

---

## PREDIKSI TINGKAT KELULUSAN PESERTA DIDIK SMK FATHUL ULUM GABUS DENGAN METODE NAIVE BAYES

Wahyudi<sup>1</sup>; Eko Supriyadi<sup>2</sup>; Andri Triyono<sup>3</sup>

---

### ABSTRACT

*The graduation of students refers to those who are able to complete and meet the graduation requirements set through a graduation meeting based on the decision letter signed by the school principal. Graduation rate data can be used to help make policies and strategies for the school to improve graduation rates in the following year. This study utilizes classification or prediction methods to analyze the graduation rates of students at SMK Fathul Ulum Gabus. The method used in this study is Naive Bayes, using variables such as practical exam scores, school exam scores, competency test scores, student attendance, and student behavior. The purpose of this study is to test the accuracy of the Naive Bayes method in predicting graduation rates based on data collected from 2019 to 2024. The research process includes data collection, data integration, and model training using Naive Bayes, which produces fairly accurate predictions with an accuracy of 94.64%. Based on this accuracy, it can be concluded that the Naive Bayes method can be used to predict graduation rates at SMK Fathul Ulum Gabus.*

**Keywords:** *Graduation rates prediction1; Naive bayes2*

---

#### **Correspondence:**

Wahyudi

Universitas An Nuur, Email; [wahyudhi2102@gmail.com](mailto:wahyudhi2102@gmail.com)

---

### PENDAHULUAN

Kelulusan peserta didik adalah mereka yang mampu menyelesaikan dan memenuhi persyaratan kelulusan yang ditetapkan dalam rapat kelulusan melalui surat keputusan dari hasil rapat yang di tanda tangani oleh Kepala Sekolah. Tingkat kelulusan digunakan untuk membuat kebijakan baru dan atau menentukan strategi sekolah sehingga meningkatkan kelulusan pada tahun selanjutnya bisa diketahui dari data yang sudah dikelola.

Untuk melakukan tingkat kelulusan pada peserta didik dalam suatu sekolah dapat dilakukan suatu klasifikasi atau prediksi.

SMK Fathul Ulum Gabus merupakan salah satu Sekolah Menengah Kejuruan yang berdiri sejak tahun 2014. Dalam masalah ini sering kali terjadi kesalahan pengolahan nilai peserta didik untuk pengolahan tingkat kelulusan peserta didik, yang dimana dalam proses yang terkadang masih salah mengakibatkan pihak sekolah kesulitan dalam pengolahan kelulusan

peserta didik tersebut dan terjadinya kesalahan yang tidak akurat.

Metode *Naïve Bayes* dapat memanfaatkan data nilai Ujian Sekolah, Uji Kompetensi Keahlian, Ujian Praktik, perilaku peserta didik, dan kehadiran peserta didik untuk menentukan faktor utama penunjang tingkat kelulusan. Dengan analisis ini, sekolah dapat memahami pola kelulusan, menyusun strategi yang tepat, meningkatkan kualitas pendidikan, dan memperkuat daya saing sekolah.

Hal ini akan membantu pihak sekolah untuk menentukan tingkat kelulusan peserta didik dan mengetahui pola kelulusan peserta didik dengan memanfaatkan nilai atau data peserta didik tersebut, sehingga membantu pihak sekolah untuk menyusun strategi yang tepat dalam meningkatkan kualitas sekolah dan menjadikan sekolah memiliki daya saing yang tinggi.

## METODE PENELITIAN

### Tempat dan Waktu

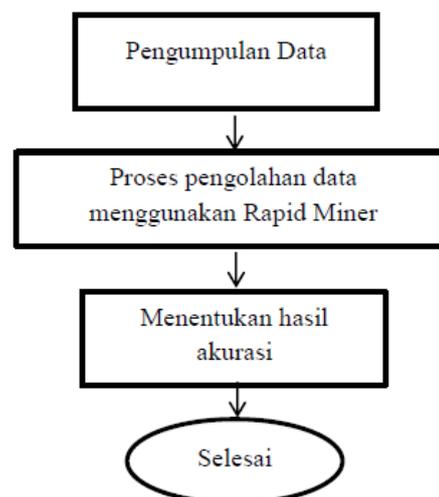
Tempat penelitian ini adalah SMK Fathul Ulum Gabus yang beralamat di Jl. Sulursari No. 20 Desa Pandanharum Kecamatan Gabus Kabupaten Grobogan. Waktu penelitian dilakukan pada rentang waktu bulan Mei – Agustus 2024.

