



INNOVATIVE: Journal Of Social Science Research

Volume 3 Nomor 2 Tahun 2023 Page 9363-9372

E-ISSN 2807-4238 and P-ISSN 2807-4246

Website: <https://j-innovative.org/index.php/Innovative>

Mengimplementasikan Metode Naive Bayes Untuk Memprediksi Tingkat Kelulusan Mahasiswa Tepat Waktu

Nur Aini H, Hengky P^{1✉}, M Amin Yusuf, Frastio, M Ali Akbar

Sistem Informasi, Universitas Prabumulih

Email: Ainihatagalung*@gmail.com^{1✉}

Abstrak

Tingkat kesuksesan Universitas untuk mempertahankan mutu pembelajaran dan kualitas belajar mengajar dapat diketahui dari banyaknya jumlah mahasiswa yang lulus dengan predikat tinggi dan tepat waktu. Penelitian ini bertujuan untuk mengidentifikasi mahasiswa yang dapat lulus tepat waktu. Fokus penelitian ini adalah pada mahasiswa program Sistem Informasi di Universitas Prabumulih. Metode yang digunakan dalam penelitian ini adalah metode Naive Bayes, yang merupakan metode klasifikasi berbasis probabilitas. Data yang digunakan dalam penelitian ini mencakup informasi tentang mahasiswa Sistem Informasi seperti IPK, jumlah SKS yang diambil per semester, kehadiran, dan atribut lain yang relevan. Analisis tingkat kelulusan adalah salah satu faktor pendukung dalam memperbaiki mutu dan layanan pendidikan pada lingkup Universitas Prabumulih, oleh karena itu penelitian ini menjadi hal yang sangat penting untuk dilakukan. Hasil penelitian ini dapat disimpulkan bahwa Visi dan Misi yang telah dijalankan oleh Universitas Prabumulih Jurusan Sistem Informasi Angkatan 2020, 2021, 2022 akan menghasilkan sarjana tepat waktu bahkan lebih cepat dari waktunya.

Kata Kunci: *Naive Bayes, Klasifikasi, Universitas Prabumulih*

Abstract

The university's level of success in maintaining the quality of learning and teaching and learning can be seen from the large number of students who graduate with high honors and on time. This study aims to identify students who can graduate on time. The focus of this research is on students of the Information Systems program at Prabumulih University. The method used in this study is the Naive Bayes method, which is a probability-based classification method. The data used in this study includes information about Information Systems students such as GPA, number of credits taken per semester, attendance, and other relevant attributes. Analysis of the graduation rate is one of the supporting factors in improving the quality and service of education at Prabumulih University, therefore this research is very important to do. The results of this study can be concluded that the Vision and Mission that have been implemented by Prabumulih University Information Systems Department Class of 2020, 2021, 2022 will produce graduates on time even faster than the time.

Keywords: *Naive Bayes, Classification, Prabumulih University*

PENDAHULUAN

Indeks prestasi kumulatif Mahasiswa pada universitas yang dikelola dan di susun secara strategis dapat memberikan sebuah informasi yang detail dan memiliki analisis yang baik tentang tingkat prestasi akademik mahasiswa yang di capai. secara khusus dan keberhasilan studi mahasiswa pada capaian tingkatan kelulusan setiap Universitas Prabumulih harus meningkatkan kualitas pendidikannya agar diharapkan dapat memiliki lulusan yang dapat mengimplementasikan ilmu yang di didapatkan ke dunia. Di suatu tujuan untuk mendapatkan hasil yang sangat memuaskan dengan tingkat kelulusan yang tepat waktu bahkan lebih cepat semua itu kembali lagi kepada individu mahasiswa itu sendiri. Bentuk analisa dari data keberhasilan studi/kelulusan mahasiswa berupa predikat kelulusan yang diprediksikan dalam kriteria "Predikat Lulus Cepat", "Predikat tepat waktu", dan " Kelulusan terlambat ". Predikat kelulusan saat ini dijadikan sebuah pedoman oleh lulusan Pencapaian predikat kelulusan mahasiswa dapat ditentukan dengan cara berikut, yaitu membuat kelompok data dari Jumlah total sks yang diambil oleh mahasiswa, indeks kumulatif dari mahasiswa. Pada penelitian ini dilakukan analisis prediksi predikat indeks prestasi kelulusan mahasiswa dengan penggunaan (Model & pengembangan sistem) yang pada akhirnya dapat dimanfaatkan oleh Universitas Prabumulih dalam mengatasi persoalan yang ada, sehingga perguruan tinggi dapat berkompetisi dengan perguruan tinggi lainnya dan dapat meningkatkan kualitas akademik dari mahasiswa pada Universitas tersebut.

