



INNOVATIVE: Journal Of Social Science Research

Volume 4 Nomor 6 Tahun 2024 Page 8417-8429

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

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

Analisa Klasifikasi Tingkat Kelulusan Mahasiswa Metode Algoritma *Naive Bayes* Menggunakan Rapidminer

Agung Firmansyah^{1✉}, Muhammad Fatchan², Anggi Muhammad Rifa'i³

Universitas Pelita Bangsa Bekasi

Email: Agungfirmansyah549@gmail.com^{1✉}

Abstrak

Perkembangan teknologi informasi menghasilkan data besar dari berbagai bidang, termasuk pendidikan. Data mining membantu menemukan pola dalam dataset agar lebih bermanfaat. Di perguruan tinggi, kelulusan tepat waktu menjadi indikator kinerja, tetapi banyak mahasiswa terlambat lulus karena faktor seperti keuangan, pekerjaan, dan kurangnya tanggung jawab. Penelitian ini menggunakan Naive Bayes Classifier untuk memprediksi kelulusan mahasiswa dan merekomendasikan langkah yang diperlukan. Analisis dengan perhitungan manual, Microsoft Excel, dan RapidMiner menunjukkan akurasi 93,33%, precision 100%, dan recall 66,67%, membuktikan metode ini efektif untuk klasifikasi data tingkat kelulusan.

Kata Kunci : *Data Mining, Kelulusan Mahasiswa, Naive Bayes, Klasifikasi*

Abstract

The advancement of information technology has generated large amounts of data across various fields, including education. Data mining helps uncover patterns within datasets to make them more useful. In higher education, on-time graduation serves as a key performance indicator, yet many students face delays due to factors such as finances, employment, and lack of responsibility. This study employs the Naive Bayes Classifier to predict student graduation outcomes and recommend necessary steps. Analysis using manual calculations, Microsoft Excel, and RapidMiner revealed an accuracy of 93.33%, precision of 100%, and recall of 66.67%, demonstrating the method's effectiveness in classifying graduation data.

Keywords: *Data Mining, Student Graduation, Naive Bayes, Classification*

PENDAHULUAN

Kemajuan teknologi informasi sudah semakin berkembang pesat disegala bidang kehidupan. Banyak sekali data yang dihasilkan oleh teknologi informasi, mulai dari bidang industri, ekonomi, ilmu dan teknologi serta berbagai bidang kehidupan lainnya. Setelah menyelesaikan penelitian setiap mahasiswa dari perguruan tinggi negeri maupun swasta baik yang menempuh jenjang Strata maupun diploma diwajibkan untuk menyelesaikan tugas akhir berupa penelitian yang disetujui sehingga pada akhirnya dapat melewati sejumlah proses dan langkah yang harus diambil seelum penelitian. (Diska dan Budayawan, 2023)

Dalam setiap proses yang dilalui mahasiswa pasti ada kesulitan salah satu kesulitan yang dihadapi adalah ketika memilih topik tugas akhir yang akan digunakan. Topik yang tepat akan memudahkan mahasiswa dalam menyelesaikan tugas akhir dan dengan topik yang tepat diharapkan dapat membantu mahasiswa mencapai hasil yang diinginkan dari tujuan tugas akhir. Oleh karna itu dibutuhkan referensi atau patokan yang dapat membantu merekomendasikan mahasiswa dalam menentukan topik tugas akhir.(Nuralia et al., 2023)

Pengolahan data yang tersembunyi dapat dimulai dengan melakukan apa yang disebut dengan data mining. Data mining adalah proses menemukan pola atau informasi yang menarik dalam data yang dipilih dengan menggunakan teknik atau metode tertentu. Data mining menentukan pola yang ditemukan dalam dataset sehingga data yang ada lebih berguna dalam kehidupan nyata. Perguruan tinggi dapat memberikan banyak informasi tentang data mahasiswa dan data jumlah gelar. Universitas Pelita Bangsa (UPB) merupakan perguruan tinggi baru hasil perubahan bentuk dari penggabungan antara Sekolah Tinggi Ilmu Ekonomi (STIE) dan Sekolah Tinggi Teknik (STT). Rektor Universitas Pelita Bangsa telah berkembang menjadi 4 fakultas: Fakultas Ekonomi dan Bisnis, Fakultas Teknik, Fakultas Pendidikan dan Ilmu Budaya, Fakultas Agama Islam dan akhir-akhir ini Fakultas Ekonomi dan Bisnis terdiri dari 5 program akademik.(Kartarina dan Sriwinarti, 2021)

