

Implementasi Metode Naïve Bayes Dalam Memprediksi Kelulusan Mahasiswa Tepat Waktu Pada Program Studi Pendidikan Teknologi Informasi

Darman^{1*}, Zila Razilu¹

¹Program Studi Pendidikan Teknologi Informasi, Universitas Muahmmadiyah Kendari, Indonesia.

Artikel Info

Kata Kunci:

Implementasi;
Kelulusan;
Naïve Bayes;
Tepat Waktu.

Keywords:

Implementation;
Graduation;
Naïve Bayes;
On time.

Riwayat Artikel:

Submitted: 21 Februari 2024
Accepted: 29 Oktober 2024
Published: 29 Oktober 2024

Abstrak: Program Studi Pendidikan Teknologi Informasi merupakan salah satu program studi baru di fakultas keguruan dan ilmu Pendidikan, Data dari pangkalan Data mahasiswa aktif sampai tahun 2023 ini berjumlah 520 orang, dari jumlah mahasiswa tersebut, ada sebanyak 27 orang mahasiswa yang telah lulus studi. Jika melihat data kelulusan berdasarkan angkatannya maka setidaknya ada 21 data lulusan mahasiswa Angkatan 2018 dari total penerimaan Angkatan pertama 97 mahasiswa, selanjutnya 6 mahasiswa lulus dari 146 mahasiswa Angkatan 2019. Penelitian ini bertujuan untuk: 1) Memodelkan prediksi kelulusan mahasiswa tepat waktu dengan melibatkan factor-faktor yang mendukung kelulusan mahasiswa tepat waktu. 2) Melakukan analisis factor-faktor penentu kelulusan mahasiswa tepat waktu dengan metode naïve bayes, dan 3) untuk mengetahui apakah metode naïve bayes dapat digunakan dalam memprediksi kelulusan mahasiswa tepat waktu. Metode penelitian yang digunakan adalah kuantitatif. Hasil dari penelitian ini adalah: Memodelkan prediksi kelulusan mahasiswa tepat waktu dan tidak tepat waktu dengan melibatkan faktor-faktor yang mendukung kelulusan mahasiswa tepat waktu dan tidak tepat waktu, Metode naïve bayes dapat digunakan dalam memprediksi kelulusan mahasiswa tepat waktu, dengan probabilitas Kelulusan Tepat Waktu Sebesar 11% dan Tidak Tepat Waktu Sebesar 89%. Dengan Akurasi 100%, dengan hasil pelatihan dan pengujian tanpa error. Ini dibuktikan pada data pengujian model melalui RapidMiner.

Abstract: *The Information Technology Education Study Program is one of the new study programs at the teaching and education faculty. Data from the database of active students until 2023 totals 520 people, of that number of students, there are 27 students who have graduated. If you look at the graduation data based on class, there are at least 21 students graduating from the Class of 2018 from the total intake of the first class of 97 students, then 6 students graduated from the 146 students from the Class of 2019. This research aims to: 1) Model predictions of student graduation on time by involving factors- factors that support student graduation on time. 2) Analyzing the factors determining student graduation on time using the Naïve Bayes method, and 3) to find out whether the Naïve Bayes method can be used to predict student graduation on time. The research method used is quantitative. The results of this research are: Modeling predictions of student graduation on time and not on time by involving factors that support student graduation on time and not on time. The Naïve Bayes method can be used to predict student graduation on time, with a probability of Graduation on Time of 11% and not on time at 89%. With 100% accuracy, with error-free training and testing results. This is proven on model test data via RapidMiner.*

Corresponding Author:

Darman

Email: darman@umkendari.ac.id

PENDAHULUAN

Program Studi Pendidikan Teknologi Informasi merupakan salah satu program studi baru di fakultas keguruan dan ilmu Pendidikan dengan nomor izin penyelenggaraan 678/KPT/I/2017 tanggal 05 Desember 2017, dan mulai menerima mahasiswa baru tahun 2018 dengan jumlah mahasiswa sebanyak 97 orang. Data dari pangkalan Data mahasiswa aktif sampai tahun 2023 ini berjumlah 520 orang, dengan rincian sebagai berikut: Mahasiswa aktif 2018 sebanyak 97 orang, Angkatan 2019 sebanyak 146 orang, Angkatan 2020 sebanyak 83 orang, Angkatan 2021 sebanyak 73 orang dan Angkatan 2022 sebanyak 121 orang. Dari jumlah mahasiswa angkatan 2018-2019 tersebut, ada sebanyak 27 orang mahasiswa yang telah lulus studi. Jika melihat data kelulusan berdasarkan angkutannya maka setidaknya ada 21 data lulusan mahasiswa Angkatan 2018 dari total penerimaan Angkatan pertama 97 mahasiswa, selanjutnya 6 mahasiswa lulus dari 146 mahasiswa Angkatan 2019.

