

Desain Aplikasi Naïve Bayes Untuk Memprediksi Kelulusan Siswa

Rahayu Mayang Sari^{*}, Virdyra Tasril, Sri Wahyuni, Shabrina Elsa Putri

Fakultas Sains dan Teknologi, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: ¹*rahayu@dosen.pancabudi.ac.id, ²virdyra@dosen.pancabudi.ac.id, ³yuke@pancabudi.ac.id, ⁴shabrina@gmail.com

Email Penulis Korespondensi: rahayu@dosen.pancabudi.ac.id

Abstrak—Setiap siswa dalam menghadapi ujian sendiri berbeda, ada siswa yang menghadapi dengan serius dan ada yang menghadapinya dengan santai. Peneliti disini merancang sebuah sistem prediksi kesiapan siswa dalam menghadapi ujian nasional berbasis dekstop, yang menggunakan teknik data mining dengan metode naïve bayes. Metode naïve bayes diterapkan untuk mencari probabilitas terbesar dari variabel yang sudah ditentukan meliputi nilai TKJ, Matematika, Bahasa Indonesia, dan Bahasa Inggris untuk mengetahui siap dan kurang siap dalam menghadapi ujian nasional. Hasil dari penelitian ini adalah suatu sistem yang membantu admin dari pihak sekolah untuk mengetahui kesiapan siswa dalam menghadapi ujian nasional. Sehingga dapat menjadikan pertimbangan pihak sekolah untuk meningkatkan minat belajar siswa.

Kata Kunci: Data Mining; Naïve Bayes; Aplikasi; Kelulusan; Siswa;

1. PENDAHULUAN

Pemanfaatan teknologi komputer dapat menjadi sumber informasi penting bagi sekolah, salah satunya informasi penelusuran siswa. Informasi tersebut akan berguna bagi sekolah dalam mendapatkan data siswa yang memilih melanjutkan pendidikan tinggi negeri atau swasta, sehingga dari data tersebut sekolah dapat mengukur kualitas pendidikan yang diberikan pada siswa selama tiga tahun belajar di SMK. Kualitas pendidikan pada siswa merupakan salah satu tujuan dalam mempromosikan sekolah agar mendapatkan kesan yang positif dan dapat dikenal oleh masyarakat luas ataupun calon siswa baru (Susanti et al., 2021).

Dimana siswa mempunyai kewajiban untuk belajar, dalam hal ini harus lebih giat lagi dalam belajar karena akan menghadapi Ujian Nasional (UN). Sebagai salah satu syarat untuk bisa melanjutkan jenjang pendidikan ke tingkat selanjutnya dan seperti yang kita ketahui, bahwa akhir-akhir ini standar kelulusan Ujian Nasional (UN) di Indonesia semakin tinggi. Oleh karena itu, sekolah seharusnya mengetahui apa yang menjadi faktor-faktor yang menentukan tingkat kelulusan siswanya.

Data yang akan digunakan adalah nilai-nilai siswa. Metode Naïve Bayes adalah metode klasifikasi statistik yang dapat memprediksi kelas suatu anggota probabilitas, algoritma ini memanfaatkan teori probabilitas yang dikemukakan oleh ilmuwan Inggris yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sekarang.

Data mining dapat diartikan sebagai proses yang menerapkan ilmu statistik, matematis, artificial intelligence, dan machine learning untuk menemukan pengetahuan berupa informasi yang berguna serta informasi yang saling berhubungan dari beberapa basis data yang besar. Data mining dapat merubah data yang banyak menjadi informasi yang bermanfaat. Informasi tersebut dapat digunakan untuk mengaplikasikan seperti analisis market, pendeteksian kriminal dan lain sebagainya (Sari et al., 2020). Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan dapat diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data (Susanti et al., 2021).