Lima Program akademik yaitu Magister Manajemen, Sarjana Manajemen, Sarjana Bisnis Digital, Sarjana Kewirausahaan, dan Diploma Akuntansi. Salah satu indikator kinerja Perguruan Tinggi adalah presentasi kelulusan mahasiswa tepat waktu namun sayangnya masih banyaknya mahasiswa yang lulus tidak tepat waktu dikarnakan ada beberapa faktor seperti halnya dalam hal ekonomi, pekerjaan, waktu serta kurang bertanggung jawabnya mahasiswa dalam menjalankan pendidikannya yang menyebabkan nilai dalam mata kuliah tersebut kurang baik sehingga mahasiswa diharuskan mengulang mata kuliah yang menyebabkan keterlambatan dalam kelulusannya dan membuat indikator kinerja perguruan tinggi dalam hal presentasi kelulusan mahasiswa tidak tepat waktu. (Azizi, 2024)

METODE PENELITIAN

Penelitian ini bertujuan untuk menentukan tingkat kelulusan mahasiswa secara sistematis melalui langkah-langkah dan atribut yang mendukung pencapaian tujuan. Metode yang digunakan adalah Knowledge Discovery Database (KDD), yang merupakan proses penemuan pola dan pengetahuan dari data dalam skala besar. Metode ini menghasilkan informasi yang valid, baru, mudah dipahami, dan bermanfaat. Penelitian didukung oleh perangkat lunak.(Sadli et al., 2018)

Data yang digunakan terdiri dari data kualitatif, seperti wawancara dan observasi, serta data kuantitatif berupa angka yang dianalisis secara statistik. Sumber data meliputi data sekunder dari kuesioner (230 data dengan 10 atribut) dan data primer yang diperoleh dari internet, GitHub, serta studi pustaka. Atribut yang digunakan untuk memprediksi kelulusan mahasiswa mencakup nama, biaya, pekerjaan, IPK, masalah pribadi, masalah pembimbing, dan Keputusan. Pengolahan data dimulai dari pembersihan data (data cleaning) untuk menghilangkan nilai kosong, duplikasi, atau inkonsistensi. Selanjutnya, dilakukan seleksi data (data selection) untuk memilih data yang relevan, yang dirangkum dalam Tabel 3.5. Tahap akhir pengolahan data adalah transformasi data (data transformation), di mana data diubah ke format yang sesuai untuk .algoritma,(Damanik, 2022)

Pemodelan dilakukan menggunakan algoritma Naive Bayes, yang menghitung probabilitas berdasarkan teorema Bayes dengan asumsi independensi atribut. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian guna memastikan kemampuan generalisasi model. Pengujian dan validasi dilakukan menggunakan perangkat lunak RapidMiner untuk mengevaluasi hasil perhitungan serta mengukur efektivitas metode yang diterapkan. Proses ini memastikan bahwa metode dan prosedur yang digunakan menghasilkan hasil yang sesuai dengan prediksi penelitian.(Sahasti, 2022)

HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Data Uji

Penelitian ini menggunakan algoritma *Naive Bayes*, untuk mengidentifikasi tingkat kelulusan mahasiswa serta prediksi yang dapat digunakan dalam pengambilan keputusan. Sumber data sebagai objek pada penelitian ini adalah data historis yang diambil dari kuisisioner mahasiswa yang nanti akan dijadikan data training. Data yang digunakan telah melalui proses pembersihan dan transformasi. Pengujian menggunakan data mahasiswa dari angkatan 2017- 2019 dengan data mahasiswa sebanyak 230 *record*, Untuk data training

digunakan 163 data dan data testing digunakan 67 data. Proses pengujian data dibagi menjadi 2 yaitu data *training* dan *testing* menggunakan *naive bayes*. Data *training* digunakan untuk menentukan tabel probabilitas dan data *testing* untuk menguji probabilitas yang telah dibentuk. (Diska dan Budayawan, 2023)