Data kelulusan yang telah diuraikan di atas masih tergolong sangat kecil jika dibandingkan dengan jumlah mahasiswa diangkutannya, terutama untuk mahasiswa angkatan 2018 dan 2019. Untuk mencapai kelulusan pada jenjang S1 di Universitas Muhammadiyah Kendari yang ditetapkan dalam Buku Pedoman Akademik pasal 56 ayat 2 point a) Telah meluluskan semua mata kuliah yang dipersyaratkan oleh kurikulum Program studi., b) Mencapai IPK. > 2,75 (dua koma tujuh lima), c) Telah memenuhi persyaratan yang ditetapkan, seperti memperoleh sertifikat baca tulis al-quran, bebas pemabayaran, bebas Pustaka, dan lain-lain, d) telah dilaporkan kelulusannya oleh universitas le LLDIKTI melalui rektor.

Persyaratan kelulusan yang ditetapkan seperti yang diuraikan pada paragraph dua, pada dasarnya merupakan persyaratan yang mudah untuk dicapai. Disamping itu persyaratan tersebut masih bersifat umum yang juga digunakan di perguruan tinggi lain, dengan kata lain pencapaian kelulusan mahasiswa tidak terlalu rumit untuk menempuhnya selama mahasiswa benar-benar mengikuti prosedur yang ditetapkan di setiap perguruan tinggi, termasuk pada program studi Pendidikan teknologi informasi Universitas Muhammadiyah Kendari. Fakta dilapangan menunjukkan ketidak selarasan data kelulusan yang berbanding jauh dengan mahasiswa baru yang masuk pada setiap tahunnya.

Ketidak selarasan itu dipandang sebagai masalah yang perlu dicarikan solusinya. Untuk mencari solusi dari permasalahan tersebut, kami mencoba memprediksi kelulusan tepat dan tidak tepat waktu. Hasil penelitian ini nantinya dapat dijadikan sebagai rujukan dalam membuat kebijakan untuk meningkatkan jumlah kelulusan yang wajar dimasa mendatang. Untuk memprediksi menggunakan metode klasifikasi naïve bayes. Naïve bayes adalah sebuah algoritma klasifikasi probabilitas yang menggunakan teorema bayes untuk memperkirakan kemungkinan suatu kelas berdasar suatu input (Yunus et al., 2023)(Alfa Saleh, 2015; Fitriani, 2019; Irmayani, 2021; Mustofa & Mahfudh, 2019). Metode klasifikasi naïve bayes ini sudah sering digunakan untuk aplikasi prediksi diantaranya penelitian yang dilakukan (Damanik et al., 2021) dengan judul Prediksi Tingkat kepuasan dalam pembelajaran daring menggunakan algoritma Naïve Bayes, Selain itu penelitian yang dilakukan oleh (Dwi Herlambang & Hadi Wijoyo, 2019) dengan judul Algoritma naïve bayes untuk klasifikasi sumber belajar berbasis teks pada mata pelajaran produktif di SMK Rumpun Teknologi informasi dan komunikasi, selain itu naïve bayes banyak juga digunakan dalam analisis sentiment dan podelan seperti penelitian yang dilakukan (Putu et al., 2021) dengan judul penelitian Analisis sentiment dan pemodelan topik pariwisata Lombok menggunakan algoritma naïve bayes dan latent Dirichlet allocation, Sama halnya yang dilakukan oleh (Felicia Watratan et al., 2020) penelitian yang berjudul Implementasi Algoritma naïve bayes untuk memprediksi tingkat penyebaran covid-19 di Indonesia, dan masih banyak lagi riset-riset lain yang menggunakan metode naïve bayes untuk prediksi. Selain itu naïve bayes juga banyak digunakan dalam penelitian lain diantaranya penelitian yang dilakukan oleh (Irmayani, 2021), dengan judul penelitian

Visualisasi data pada data mining menggunakan metode klasifikasi naïve bayes. Penelitian yang dilakukan oleh (Sihombing et al., 2021) dengan judul Analisis Keberhasilan pembelajaran Daring pada masa pandemic covid 19 menggunakan algoritma C4.5 dan Naïve Bayes.