Metode klasifikasi adalah pendekatan Mengkategorikan fungsi data. Data mining adalah suatu proses pengerukan atau pengumpulan informasi penting dari suatu data yang besar. Teknologi Klasifikasi ini juga dapat digunakan untuk Membuat prediksi berdasarkan informasi yang tidak lengkap tahu sebelumnya. beberapa algoritma Salah satu hal yang dapat digunakan adalah algoritma Pohon Keputusan C.45, Jaringan Syaraf Tiruan (ANN), K-Nearest Neighbors (KNN), Algoritma Naive Bayes, Algoritma Genetika, Rough Set, pendekatan berbasis aturan, berbasis memori Mesin Vektor Inferensi dan Dukungan (Subariah & Zein, 2023).

Naive bayes merupakan salah satu metode data mining yang digunakan pada persoalan klasifikasi berdasarkan pada penerapan teorema bayes. Naive bayes akan menghitung probabilitas posterior untuk setiap nilai kejadian dari atribut target pada setiap sampel data. Selanjutnya, Naive bayes akan mengklasifikasikan sampel data tersebut ke kelas yang mempunyai nilai probabilitas posterior tertinggi (Sari et al., 2022).

Naive Bayes digunakan untuk teknik pengklasifikasian data dengan metode kemungkinan dan statistik yang memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antara atribut yang satu dengan atribut yang lainnya itu saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak adanya ciri-ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas yang lainnya (Apridiansyah et al., 2021).

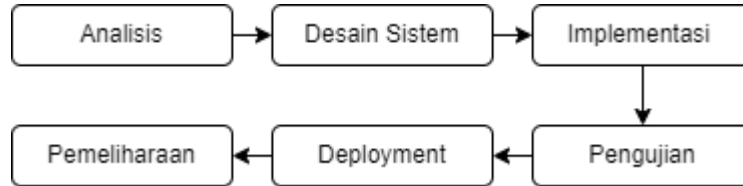
Algoritma yang terdapat pada teknik data mining klasifikasi. Algoritma naive bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris yaitu Thomas bayes, naive bayes memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya, sehingga dikenal dengan teorema bayes. Penelitian dengan menggunakan naïve bayes didasarkan pada nilai probabilitas.

2. METODE PENELITIAN

2.1 Kerangka Dasar Penelitian

Lokasi Penelitian ini dilakukan di SMKS PAB 5 yang berlokasi di Desa Klambir V Kebun, Kecamatan Hamparan Perak, Kabupaten Deli Serdang, Sumatera Utara. Waktu penelitian dilakukan dari bulan Januari - Mei 2024.

2.2 Tahapan Penelitian

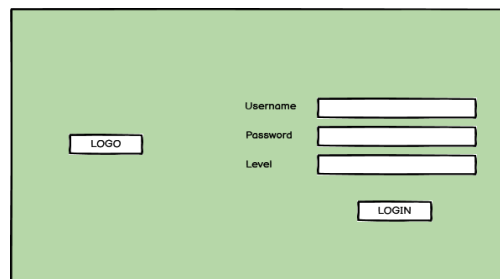


Gambar 1. Tahapan penelitian

3. HASIL DAN PEMBAHASAN

a. Tampilan Login

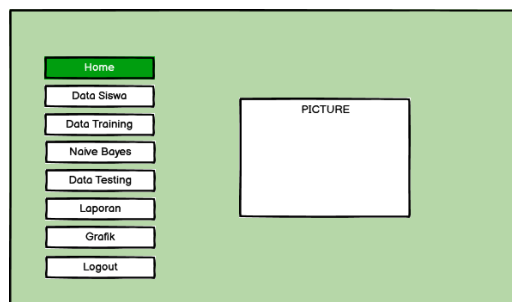
Tampilan login berfungsi untuk verifikasi pengguna yang berhak menggunakan system. Adapun tampilan login adalah sebagai berikut :



Gambar 2. Tampilan Login

b. Tampilan Home

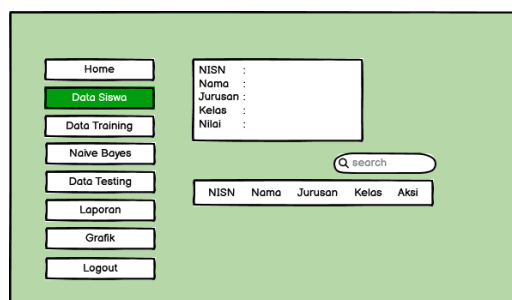
Tampilan home berfungsi untuk menampilkan seluruh isi form-form yang terdapat didalam system. Adapun tampilan home adalah sebagai berikut :