2. Perhitungan Manual Algoritma *Naive Bayes*

a). Prior Probabilitas

Proses untuk melakukan perhitungan kelulusan mahasiswa tepat waktu menggunakan data yang telah dipilih. Langkah pertama mencari nilai prior probabilitas untuk semua kasus yang Tepat dan Terlambat. Jumlah Tepat sebanyak 94 dan jumlah Terlambat sebanyak 69 sehingga jumlah data yang dipakai adalah 163 data.

$$P(\text{tepat}) = 94/163 = 0,58 \quad P(\text{terlambat}) = 69/163 = 0,42$$

b). Menghitung Probabilitas Atribut

Probabilitas tiap atribut adalah membandingkan atribut dari data testing dengan atribut dari data training. Jumlah atribut dengan kelas tepat waktu yang ada pada data training dibagi dengan probabilitas tepat waktu, begitu juga probabilitas tidak tepat waktu.

a) Atribut Jenis Kelamin

Tabel 4. 1 Jenis Kelamin

Jenis Kelamin	Tepat Waktu	Tidak Tepat Waktu
Laki-laki	59/94	46/69
Perempuan	35/94	23/69

Berdasarkan tabel diatas probabilitas atribut jenis kelamin dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{Jenis Kelamin} = \text{Laki-laki} \mid \text{class} = \text{Tepat Waktu}) = 59/94 = 0,62$$

$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{class} = \text{Tepat Waktu}) = 35/94 = 0,37$$

$$P(\text{Jenis Kelamin} = \text{Laki-laki} \mid \text{class} = \text{Tidak Tepat Waktu}) = 46/69 = 0,67$$

$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{class} = \text{Tidak Tepat Waktu}) = 23/69 = 0,33$$

b) Atribut Biaya

Tabel 4. 2 Biaya

Biaya	Tepat Waktu	Tidak Tepat Waktu
Setuju	65/94	27/69
Sangat Setuju	27/94	38/69
Tidak Setuju	2/94	4/69
Sangat Tidak Setuju	0	0

Berdasarkan tabel diatas probabilitas atribut biaya dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{Biaya} = \text{Setuju} \mid \text{class} = \text{Tepat Waktu}) = 65/94 = 0,69$$

$$P(\text{Biaya} = \text{Sangat Setuju} \mid \text{class} = \text{Tepat Waktu}) = 27/94 = 0,29$$

$$P(\text{Biaya} = \text{Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 2/94 = 0,02$$

$$P(\text{Biaya} = \text{Sangat Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 0$$

$$P(\text{Biaya} = \text{Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 27/69 = 0,39$$

$$P(\text{Biaya} = \text{Sangat Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 38/69 = 0,55$$

$$P(\text{Biaya} = \text{Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 4/69 = 0,05$$

$$P(\text{Biaya} = \text{Sangat Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 0$$

c) Atribut Pekerjaan

Tabel 4. 3 Pekerjaan

Pekerjaan	Tepat Waktu	Tidak Tepat Waktu
Setuju	60/94	24/69
Sangat Setuju	29/94	37/69
Tidak Setuju	5/94	8/69
Sangat Tidak Setuju	0	0

Berdasarkan tabel diatas probabilitas atribut pekerjaan dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{Pekerjaan} = \text{Setuju} \mid \text{class} = \text{Tepat Waktu}) = 60/94 = 0,63$$

$$P(\text{Pekerjaan} = \text{Sangat Setuju} \mid \text{class} = \text{Tepat Waktu}) = 29/94 = 0,30$$

$$P(\text{Pekerjaan} = \text{Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 5/94 = 0,05$$

$$P(\text{Pekerjaan} = \text{Sangat Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 0$$

$$P(\text{Pekerjaan} = \text{Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 24/69 = 0,34$$