### Bahan dan Alat Penelitian

Hardware berupa Laptop merk Acer dan Hp merk OPPO A12. Software berupa *Microsoft Windows 10, Google Chrome, RapidMiner*.

### Desain Penelitian

Penelitian ini menggunakan metode *naïve bayes* yang dirancang untuk digunakan untuk mengklasifikasikan data, dengan tahapan desain penelitian sebagai berikut :



Gambar 1. Metode penelitian

Metode penelitian ini di mulai dengan pengambilan data pada sekolah SMK Fathul Ulum Gabus, pengambilan data di lakukan secara wawancara secara langsung kepada narasumber dari salah satu guru di sekolah tersebut. Data data yang akan dimasukan dan dilakukan pengolahan

yaitu data Nilai Praktik, Nilai US (Ujian Sekolah), Nilai UKK (Uji Kompetensi Keahlian), Perilaku Peserta didik dan Kehadiran Peserta didik dalam bentuk microsoft excel. Data tersebut lalu akan di olah terlebih dahulu melalui tahap penggabungan data dimana data tersebut dijadikan satu file sehingga dapat di olah pada aplikasi rapid miner.

Setelah data di olah pada aplikasi rapid miner langkah selanjutnya yaitu tahap pemilihan atribut, setelah pemilihan atribut selesai dilakukan, langkah selanjutnya yaitu pengujian model dari data yang didapatkan menggunakan metode naive bayes. Setelah data di olah akan di dapatkan tabel confusion matrix yang menunjukkan akurasi dari pengolahan data.

## HASIL DAN PEMBAHASAN

### Tahap Pengumpulan Data

Pengumpulan data dilakukan dengan melakukan observasi langsung sekolah SMK Fathul Ulum Gabus data yang di dapatkan berupa data kelulusan dari tahun 2019 sampai 2024. Data yang didapat berupa 5 file dokumen, dengan setiap file nya berisi data kelulusan peserta didik setiap tahun yang mencakup nilai ujian sekolah, nilai ujian praktik, nilai UKK (Uji Kompetensi Keahlian), perilaku peserta didik, kehadiran peserta didik dan

keterangan lulus/tidak lulus dengan jumlah data 187 data peserta didik. Gambar di atas merupakan salah satu file dokumen mentah yang di dapat dari tenaga administrasi berbentuk Excel.

### Tahap Penggabungan Data

Pada penggabungan data peneliti menggabungkan dari 5 file dokumen yang berbeda menjadi 1 file dokumen untuk memudahkan peneliti dalam mengolah data.

Tabel 1. Penggabungan data

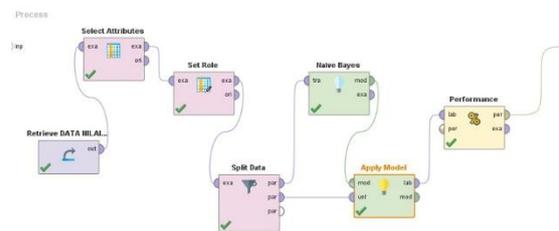
No	Nama	Niali Ujian	Nilai Ujian	Nilai UKK	Perilaku Peserta	Rata-Rata	Kehadiran	Keterangan
1	Aditya Danu Pratama	8 8 5	8 8 5	8 8 2	8 8 2	8 3, 5	1	LU LU S
2	Ahmad Arif Muzaka	8 8 7	8 8 3	8 8 6	8 8 0	8 8 4	9	LU LU S
3	Ahmad Aurelio Fahrezi	8 8 6	8 8 5	9 9 2	8 8 2	8 6, 3	1 0	TI DA K LU LU S
4	Ahmad Mukht	8 8	8 5	9 0	8 4	8 6, 8	7	LU LU S

	ar Jamil							
5	Ahmad Sarifudin	8 6	8 5	8 4	8 1	8 4	2	LU LU S
6	Kurnia Bayu Afandi	8 6	8 7	8 3	8 4	8 5	6	LU LU S
7	Miftah ul Munir	8 8	9 0	9 0	8 5	8 3	5	LU LU S
8	Mohammad Alwy Jawahirul Bahri	8 6	8 5	8 7	8 2	8 5	5	LU LU S
9	Muhammad Aldianto	8 6	8 5	8 8	8 2	8 3	5	LU LU S
10	Muhammad Faisal Rizki	8 5	8 3	8 5	8 1	8 3	2	LU LU S
11	.....	...	...	...	...	...	...	.....
12	.....	...	...	...	...	...	...	.....
13	.....	...	...	...	...	...	...	.....
...	.....	...	...	...	...	...	...	.....
...	.....	...	...	...	...	...	...	.....
187	Wahid	8 9	8 7	8 0	8 5	8 3	2	LU LU S