1.1. Tingkat Kelulusan Mahasiswa Tepat Waktu

Tingkat kelulusan mahasiswa tepat waktu merupakan salah satu indikator penting dalam sistem pendidikan tinggi. Kelulusan tepat waktu mengacu pada mahasiswa yang berhasil menyelesaikan semua persyaratan akademik yang ditetapkan dalam kurun waktu yang telah ditentukan, seperti dalam jangka waktu standar untuk menyelesaikan suatu program studi. Tingkat kelulusan tepat waktu yang tinggi mencerminkan efisiensi pendidikan, sementara tingkat kelulusan yang rendah dapat menunjukkan adanya masalah dalam sistem pendidikan.

Terdapat beberapa faktor yang mempengaruhi tingkat kelulusan mahasiswa tepat waktu. Beban studi yang berlebihan atau tidak teratur dapat menjadi hambatan bagi mahasiswa dalam menyelesaikan program studi tepat waktu. Selain itu, faktor keuangan juga dapat memengaruhi kelulusan tepat waktu, seperti kesulitan dalam membiayai pendidikan atau pekerjaan paruh waktu yang mengganggu fokus pada studi. Dukungan akademik yang memadai dari universitas dan fakultas juga berperan penting, termasuk bimbingan akademik, layanan konseling, dan mentoring.

Kualitas pengajaran juga memiliki dampak yang signifikan terhadap tingkat kelulusan tepat waktu. Pengajaran yang baik dan efektif dapat membantu mahasiswa memahami materi dengan lebih baik, meningkatkan motivasi, dan mengurangi tingkat pengulangan mata kuliah. Faktor pribadi mahasiswa, seperti motivasi, disiplin, manajemen waktu, dan kemampuan belajar, juga berperan dalam tingkat kelulusan tepat waktu. Selain faktor internal mahasiswa, kondisi lingkungan juga dapat mempengaruhi tingkat kelulusan mahasiswa tepat waktu. Faktor-faktor seperti dukungan keluarga, lingkungan sosial, dan faktor eksternal seperti kesempatan kerja atau kesempatan magang juga dapat memengaruhi prioritas dan komitmen mahasiswa terhadap studi. Memahami faktor-faktor ini dan mengidentifikasi faktor-faktor yang paling berpengaruh dalam tingkat kelulusan mahasiswa tepat waktu sangat penting. Dengan pemahaman yang baik tentang faktor-faktor tersebut, universitas dan fakultas dapat mengimplementasikan strategi dan intervensi yang tepat untuk meningkatkan tingkat kelulusan mahasiswa tepat waktu.

1.2. Metode Naive Bayes

Metode Naive Bayes adalah salah satu metode klasifikasi yang populer dalam bidang pembelajaran mesin. Metode ini didasarkan pada teorema Bayes dan asumsi independensi antar-variabel. Naive Bayes bekerja dengan menghitung probabilitas kelas yang paling