$$P(\text{Pekerjaan} = \text{Sangat Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 37/69 = 0,53$$

$$P(\text{Pekerjaan} = \text{Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 8/69 = 0,11$$

$$P(\text{Pekerjaan} = \text{Sangat Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 0$$

d) Atribut IPK

Tabel 4. 4 IPK

IPK	Tepat Waktu	Tidak Tepat Waktu
1	0	0
2	12/94	40/69
3	82/94	26/69

Berdasarkan tabel diatas probabilitas atribut IPK dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{IPK} = 1 \mid \text{class} = \text{Tepat Waktu}) = 0$$

$$P(\text{IPK} = 2 \mid \text{class} = \text{Tepat Waktu}) = 12/94 = 0,12$$

$$P(\text{IPK} = 3 \mid \text{class} = \text{Tepat Waktu}) = 82/94 = 0,87$$

$$P(\text{IPK} = 1 \mid \text{class} = \text{Tidak Tepat Waktu}) = 0$$

$$P(\text{IPK} = 2 \mid \text{class} = \text{Tidak Tepat Waktu}) = 40/69 = 0,57$$

$$P(\text{IPK} = 3 \mid \text{class} = \text{Tidak Tepat Waktu}) = 29/69 = 0,42$$

e) Atribut Masalah Pribadi

Tabel 4. 5 Masalah Pribadi

Pekerjaan	Tepat Waktu	Tidak Tepat Waktu
Setuju	32/94	48/69
Sangat Setuju	58/94	14/69
Tidak Setuju	3/94	6/69
Sangat Tidak Setuju	1/94	1/69

Berdasarkan tabel diatas probabilitas atribut Masalah Pribadi dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{Masalah Pribadi} = \text{Setuju} \mid \text{class} = \text{Tepat Waktu}) = 32/94 = 0,34$$

$$P(\text{Masalah Pribadi} = \text{Sanangat Setuju} \mid \text{class} = \text{Tepat Waktu}) = 58/94 = 0,61$$

$$P(\text{Masalah Pribadi} = \text{Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 3/94 = 0,03$$

$$P(\text{Masalah Pribadi} = \text{Sangat TidakSetuju} \mid \text{class} = \text{Tepat Waktu}) = 1/94 = 0,01$$

$$P(\text{Masalah Pribadi} = \text{Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 48/69 = 0,70$$

$$P(\text{Masalah Pribadi} = \text{Sanangat Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 14/69 = 0,20$$

$$P(\text{Masalah Pribadi} = \text{Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 6/69 = 0,07$$

$$P(\text{Masalah Pribadi} = \text{Sangat TidakSetuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 1/69 = 0,01$$

f) Atribut Masalah Pembimbing

Tabel 4. 6 Masalah Pembimbing

Masalah Pembimbing	Tepat Waktu	Tidak Tepat Waktu
Setuju	51/94	49/69
Sangat Setuju	36/94	14/69
Tidak Setuju	6/94	5/69
Sangat Tidak Setuju	1/94	1/69

Berdasarkan tabel diatas probabilitas atribut Masalah Pembimbing dengan kelas tepat waktu dan tidak tepat waktu menghasilkan data sebagai berikut:

$$P(\text{Masalah Pembimbing} = \text{Setuju} \mid \text{class} = \text{Tepat Waktu}) = 51/94 = 0,79$$

$$P(\text{Masalah Pembimbing} = \text{Sanangat Setuju} \mid \text{class} = \text{Tepat Waktu}) = 36/94 = 0,38$$

$$P(\text{Masalah Pembimbing} = \text{Tidak Setuju} \mid \text{class} = \text{Tepat Waktu}) = 6/94 = 0,06$$

$$P(\text{Masalah Pembimbing} = \text{Sangat TidakSetuju} \mid \text{class} = \text{Tepat Waktu}) = 1/94 = 0,01$$

$$P(\text{Masalah Pembimbing} = \text{Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 49/69 = 0,01$$

$P(\text{Masalah Pembimbing} = \text{Sanagat Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 14/69 = 0,20$

$P(\text{Masalah Pembimbing} = \text{Tidak Setuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 5/69 = 0,07$

