	Tidak Ada
123	Dani	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.15	3.4	3.45	3.4	3.4	3.4	Tidak Ada	Tidak Ada
124	Dadi	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.4	3.45	3.45	3.45	3.45	3.44	Tidak Ada	Tidak Ada
125	Diki	Laki-laki	20 Tahun	5 (lima)	Teknik Informatika	3.15	3.2	3.25	3.2	3.2	3.2	Tidak Ada	Tidak Ada

Gambar 2. Data Training



Setelah menyiapkan data, langkah selanjutnya adalah membersihkannya untuk memastikan tidak ada data yang kosong atau tidak relevan. Validasi data dilakukan dengan teknik 10x cross-validation yang membagi dataset menjadi 70% untuk pelatihan dan 30% untuk pengujian.

3.2. Halaman Sistem Prediksi

Halaman ini dibuat untuk memberikan administrator kemampuan untuk memprediksi hasil mahasiswa berdasarkan data yang sudah tersedia. Administrator memiliki kemampuan untuk memasukkan informasi siswa dan memprediksi lalu menyimpan hasilnya. Gambar di bawah mengilustrasikan desain halaman prediksi ini:



Gambar 3. Halaman Prediksi

3.3. Hasil Perhitungan Naïve Bayes

Tabel 1. Tabel Ketentuan

No.	IPK	Organisasi	Piagam Penghargaan	Keterangan
1	>3.50	Ada	Ada	Berprestasi
2	>3.50	Ada	Tidak Ada	Berprestasi
3	>3.50	Tidak Ada	Ada	Berprestasi
4	<3.49	Ada	Tidak Ada	Tidak Berprestasi
5	<3.49	Tidak Ada	Ada	Tidak Berprestasi
6	<3.49	Ada	Ada	Berprestasi
7	<3.49	Tidak Ada	Tidak Ada	Tidak Berprestasi
8	<3.49	Tidak Ada	Tidak Ada	Tidak Berprestasi

Penyelesaian:

1. Hipotesis

- $X = \text{IPK} > 3.65$, dengan keikutsertaan dalam organisasi dan memiliki piagam penghargaan.
- Apakah benar objek data X dugaan atau hipotesisnya berada pada kelas berprestasi (H)?

2. Menghitung Probabilitas Awal (Prior Probability)

- $P(\text{Berprestasi}) = 4/8 = 0.5$
- $P(\text{Tidak Berprestasi}) = 4/8 = 0.5$

3. Menghitung Probabilitas Bersyarat (Conditional Probability)



- $P(IPK = ">3.65" | Berprestasi) = 4/8$
 - $P(IPK = ">3.65" | Tidak Berprestasi) = 0/8$
 - $P(Organisasi = "Ada" | Berprestasi) = 2/8$
 - $P(Organisasi = "Ada" | Tidak Berprestasi) = 2/8$
 - $P(Piagam = "Ada" | Berprestasi) = 2/8$
 - $P(Piagam = "Ada" | Tidak Berprestasi) = 2/8$
4. Menghitung Probabilitas Posterior (Posterior Probability)
- $P(X | Berprestasi) = P(IPK = ">3.65" | Berprestasi) * P(Organisasi = "Ada" | Berprestasi) * P(Piagam = "Ada" | Berprestasi)$
 - $= (4/8) * (2/8) * (2/8)$
 - $= 0.03125$
 - $P(X | Tidak Berprestasi) = P(IPK = ">3.65" | Tidak Berprestasi) * P(Organisasi = "Ada" | Tidak Berprestasi) * P(Piagam = "Ada" | Tidak Berprestasi)$
 - $= 0 * (2/8) * (2/8)$
 - $= 0$
5. Menghitung Nilai Akhir Untuk Prediksi
- $P(Berprestasi | X) = P(X | Berprestasi) * P(Berprestasi)$
 - $= 0.03125 * 0.5$
 - $= 0.015625$
 - $P(Tidak Berprestasi | X) = P(X | Tidak Berprestasi) * P(Tidak Berprestasi)$
 - $= 0 * 0.5$
 - $= 0$
6. Kesimpulan
- $P(X) = Berprestasi > Tidak Berprestasi (0.015625 > 0)$
 - Dengan demikian, objek data X diklasifikasikan sebagai mahasiswa berprestasi.

4. KESIMPULAN

4.1 Kesimpulan

Penelitian ini menunjukkan bahwa metode Naïve Bayes dapat memprediksi prestasi akademik mahasiswa secara akurat. Penggunaan algoritma ini membantu dalam mengidentifikasi mahasiswa yang berpotensi berprestasi atau tidak, yang dapat digunakan untuk intervensi lebih lanjut oleh universitas.

4.2 Saran

Untuk penelitian selanjutnya, disarankan untuk meningkatkan ketepatan model ini dengan menerapkan beberapa metodologi, misalnya memilih karakteristik yang lebih tepat atau menambahkan informasi yang lebih besar. Selain itu, memilih indikator yang tepat dan mengungkapkan hubungan antar indikator adalah subjek penting untuk meningkatkan akurasi prediksi.



JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 4, September Tahun 2024
ISSN 3025-0919 (media online)
Hal 591-596

REFERENCES

- Annisa, R., & Sasongko, A. (2020). Prediksi Nilai Akademik Mahasiswa Menggunakan Algoritma Naive Bayes. *Jurnal Sains dan Teknologi*.
- Chairina, N. (2023). Penggunaan Metode Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Pendidikan Teknologi Informasi Menggunakan Rapid Miner. (*Skripsi Sarjana, Universitas Islam Negeri Ar-Raniry*, Diakses dari <https://repository.ar-raniry.ac.id>).
- Farida, I. N., & Niswatin, R. K. (2017). Penggunaan Algoritma Naive Bayes Untuk Mengevaluasi Prestasi Akademik Mahasiswa Universitas Nusantara PGRI Kediri. *Jurnal Sains dan Informatika*, 122-127.
- Firdaus, Y. M. (2019). Penerapan Metode Naive Bayes Classifier Untuk Mengklasifikasi Tingkat Prestasi Akademik Santri Pondok Pesantren Mahasiswa (PPM) Baitul Jannah Malang. *JATI Jurnal Mahasiswa Teknik Informatika*, 327-336.
- Hasudungan, R., & Pranoto, W. J. (2021). Implementasi Teorema Naive Bayes Pada Prediksi Prestasi Mahasiswa. *JURTI*, 10-16.
- Puspita, D., & Aminah, S. (2022). Implementasi Naive Bayes Untuk Sistem Prediksi Mahasiswa Berprestasi. *Jurnal Ilmiah Teknosains*, 14-19.