mungkin berdasarkan nilai-nilai fitur yang diamati. Meskipun memiliki asumsi yang sederhana, metode Naive Bayes telah terbukti efektif dalam banyak aplikasi, termasuk dalam prediksi klasifikasi. Metode Naive Bayes cocok untuk digunakan dalam konteks prediksi tingkat kelulusan mahasiswa tepat waktu. Dalam penggunaannya, metode ini membutuhkan data historis tentang mahasiswa dan variabel-variabel yang berpotensi mempengaruhi kelulusan. Variabel-variabel ini dapat mencakup indeks prestasi, jumlah mata kuliah yang diambil per semester, kehadiran, dan faktor lain yang relevan. Langkah pertama dalam mengimplementasikan metode Naive Bayes adalah mengumpulkan data historis tentang mahasiswa yang telah lulus tepat waktu dan yang tidak lulus tepat waktu. Data ini kemudian akan digunakan untuk melatih model prediktif. Data tersebut perlu diproses secara pra-pemrosesan, termasuk dalam hal pembersihan data dan penyeimbangan dataset jika diperlukan. Setelah pra-pemrosesan data, langkah berikutnya adalah menerapkan metode Naive Bayes sebagai algoritma klasifikasi. Metode ini mengasumsikan independensi antar-variabel, yang berarti variabel-variabel yang digunakan dalam prediksi dianggap saling independen satu sama lain. Dalam tahap pelatihan, data historis yang diketahui status kelulusannya digunakan untuk melatih model Naive Bayes.

Proses pelatihan ini melibatkan perhitungan probabilitas kelas berdasarkan nilai-nilai fitur yang diamati. Setelah model pelatihan selesai, model tersebut dapat digunakan untuk memprediksi tingkat kelulusan mahasiswa baru berdasarkan data masukan yang diberikan. Model Naive Bayes akan menghitung probabilitas kelas yang paling mungkin berdasarkan nilai-nilai fitur mahasiswa baru. Mahasiswa akan diklasifikasikan sebagai lulus tepat waktu atau tidak lulus tepat waktu berdasarkan probabilitas kelas yang diperoleh. Evaluasi kinerja model Naive Bayes dilakukan dengan menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Metrik-metrik ini memberikan gambaran tentang sejauh mana model dapat memprediksi dengan benar tingkat kelulusan mahasiswa tepat waktu. Analisis sensitivitas juga dapat dilakukan untuk memahami pengaruh variabel-variabel input terhadap prediksi tingkat kelulusan. Metode Naive Bayes merupakan alat yang kuat dalam prediksi tingkat kelulusan mahasiswa tepat waktu. Kelebihan metode ini meliputi kemampuannya untuk menangani data yang besar, skalabilitas yang baik, serta interpretabilitas yang tinggi. Namun, perlu diingat bahwa metode Naive Bayes memiliki asumsi independensi antar-variabel

Teorema Bayes memiliki bentuk umum sebagai berikut (Jananto, 2013):

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

Keterangan:

X= Data dengan class yang belum diketahui

H= Hipotesis data X merupakan suatu kelas spesifik

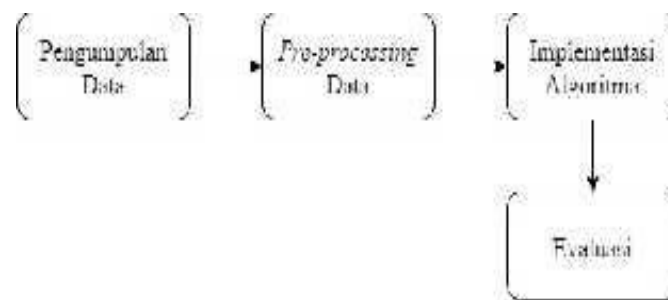
P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X

P(H)= Probabilitas hipotesis H

METODE PENELITIAN

Metodologi yang digunakan secara keseluruhan dapat dilihat pada Gambar 1. Terdapat beberapa tahap yaitu Pengumpulan Data, Pre-processing Data dan Implementasi Algoritma

Gambar 1 Metodologi Penelitian



A. Pengumpulan Data

Dataset yang digunakan pada penelitian ini adalah data mahasiswa Sistem Informasi Universitas Prabumulih tahun angkatan 2020 berjumlah 36 data

Contoh dataset yang digunakan dapat dilihat pada tabel 1.