$P(\text{Masalah Pembimbing} = \text{Sangat TidakSetuju} \mid \text{class} = \text{Tidak Tepat Waktu}) = 1/69 = 0,01$

3. Menghitung Probabilitas Akhir

Menghitung probabilitas akhir setiap kelas dapat menggunakan data traning dan mengubahnya menjadi nilai yang sudah ditentukan pada proses menghitung probabilitas atribut. Probabilitas kelas dikalikan dengan tiap atributnya, kedua hasil yang sudah ditentukan pada masing-masing kelas dibandingkan dengan nilai yang tertinggi. Berikut contoh data testing dalam pemahaman *naïve bayes* secara manual menggunakan *Rapid Miner*.

Tabel 4. 7 Data Testing

Jenis Kelamin	Biaya	Pekerjaan	IPK	Masalah Pribadi	Masalah Pembimbing	Keputusan
Perempuan	Setuju	Setuju	3,00	Setuju	Setuju	?
Laki-laki	Setuju	Tidak Setuju	3,00	Setuju	Setuju	?

Berdasarkan tabel diatas langkah selanjutnya adalah kalikan semua variabel tepat waktu dan tidak tepat waktu dan menghasilkan data sebagai berikut:

$$\begin{aligned} &P(\text{Perempuan} \mid \text{Tepat Waktu}) \times P(\text{Biaya} \mid \text{Tepat Waktu}) \times P(\text{Pekerjaan} \mid \text{Tepat Waktu}) \\ &\times P(\text{IPK } 3 \mid \text{Tepat Waktu}) \times P(\text{Masalah Pribadi} \mid \text{Tepat Waktu}) \times P(\text{Masalah} \\ &\text{Pembimbing} \mid \text{Tepat Waktu}) \\ &= 0,37 \times 0,69 \times 0,63 \times 0,87 \times 0,34 \times 0,54 \\ &= 0,026 \end{aligned}$$

$$\begin{aligned} &P(\text{Perempuan} \mid \text{Tidak Tepat Waktu}) \times P(\text{Biaya} \mid \text{Tidak Tepat Waktu}) \times P(\text{Pekerjaan} \mid \text{Tidak Tepat} \\ &\text{Waktu}) \times P(\text{IPK} \mid \text{Tidak Tepat Waktu}) \times P(\text{Masalah Pribadi} \mid \text{Tidak Tepat Waktu}) \times P(\text{Masalah} \\ &\text{Pembimbing} \mid \text{Tidak Tepat Waktu}) \\ &= 0,33 \times 0,39 \times 0,35 \times 0,42 \times 0,70 \times 0,71 \\ &= 0,009 \end{aligned}$$

Bahwa dalam perhitungan *naïve bayes* terlihat probabilitas jenis kelamin perempuan tertinggi ada pada $(P \mid \text{Tepat Waktu})$.

$$\begin{aligned} &P(\text{Laki-laki} \mid \text{Tepat Waktu}) \times P(\text{Biaya} \mid \text{Tepat Waktu}) \times P(\text{Pekerjaan} \mid \text{Tepat Waktu}) \times P(\text{IPK} \\ &3 \mid \text{Tepat Waktu}) \times P(\text{Masalah Pribadi} \mid \text{Tepat Waktu}) \times P(\text{Masalah Pembimbing} \mid \text{Tepat Waktu}) \\ &= 0,62 \times 0,69 \times 0,05 \times 0,87 \times 0,34 \times 0,54 \end{aligned}$$

= 0,003

$P(\text{Laki-laki}|\text{Tidak Tepat Waktu}) \times P(\text{Biaya}|\text{Tidak Tepat Waktu}) \times P(\text{Pekerjaan}|\text{Tidak Tepat Waktu}) \times P(\text{IPK}|\text{Tidak Tepat Waktu}) \times P(\text{Masalah Pribadi}|\text{Tidak Tepat Waktu}) \times P(\text{Masalah Pembimbing}|\text{Tidak Tepat Waktu})$

= 0,67 x 0,39 x 0,11 x 0,42 x 0,70 x 0,71

= 0,005

Bahwa dalam perhitungan *naïve bayes* terlihat probabilitas jenis kelamin Laki-laki tertinggi ada pada (P | Tidak Tepat Waktu). Selanjutnya dapat disimpulkan bahwa dalam perhitungan *naïve bayes* terlihat probabilitas tertinggi ada pada Jenis kelamin perempuan (P | Tepat Waktu) sehingga dapat disimpulkan mahasiswa dengan jenis kelamin perempuan memiliki estimasi tertinggi lulus tepat waktu.