	Romad hona						
--	---------------	--	--	--	--	--	--

Data yang didapat berupa 5 file dokumen, dengan setiap file nya berisi data kelulusan peserta didik setiap tahun yang mencakup nilai ujian sekolah, nilai ujian praktik, nilai UKK (Uji Kompetensi Keahlian), perilaku peserta didik, kehadiran peserta didik dan keterangan lulus/tidak lulus dengan jumlah data 187 data peserta didik.

**Model Rapid Miner**



Gambar 2. Model rapid miner

Gambar di atas merupakan alur proses pada aplikasi rapid miner, prosesnya adalah sebagai berikut *Retrieve data, Select Attributes, Set Role, Split Data, Naive Bayes, Apply Model* dan *Perfomance*.

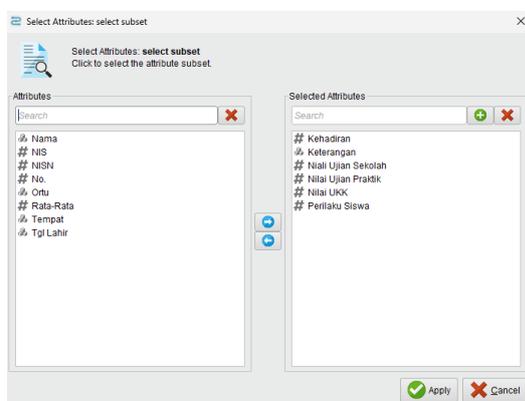
**Retrieve Data**

*Retrieve data* merupakan proses mengambil dan memlih data dari berbagai sumber dan memasukkannya ke dalam *RapidMiner* untuk analisis dan pemrosesan dengan menggunakan algoritma *naive*

bayes. Data yang di masukan adalah data yang sudah di gabungkan dari 5 file data yang berbeda.

### Tahap Pemilihan Atribut

Pada tahap ini akan di lakukan pemilhan atribut yang di olah nantinya menggunakan algoritma *naïve bayes*, atribut yang di pilih yaitu nilai ujian sekolah, niali ujian praktik, nilai UKK (Uji Kompetensi Keahlian), perilaku peserta didik, kehadiran peserta didik dan keterangan sebagai labelnya.

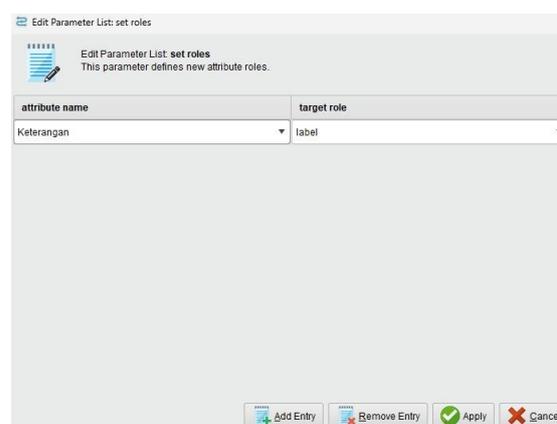


Gambar 3. Pemilihan atribut

Gambar diatas merupakan proses pemilihan atribut/variabel, kolom di sebelah kiri yaitu nama, nis, nisn, no, ortu, rata-rata, tempat dan tgl lahir merupakan data yang tidak di gunakan dan kolom sebelah kanan yaitu nilai ujian sekolah, nilai ujian praktik, nilai UKK, kehadiran dan perilaku peserta didik adalah variabel yang akan digunakan, berikut ini adalah hasil dari select attributes.