Tabel 1. Contoh Dataset

| No | Atribut | Nilai |
|----|----------------|--------|
| 1 | Tahun Masuk | Int |
| 2 | Jenis Kelamin | L P |
| 3 | SKS Semester 1 | Int |
| 4 | SKS Semester 2 | Int |
| 5 | SKS Semester 3 | Int |

| | | |
|---|----------------|-------------|
| 6 | SKS Semester 4 | Int |
| 7 | SKS Semester 5 | Int |
| 8 | SKS Semester 6 | Int |
| | | Cepat |
| 9 | IPK Semester 6 | Tepat Waktu |
| | | Terlambat |

B. Pre-processing Data

Jenis Pre-processing data yang digunakan adalah transformasi data. Nilai IPS dari semester 1 sampai 6 dihitung rata-rata untuk mendapatkan IPK semester 6. Kemudian data dirubah nilainya menjadi "Cepat", "Tepat Waktu", atau "Terlambat" berdasarkan hasilnya. Transformasi data dapat dilihat pada tabel 2.

Tabel 2. Transformasi Data

| IPK Semester 6 | Konveksi |
|----------------|-------------|
| IPK 1,0 – 2,99 | Terlambat |
| IPK 3,0 – 3.49 | Tepat Waktu |
| IPK 3,50 – 4,0 | Cepat |

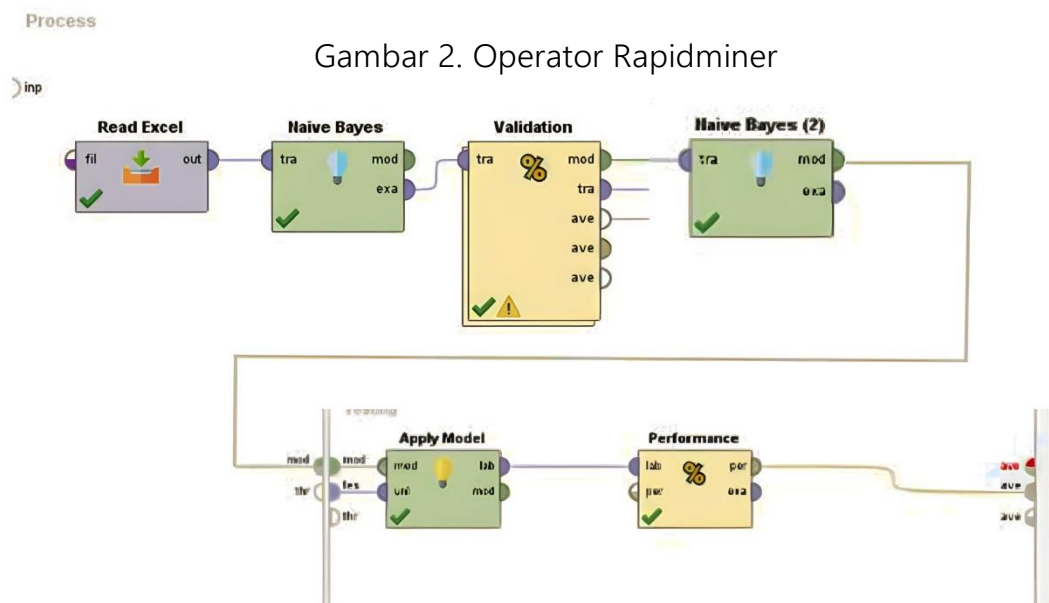
Tabel 3. Dataset

| Nama | JJ K | Kela s | Tahun Masu k | SKS Smstr 1 | SKS Smstr 2 | SKS Smstr 3 | SKS Smstr 4 | SKS Smstr 5 | IPK Smstr 6 |
|----------------------|---------|-----------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|----------------|
| Muhammad Amin Yusuf | L | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Frastio Okti Rachmat | L | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Reisya Kurnia | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Cepat |

| | | | | | | | | | |
|----------------|---|---|------|----|----|----|----|----|-------------|
| Riski Ulyani | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Cepat |
| Niken Rahmanti | P | B | 2020 | 20 | 24 | 22 | 24 | 24 | Tepat Waktu |
| Nesya Olivia | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Fadilah Riski | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Dhea Natasya | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Desnia Angelia | P | B | 2020 | 20 | 24 | 24 | 24 | 24 | Tepat Waktu |
| Carles | L | B | 2020 | 20 | 24 | 24 | 20 | 24 | Terlambat |

C. Implementasi Algoritma

Pada tahap ini, akan dilakukan implementasi algoritma Naïve Bayes pada data mahasiswa tahun angkatan 2020 menggunakan RapidMiner. Proses implementasi algoritma Naïve Bayes dapat dilihat pada gambar 2.