B. Pembahasan

1. Proses Pengujian Data (Rapidminer)

Setelah dilakukannya perhitungan manual maka proses selanjutnya untuk mengklasifikasi tingkat kelulusan dalam pengujian ini menggunakan *tools RapidMiner*. *Dataset* yang wajib dilakukan penyesuaian pola data serta sesuai menggunakan tujuan asal pengolahan data. Dalam rangkaian proses data mining memakai *tools RapidMiner* data yang dipergunakan merupakan data yang sudah di transformasi kedalam format *Microsoft Excel*. Pengujian ini untuk mengetahui nilai akurasi dan prediksi data. Pengujian dilakukan menggunakan metode *algoritma Naive Bayes* dan data yang digunakan sebanyak 230 data, data training 163 dan data testing sebanyak 67 data.

2. Tahap Pengujian

Pertama, peneliti membuka alat rapidminer dan menggunakan file pembinaan data untuk mengimportnya. Selanjutnya, mereka menentukan fungsi dan atribut yang ingin dilabelkan.

1. *Import Data Training*

Sebuah model classifier dibuat dengan data training untuk memprediksi kelas data baru.

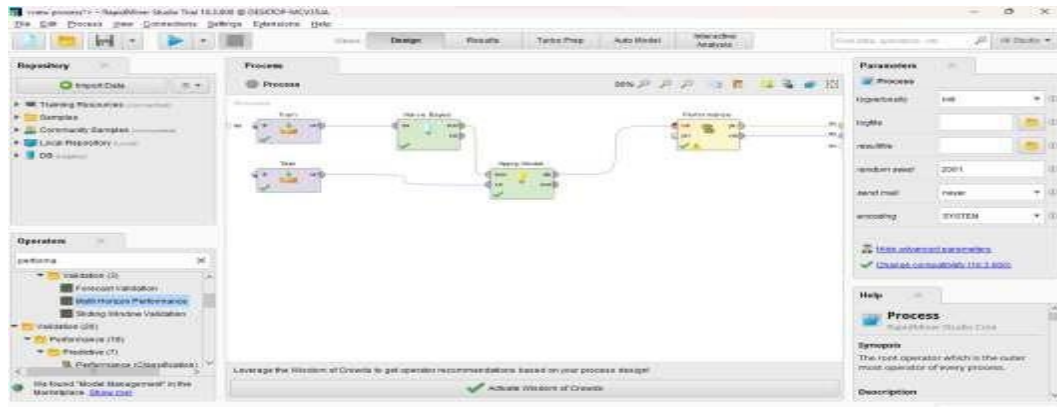
2. *Import Data Testing*

Data testing digunakan untuk mengetahui sejauh mana prediksi berhasil untuk melakukan prediksi kualitas dengan tepat.

3. Proses Pengujian

Proses pembagian terstruktur mengenai di *RapidMiner* memakai metode *Naive Bayes* untuk membandingkan data training serta *data testing* yang telah diketahui rulenya, di tampilan proses masukan operator *read excel input* masing masing data training serta data

testing, lalu masukan operator *Naive Bayes*, *Apply model* dan *Performance* serta sambungkan kabel kemudian klik *button run*.