### Set Role

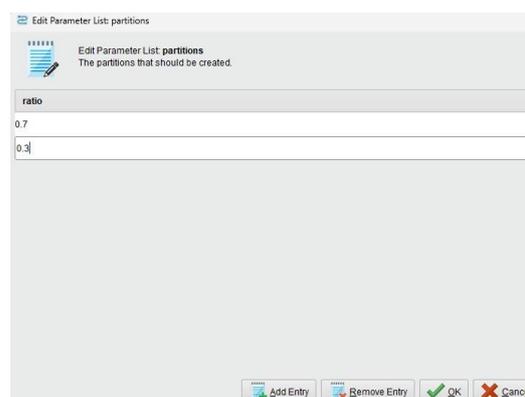
*Set role* di lakukan pada variabel dependen, hal ini di lakukan agar aplikasi rapid miner mengetahui variabel dependen adalah variabel target yang akan di lakukan prediksi. Berikut adalah penyetulan label pada variabel keterangan.



Gambar 4. Pemilihan label

### Split Data

*Split data* merupakan proses pemisahan data menjadi dua bagian yaitu memisahkan data yang berjumlah 187 baris menjadi dua set data, data tersebut akan di pisah dengan rasio 70% untuk training dan 30% untuk testing.



Gambar 5. Pemisahan data

Gambar di atas adalah pengaturan rasio pemisahan data pada rapid miner, dimana data training akan di olah dengan algoritma *naive bayes* dan data testing di gunakan untuk membandingkan hasil olahan dan data asli sehingga menghasilkan akurasi.

### Penerapan *Naïve Bayes*

Pada tahap ini di lakukan pemrosesan data dengan algoritma *naïve bayes*.

### SimpleDistribution

Distribution model for label attribute Keterangan

```
Class LULUS (0.924)
5 distributions

Class TIDAK LULUS (0.076)
5 distributions
```

Gambar 6. Hasil pemrosesan algoritma *naive bayes*

Dari gambar di atas, dapat di intepretasikan dari semua data training yang di modelkan, 0,924 (92,4%) dari peserta didik termasuk dalam kelas “Lulus”, Sedangkan 0,076 (7,6%) dari semua data yang dimodelkan termasuk dalam kelas “Tidak Lulus” dengan menggunakan empat variabel yang berbeda.

### Apply Model

*Apply model* adalah perbandingan antara data prediksi dan data testing untuk

menghasilkan peforma dari algoritma yang di gunakan.

Row No.	Keterangan	prediction(K...	confidence(...	confidence(...
1	LULUS	LULUS	1.000	0.000
2	TIDAK LULUS	LULUS	0.990	0.010
3	LULUS	LULUS	1.000	0.000
4	LULUS	LULUS	1.000	0.000
5	LULUS	LULUS	1.000	0.000
6	LULUS	LULUS	0.999	0.001
7	LULUS	LULUS	0.938	0.062
8	LULUS	LULUS	1.000	0.000
9	LULUS	LULUS	1.000	0.000
10	LULUS	LULUS	0.999	0.001
11	LULUS	LULUS	1.000	0.000
12	LULUS	LULUS	0.998	0.002
13	LULUS	LULUS	1.000	0.000
14	LULUS	LULUS	0.997	0.003

Gambar 7. Perbandingan data testing dan prediksi

### Perfomance

Setelah data di bandingkan akan di dapatkan tabel confusion matrix sebagai berikut.

accuracy: 94.64%

	true LULUS	true TIDAK LULUS	class precision
pred: LULUS	50	1	98.04%
pred: TIDAK LULUS	2	3	60.00%
class recall	96.15%	75.00%	

Gambar 8. *Perfomance confusion matrix*

Dari gambar diatas, dapat di simpulkan akurasi dari proses penerapan algoritma *naive bayes* untuk memprediksi tingkat kelulusan peserta didik meiliki nilai 94,64%. Dengan hasil ini, maka algoritma *naive bayes* dapat memprediksi tingkat kelulusan.

## SIMPULAN DAN SARAN

### SIMPULAN

Berdasarkan hasil penelitian terhadap tingkat kelulusan peserta didik di SMK Fathul Ulum Gabus yang dilakukan dapat di ambil kesimpulan data kelulusan peserta didik SMK Fathul Ulum Gabus dapat di prediksi menggunakan algoritma naïve bayes dan mendapatkan akurasi sebesar 94,64%.

### SARAN

Peneliti selanjutnya diharapkan untuk mengkaji lebih banyak sumber maupun referensi yang terkait dalam penelitian ini, serta bisa mengembangkan dengan cara menggabungkan atau membandingkannya dengan algoritma yang lain untuk mendapatkan hasil prediksi yang lebih akurat.