- Read Excel
Operator ini digunakan untuk membaca data dari Microsoft Excel di RapidMiner. Pada operator ini user dapat memilih data dan menentukan atribut masing-masing data yang akan diuji.
- Naïve Bayes
Operator ini digunakan untuk mengaplikasikan algoritma Naïve Bayes pada data training yang diambil dari Split Data. Operator ini mencari pola pada data training untuk selanjutnya digunakan pada data testing.
- Apply Model
Operator ini digunakan untuk mengimplementasikan model yang sudah diolah pada data training. Data yang diimplementasikan pada model memiliki atribut yang sama dengan model yang dihasilkan.
- Performance
Performance Operator adalah operator evaluasi yang dapat digunakan untuk mengukur performa model. Operator ini akan menentukan jenis pemodelan secara otomatis dan menghitung performa dari model tersebut.

HASIL DAN PEMBAHASAN

Metode Naïve Bayes mencari pola pada data training untuk diterapkan pada data testing. Prediksi kelulusan mahasiswa ditentukan dengan atribut Status Kelulusan dengan nilai Cepat, Tepat, dan Terlambat. Probabilitas posterior data testing yang dihasilkan pada metode Naïve Bayes menggunakan RapidMiner dapat dilihat pada Gambar 3.

Gambar 3. Probabilitas posterior

SimpleDistribution

Distribution model for label attribute IPK Semester 6

Class Tepat Waktu (0.658)
11 distributions

Dari gam Class Cepat (0.079)
memiliki 11 distributions

ulus dibawah 4 tahun)

memiliki Class Terlambat (0.211)
memiliki 11 distributions

lulus diatas 4 tahun)

asiswa lulus 4 tahun)

memiliki probabilitas posterior: 0,079.

SIMPULAN

Data yang digunakan adalah data mahasiswa angkatan 2020 dengan 36 data sebagai data training dan data testing. Dari hasil pengujian, ditemukan bahwa probabilitas posterior untuk class Cepat adalah 0,079, probabilitas posterior untuk class Terlambat adalah 0,211, dan probabilitas posterior untuk class Tepat adalah 0,658. Ini menunjukkan bahwa jika data testing memiliki atribut yang mengarah pada kelas Cepat, kemungkinan besar mahasiswa tersebut akan lulus dalam waktu kurang dari 4 tahun. Jika data testing memiliki atribut yang mengarah pada kelas Terlambat, kemungkinan besar mahasiswa tersebut akan lulus setelah lebih dari 4 tahun. Sementara itu, jika data testing memiliki atribut yang mengarah pada kelas Tepat, kemungkinan besar mahasiswa tersebut akan lulus dalam waktu tepat 4 tahun dan dapat disimpulkan bahwa metode Naïve Bayes dengan menggunakan RapidMiner dapat digunakan sebagai alat yang efektif dalam melakukan prediksi kelulusan mahasiswa berdasarkan atribut yang relevan. Atibrutnya ntara lain adalah nama mahasiswa, jenis kelamin, kelas, tahun masuk, sks semester, dan IPK. Hasil penelitian yang diperoleh dapat dijadikan sebagai pendukung pengambilan keputusan dalam peningkatan sarana dan prasarana untuk meningkatkan efektifitas kegiatan belajar mengajar di Universitas Prabumulih.

DAFTAR PUSTAKA

- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: "Concepts and Techniques". San Francisco, CA, itd: Morgan Kaufmann.
- Salmu, S., & Solichin, A. (2017). "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu" Menggunakan Naïve Bayes : Studi Kasus UIN Syarif Hidayatullah Jakarta, (April), 701–709

- harma, T., Sharma, A., & Mansotra, P. V. (2016). "*Performance Analysis of Data Mining Classification Techniques on Public Health Care*" Data, 11381–11386.
- Artaye, K. (2015). "*Implementation of Naïve Bayes Classification Method to Predict Graduation Time of IBI Darmajaya Scholar*", (August), 284–290.