Gambar 3. Tampilan Home

c. Tampilan Data Siswa

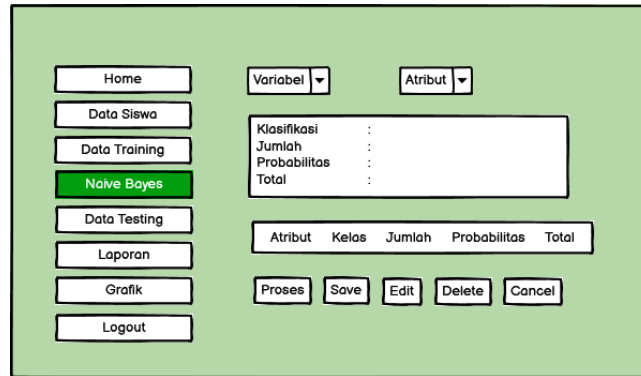
Adapun tampilan data siswa adalah sebagai berikut :



Gambar 4. Tampilan Data Siswa

d. Tampilan Naïve Bayes

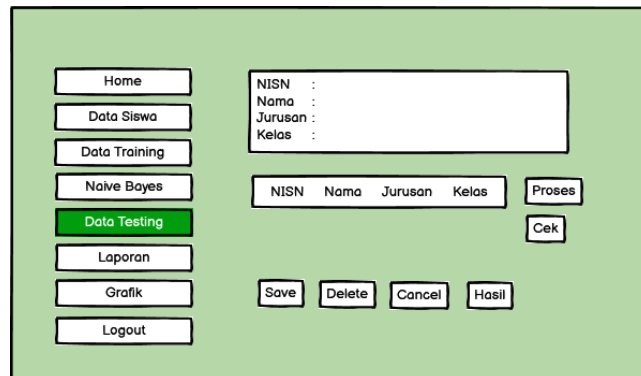
Adapun tampilan pada naïve bayes adalah sebagai berikut :



Gambar 5. Tampilan Naïve Bayes

e. Tampilan Data Testing

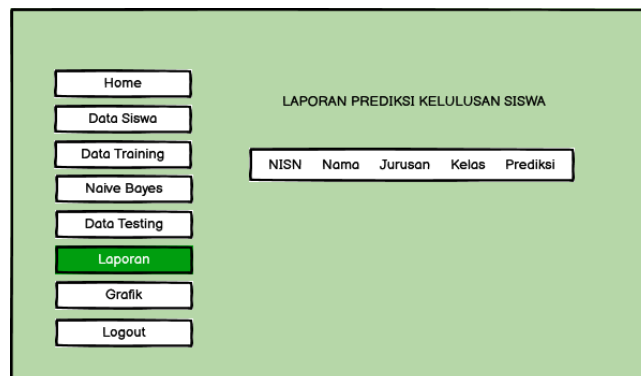
Adapun tampilan data testing adalah sebagai berikut :



Gambar 6 Tampilan Data Testing

f. Tampilan Laporan

Adapun tampilan pada laporan adalah sebagai berikut :



Gambar 7 Tampilan Laporan

3.1 Pembahasan

Algoritma ini memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Dua kelompok peneliti, satu oleh Pantel dan Lin, dan yang lain oleh Microsoft Research memperkenalkan metode statistik Bayesian ini pada teknologi anti spamfilter. Tetapi yang membuat algoritma Bayesian filtering ini populer adalah pendekatan yang dilakukan oleh Paul Graham.

Dasar dari teorema naïve digunakan dalam pemrograman adalah rumus berikut ini : [9] $P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$. Artinya Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B. Contoh penggunaan Algoritma Naïve Bayes antara lain :

1. Untuk klasifikasi dokumen
2. Untuk deteksi SPAM atau fitering SPAM

3. Dan masalah klasifikasi lainnya

Teorema Bayes: $P(C|X) = P(X|C) \cdot P(C) / P(X)$

Di mana :

- a. $P(X)$ bernilai konstan untuk semua class
- b. $P(C)$ merupakan frek relatif sample class C

Dicari $P(C|X)$ bernilai maksimum, sama halnya dengan $P(X|C) \cdot P(C)$ juga bernilai maksimum.