### DAFTAR PUSTAKA

- Amri, S. (2020). *Information Science and Library Perbandingan Kerangka Model Klasifikasi untuk Pemilihan Metode Kontrasepsi dengan Pendekatan CRIPS-DM* Info Artikel \_\_\_\_\_ e-ISSN. 1(1), 14–23.  
<https://doi.org/10.26623/jisl.v1i1>
- Amrulloh, A., Permana Wibowo Program Studi Magister Teknologi Informasi, A., & Teknologi Yogyakarta Jl Siliwangi Jombor, U. (2019). IMPLEMENTASI ALGORITMA DECISSION TREE UNTUK MENGLASIFIKASI KONDISI KESUBURAN PRIA. In *JASIEK* (Vol. 1, Issue 1).
- Bagus, I., Peling, A., Arnawan, N., Putu, I., Arthawan, A., & Janardana, I. (n.d.). Implementation of Data Mining To Predict Period of Students

Study Using Naive Bayes Algorithm. In *International Journal of Engineering and Emerging Technology* (Vol. 2, Issue 1).

- Bidang Komputer Sains dan Pendidikan Informatika, P., Akademi Perkam dan Informasi Kesehatan Iris Padang Jl Gajah Mada No, D., & Barat, S. (n.d.). *Jurnal Edik Informatika Data Mining: Klasifikasi Menggunakan Algoritma C4.5 Yuli Mardi*.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Farid Rifai, M., Jatnika, H., Valentino, B., & Tinggi Teknik PLN, S. (2019). *Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS)*. 12(2).
- Harapan, U., & Rismayanti, M. ). (2018). Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika (Studi Kasus: Fakultas Teknik dan Komputer. *Jurnal Sistem Informasi*, 1.
- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3).  
<https://doi.org/10.30998/faktorexacta.v11i3.2777>
- Indraswari, N. R., & Kurniawan, Y. I. (2018). APLIKASI PREDIKSI USIA KELAHIRAN DENGAN METODE NAIVE BAYES. *Jurnal SIMETRIS*, 9(1).
- Marathe, N., Gawade, S., & Kanekar, A. (2021). Prediction of Heart Disease and Diabetes Using Naive Bayes Algorithm. *International Journal of Scientific Research in Computer*

- Science, Engineering and Information Technology*, 447–453.  
<https://doi.org/10.32628/cseit217399>
- Meilani, B. D., & Susanti, N. (2015). APLIKASI DATA MINING UNTUK MENGHASILKAN POLA KELULUSAN SISWA DENGAN METODE NAÏVE BAYES. In *Jurnal Ilmiah NERO* (Vol. 1, Issue 3).
- Mergono Adi Ningrat, Carolina Diana Mual, & Yohanis Yan Makabori. (2021). Pertumbuhan dan Hasil Tanaman Padi (*Oryza sativa* L.) pada Berbagai Sistem Tanam di Kampung Desay, Distrik Prafi, Kabupaten Manokwari. *Prosiding Seminar Nasional Pembangunan Dan Pendidikan Vokasi Pertanian*, 2(1), 325–332.  
<https://doi.org/10.47687/snppvp.v2i1.191>
- Moh. Rodiaminollah, & Lu'lu' Nurul Qomariyah. (2023). PEMBANGUNAN SEKTOR PERTANIAN DALAM PERENCANAAN PEMBANGUNAN EKONOMI. *Istisodina Jurnal Ekonomi Syariah & Hukum Islam*, 6(1), 49–55.
- Mulyati, S., Maulana Husein, S., & Kunci, K. (2020). RANCANG BANGUN APLIKASI DATA MINING PREDIKSI KELULUSAN UJIAN NASIONAL MENGGUNAKAN ALGORITMA (KNN) K-NEAREST NEIGHBOR DENGAN METODE EUCLIDEAN DISTANCE PADA SMPN 2 PAGEDANGAN sistem dapat memprediksi dan mengklasifikasikan dengan baik dan cepat. 65–73.
- Pasaribu, L. (2019). SISTEM PAKAR MENDIAGNOSA HAMA DAN PENYAKIT TANAMAN MENTIMUN MENGGUNAKAN METODE NAÏVE BAYES. *Jurnal Pelita Informatika*, 7(3).