Gambar 4. 1 Tahap pengujian

Pengujian pada tahap ini dilakukan implementasi prosedur pemecahan *Naive Bayes* menggunakan *tools RapidMiner*, yaitu menentukan *accuracy*, *precision* dan *recall*, serta memilih *performance*. acara yang telah dipergunakan untuk mengolah data yang sudah ada. Pada tahap ini pula dilakukan penilaian hasil pengolahan data menggunakan acara yang digunakan yaitu di *tools RapidMiner*, pada pemodelan ini bagian data *training* (digunakan untuk prosedur pemecahan klasifikasi) serta data *testing* (memakai fitur pada *Apply model* buat mengaplikasikan model pada data testing), serta *tools performance* ditujukan untuk menampilkan *confusion table*, yang dipergunakan untuk menampilkan hasil dari *accuracy*, *precision* dan *recall*. (Anjarsari et al., 2020)

Pembahasan ini dilakukan untuk mendapatkan nilai *accuracy* dan *precision*, masing-masing. *Accuracy* adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem, dan *precision* adalah tingkat kedekatan antara hasil prediksi dan hasil fakta. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali informasi.

PerformanceVector

```
PerformanceVector:
accuracy: 86.57%
ConfusionMatrix:
True:   Tepat   Tidak Tepat
Tepat:  37      1
Tidak Tepat:  8      21
precision: 97.37% (positive class: Tepat)
ConfusionMatrix:
True:   Tidak Tepat   Tepat
Tidak Tepat:  21      8
Tepat:  1      37
recall: 82.22% (positive class: Tepat)
ConfusionMatrix:
True:   Tidak Tepat   Tepat
Tidak Tepat:  21      8
Tepat:  1      37
```

Gambar 4. 2 Hasil *Performance Vector*

1. Accuracy

Dengan mengetahui data yang diklasifikasi secara benar maka dapat diketahui hasil

accuracy prediksinya yaitu 86.57%

Table View Plot View

accuracy: 86.57%

	true Tepat	true Tidak Tepat	class precision
pred. Tepat	37	1	97.37%
pred. Tidak Tepat	8	21	72.41%
class recall	82.22%	95.45%	

Gambar 4. 3 Hasil Accuracy

2. Precision

Hasil dari pengujian nilai *precision* yaitu 82.22% untuk kelas *true* (tepat) dan 95.45% untuk kelas *false* (tidak tepat).

Table View Plot View

precision: 97.37% (positive class: Tepat)

	true Tidak Tepat	true Tepat	class precision
pred. Tidak Tepat	21	8	72.41%
pred. Tepat	1	37	97.37%
class recall	95.45%	82.22%	

Gambar 4. 4 Hasil Precision

3. Recall

Dari hasil pengujian *recall* yaitu 95.45% pada kelas *true* (tepat) dan nilai kelas *false* (tidak tepat) yaitu 82.22%.

Table View Plot View

recall: 82.22% (positive class: Tepat)

	true Tidak Tepat	true Tepat	class precision
pred. Tidak Tepat	21	8	72.41%
pred. Tepat	1	37	97.37%
class recall	95.45%	82.22%	

Gambar 4. 5 Hasil Recall

Analisa Hasil Penelitian

Sesudah dilakukan analisa pada bab sebelumnya, penulis melanjutkan pembahasan tentang hasil dari pengujian yang telah dilakukan untuk mengetahui nilai *accuracy*, *precision* dan *recall* dan prediksi yang dapat digunakan untuk mengetahui apakah mahasiswa dapat lulus tepat waktu atau tidak, prosedur pemecahan ini memakai *algoritma Naive Bayes*.(Dwi Amalia dan Mustofa Lestyanto, 2021)

Berdasarkan hasil dari pengukuran *Performance* yang dihasilkan oleh *algoritma*

naive bayes menggunakan *rapid miner* maka dapat diketahui hasil *accuracy* prediksinya yaitu 86.57% hasil dari pengujian nilai *precision* yaitu 82.22% untuk kelas *true* (tepat) dan 95.45% untuk kelas *false* (tidak tepat) sedangkan hasil dari pengujian *recall* yaitu 95.45% pada kelas *true* (tepat) dan nilai kelas *false* (tidak tepat) yaitu 82.22% .(Jamal, 2014)

Tabel 4. 8 Analisa Hasil Penelitian

<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
86.57%	95.45%	82.22%.