Dalam *Bayes* (terutama *Naive Bayes*) maksud independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidaknya fitur lain dalam data yang sama. Contohnya pada kasus klasifikasi hewan dengan fitur penutup kulit, melahirkan, berat, dan menyusui. Dalam dunia nyata, hewan yang berkembang biak dengan cara melahirkan dTKJstikan juga menyusui. Disini ada ketergantungan pada fitur menyusui karena hewan yang menyusui biasanya melahirkan, atau hewan yang bertelur biasanya tidak menyusui. Dalam *Bayeshal* tersebut tidak dTKJndang sehingga masing-masing fitur seolah tidak memiliki hubungan apapun.

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (1)$$

Tabel 1. Penjelasan Rumus Teorema Bayes

Parameter	Keterangan
$P(H E)$	Probabilitas akhir bersyarat (<i>conditional probability</i>) suatu hipotesis H terjadi jika diberikan bukti (<i>evidence</i>) E terjadi.
$P(E H)$	Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H
$P(H)$	Probabilitas awal (priopri) hipotesis H terjadi tanpa memandang bukti apapun.
$P(E)$	Probabilitas awal (priopri) bukti E terjadi tanpa memandang hipotesis/bukti yang lain.

Ide dasar dari aturan bayes adalah bahwa hasil dari hipotesis (H) dapat diperkirakan berdasarkan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan *Bayes* tersebut , yaitu

1. Sebuah probabilitas awal/priopri H atau $P(H)$ adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2. Sebuah Probabilitas akhir H atau $P(H|E)$ adalah probabilitas dari suatu hipotesis setelah bukti diamati.

Tabel 2 Klasifikasi Nilai Semester 1

Semester 1	Kejadian Klasifikasi	
	Siap	Tidak Siap
A5	0	0
A4	0	0
A3	210	51
A2	0	0
A1	0	0
Jumlah	210	51

Keterangan Probabilitas :

Probabilitas

A5 - Siap

A5 - Tidak Siap

A4 - Siap

A4 - Tidak Siap

A3 - Siap

A3 - Tidak Siap

A2 - Siap

A2 - Tidak Siap

A1 - Siap

A1 - Tidak Siap

Tabel 3 Klasifikasi Nilai Semester 2

Semester 2	Kejadian Klasifikasi	
	Siap	Tidak Siap
B5	0	0
B4	0	0
B3	0	0
B2	0	0
B1	0	0
Jumlah	0	0

Keterangan Probabilitas :

Probabilitas

B5 - Siap

B5 - Tidak Siap

B4 - Siap

B4 - Tidak Siap

B3 - Siap

B3 - Tidak Siap

B2 - Siap

B2 - Tidak Siap

B1 - Siap

B1 - Tidak Siap

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat diambil kesimpulan bahwa Metode Naïve Bayes yang diterapkan pada system yang dibuat dapat menghasilkan nilai probabilitas yaitu Siap dan Tidak Siap dan dengan adanya system prediksi menggunakan metode naïve bayes dapat membantu pihak SMKS PAB 5 Klambir V untuk memprediksi kesiapan siswa untuk menghadapi Ujian Nasional. Diharapkan system ini dapat dikembangkan dengan berbasis webiste atau mobile.

REFERENCES

- Amalia, R. (2020). Penerapan Data Mining Untuk Memprediksi Hasil Kelulusan Siswa menggunakan Metode Naïve Bayes. *Jurnal Informatika Dan Sistem Informasi*, 6(1), 33–42.
- Apridiansyah, Y., Veronika, N. D. M., & Putra, E. D. (2021). Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode Naive Bayes. *JSAI (Journal Scientific and Applied Informatics)*, 4(2), 236–247. <https://doi.org/10.36085/jsai.v4i2.1701>
- Efendi, R., Fitriyani, A., & Purnomo, W. A. (2023). Perancangan Aplikasi Inventory pada Toko Grosir Empat Putra. *INNOVATIVE: Journal Of Social Science Research*, 3(4), 5075–5089. <https://doi.org/10.31004/innovative.v3i4.3990>
- Mubarak, A. (2019). Rancang Bangun Aplikasi Web Sekolah Menggunakan Uml (Unified Modeling Language) Dan Bahasa Pemrograman Php (Php Hypertext Preprocessor) Berorientasi Objek. *JIKO (Jurnal Informatika Dan Komputer)*, 2(1), 19–25. <https://doi.org/10.33387/jiko.v2i1.1052>
- Prasetya, A. F., Sintia, & Putri, U. L. D. (2022). Perancangan Aplikasi Rental Mobil Menggunakan Diagram UML (Unified Modelling Language). *Jurnal Ilmiah Komputer Terapan Dan Informasi*, 1(1), 14–18.
- Ridwan, N. S., & Mardiana, A. (2020). *Perancangan Website Jurnal Dan Laporan Kuliah Kerja Nyata Mahasiswa (Knm) Universitas Majalengka Tahun 2019*. 37–47.
- Sari, R. M., Sains, F., Komputer, J. S., Pembangunan, U., Budi, P., & Medan, K. (2022). *Implementasi Data Mining Untuk Memprediksi Penjualan Menggunakan Metode Penelitian Pendahuluan Pengumpulan Data Analisis dan Perancangan Implementasi Pengujian*. 74–82.
- Sari, R. M., Tasril, V., & M, Y. A. (2020). Prediksi Jumlah APBD Kota Payakumbuh dengan Metode K-Means. *IPTEKS Terapan*, 14(1), 45–50.
- Sari, R. M., Wahyuni, S., & Rizka, A. (2023). Implementasi Algoritma K-Means Untuk Guru Berprestasi. *Jurnal Minfo Polgan*, 12(1), 531–538. <https://doi.org/10.33395/jmp.v12i1.12453>
- Situmorang, N., & Sirait, G. (2020). Implementasi Data Mining untuk Memprediksi Kelulusan Siswa dengan Metode Naive Bayes. *Comasie*, 6(2), 107–118.
- Subariah, R., & Zein, A. (2023). Penerapan Metode Naive Bayes Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu Di Prodi Sistem Informasi Universitas Pamulang. *Sainstech: Jurnal Penelitian Dan Pengkajian Sains Dan Teknologi*, 33(2), 47–51. <https://doi.org/10.37277/stch.v33i2.1585>
- Susanti, L., Rokhmah, I. T., & Stianingsih, L. (2021). *Prediksi Siswa SMK Al-Hidayah yang Masuk Perguruan Tinggi dengan Metode Klasifikasi*. 8(1), 26–33

Prosiding Seminar Nasional Teknologi Inovasi dan Kolaborasi Disiplin Ilmu

Vol 1, No 1, Oktober 2024, Hal 1-6

ISSN 3090-3505 (Media Online)

Website <https://journal.fkpt.org/index.php/sinekad>

- Suwarno, A., Ferawati, N., Sari, P. A., Studi, P., Informatika, T., Teknik, F., Pelita, U., Orientasi, O., & Barang, S. (2021). Jurnal Teknologi Pelita Bangsa. *Jurnal Teknologi Pelita Bangsa*, 12(4), 33–40.
- Wahyuni, S., Saragih, kana S., & Perangin-angin, M. I. (2018). Implemntasi Metode Decision Tree C4.5 Untuk Menganalisa Mahasiswa Dop Out. *Ethos*, 6(1), 42–51. https://web.archive.org/web/20180414185850id_/https://ejournal.unisba.ac.id/index.php/ethos/article/viewFile/3252/pdf
- Nuranisah, & Yanti Yusman. (2023). Naïve Bayes Classifier dengan Particle Optimize Weight Forward pada Dataset. *Bulletin of Computer Science Research*, 3(6), 442–447. <https://doi.org/10.47065/bulletincsr.v3i6.290>