Hasil dari data tingkat *kelulusan* menyatakan bahwa tingkat *accuracy* nya sebesar 86.57%, *precision* 95.45% sedangkan *recall* 82.22%.

SIMPULAN

Hasil pembahasan yang telah diuraikan oleh penulis tentang penerapan *Algoritma Naive Bayes* untuk menentukan klasifikasi analisa tingkat kelulusan, maka penulis dapat mengambil kesimpulan yaitu sebagai berikut:

- a. Berdasarkan hasil yang telah dilakukan oleh penulis dengan perhitungan menggunakan *Algoritma Naive Bayes* yang dihitung secara manual, *Microsoft Excel*, dan *tool Rapid Miner* menghasilkan data yang sama sehingga data mining *algoritma* metode *Naive Bayes* sangat tepat digunakan untuk menghitung klasifikasi data tingkat kelulusan karena menghasilkan data yang akurat.
- b. Berdasarkan hasil perhitungan *Metode Naive Bayes* dengan 163 data training menggunakan *Rapid Miner* diketahui bahwa nilai *accuracy*: 93.33%, *Precision*: 100%, sedangkan *Recall*: 66.67%.
- c. Semakin banyak data sampel yang digunakan maka semakin besar juga kebenaran prediksinya, sehingga algoritma yang dibuat dapat digunakan untuk memprediksi kelas yang belum diketahui yaitu memprediksi kelulusan mahasiswa dapat tepat waktu atau tidak.

DAFTAR PUSTAKA

- Anjarsari, E., Farisdianto, D.D., Asadullah, A.W., 2020. Pengembangan Media Audiovisual Powtoon Pada Pembelajaran Matematika Untuk Siswa Sekolah Dasar. *Jmpm J. Mat. Dan Pendidik. Mat.* 5, 40–50.
- Azizi, R., 2024. Comparative Analysis Of Machine Learning Algorithms In Predicting Smartphone Prices. *Tugas_Akhir Int. J. Sci. Technol. Manag.* 5, 1030–1035.
- Damanik, B.E., 2022. Pengaruh Minat Baca Dan Peran Dosen Pembimbing Terhadap Keberhasilan Penulisan Tugas Akhir. *J. Manaj. Pendidik.* 11.
- Diska, K.R., Budayawan, K., 2023. Sistem Informasi Prediksi Kelulusan Menggunakan Metode

- Naive Bayes Classifier (Studi Kasus: Prodi Pendidikan Teknik Informatika). *J. Pendidik. Tambusai* 7, 936–943.
- Dwi Amalia, A., Mustofa Lestyanto, L., 2021. Lks Berbasis Saintifik Berbantuan Live Worksheets Untuk Memahami Konsep Matematis Pada Aritmetika Sosial. *J. Cendekia J. Pendidik. Mat.* 05, 2911–2933.
- Jamal, F., 2014. Analisis Kesulitan Belajar Siswa Pada Materi Peluang | 18. *J. Maju (Jurnal Pendidik. Mat.* 1, 18–36.
- Kartarina, K., Sriwinarti, N.K., 2021. Analisis Metode K-Nearest Neighbors (K-Nn) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa. *Jtim J. Teknol. Inf. Dan Multimed.* 3, 106–112.
- Nuralia, S., Harliana, H., Prabowo, T., 2023. Implementasi Naive Bayes Classifier Dalam Memprediksi Kelulusan Mahasiswa. *J. Autom. Comput. Inf. Syst.* 3, 63–72.
- Sadli, M., Fajriana, F., Fuadi, W., Ermatita, E., Pahendra, I., 2018. Analisis Model Naive Bayes Untuk Identifikasi Penggolongan Daya Listrik Di Kota Lhoksumawe. *Komik (Konferensi Nas. Teknol. Inf. Dan Komputer)* 2.
- Sahasti, J.P., 2022. Analisis Kebutuhan Pengembangan Buku Saku Tata Penulisan Tugas Akhir Mahasiswa Fakultas Dakwah Dan Komunikasi. *Ed-Humanistics J. Ilmu Pendidik.* 7, 887–893.