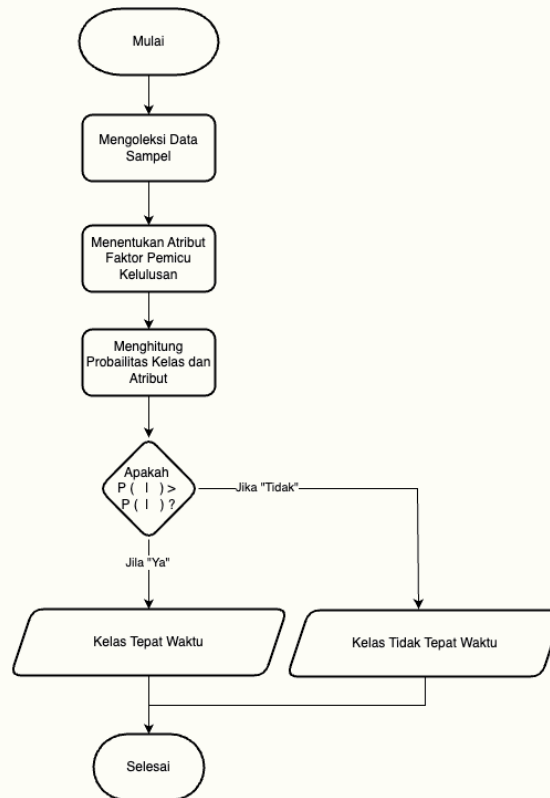
Penelitian ini sangat penting dilakukan untuk mengetahui guna memprediksi kelulusan mahasiswa tepat waktu dan tidak tepat waktu dengan melihat factor-faktor yang mendukung kelulusan mahasiswa tepat waktu dan tidak tepat waktu. Kebaruan dalam penelitian ini adalah dengan menambahkan beberapa indikator yang belum pernah dilakukan penelitian sebelumnya dimana indikator tersebut adalah Jalur masuk, Keaktifan dalam Organisasi, Sumber biaya, dan Waktu Pengusulan judul Skripsi.

Penelitian ini bertujuan untuk: 1) Memodelkan prediksi kelulusan mahasiswa tepat waktu dengan. Melibatkan factor-faktor yang mendukung kelulusan mahasiswa tepat waktu. 2) Melakukan analisis factor-faktor penentu kelulusan mahasiswa tepat waktu dengan metode naïve bayes, dan 3) untuk mengetahui apakah metode naïve bayes dapat digunakan dalam memprediksi kelulusan mahasiswa tepat waktu.

METODE

Penelitian ini berupa prediksi kelulusan mahasiswa tepat waktu, dimana metode penelitiannya adalah kuantitatif yang digunakan sebagai pendekatan penelitian yang berlandaskan pada filsafat positifisme digunakan untuk meneliti pada populasi atau sampel tertentu, pengumpulan data menggunakan instrument penelitian, analisis data bersifat kuantitatif dengan tujuan untuk menguji hipotesis yang ditetapkan. Pendekatan kuantitatif pada penelitian ini mengacu pada metode klasifikasi naïve bayes yang digunakan untuk menganalisis data untuk menghasilkan keluaran berupa klasifikasi.

Populasi dalam penelitian ini adalah data mahasiswa Angkatan 2018-2019 sedangkan yang menjadi sampel dalam penelitian ini adalah mahasiswa lulus Angkatan 2018-2019. Variabel penelitian dalam penelitian ini terdiri variable dependen yaitu kelulusan mahasiswa yakni kelulusan mahasiswa tepat waktu, dan terlambat (tidak tepat waktu). Dimana indikator kelulusan tepat dan terlambat waktu dilihat dari lama studi mahasiswa. dan variabel independent adalah berupa Data mahasiswa berupa jenis kelamin (JK), indeks prestasi yang diperoleh dari semester 1 hingga 5 (IPS1-IPS5), Jumlah Kredit Kumulatif yang ditempuh setiap mahasiswa pada semester 5 (JKK), lama masa tugas akhir (LTA), Jalur masuk (Jalur), Aktif organisasi kampus (ORG), Biaya, dan Waktu Pengusulan Tugas Akhir (WPTA). Pengumpul data dalam penelitian ini mendapatkan data dari pihak pengumpul data pertama yaitu data sampel didapat dari database program studi Pendidikan Teknologi Informasi Universitas Muhammadiyah Kendari. Pengolahan data pada penelitian ini menggunakan metode klasifikasi naïve bayes. Analisis data bertujuan untuk mengklasifikasikan mahasiswa yang termasuk kedalam kelompok kelas dengan kelulusan tepat waktu dan kelompok kelas terlambat (tidak tepat waktu). Analisis metode klasifikasi naïve bayes berupa pemodelan matematis yang menghasilkan model algoritma untuk memprediksi mahasiswa yang memungkinkan dapat lulus tepat waktu dan mahasiswa yang memungkinkan tidak tepat waktu (terlambat) (Mustofa & Mahfudh, 2019). Model algoritma naïve bayes dapat diterapkan pula dalam bentuk aplikasi computer untuk memprediksi kelulusan mahasiswa dimasa mendatang. Adapun tahapan model klasifikasi naïve bayes adalah sebagai berikut:



Gambar 1. Bagan Cara Kerja Naïve Bayes (Armansyah & Kurniawan, 2021)

HASIL DAN PEMBAHASAN

Hasil Penelitian

1. Analisis Kebutuhan

Analisis kebutuhan dilakukan dengan menganalisis kebutuhan data pelatihan (training) dan data pengujian (testing). Baik data pelatihan maupun data pengujian adalah 27 mahasiswa program studi Pendidikan Teknologi Informasi Universitas Muhammadiyah Kendari Angkatan tahun 2018 dan 2019, yang lulus pada 2 waktu yaitu pada semester ganjil 2022/2023 dan semester genap 2022/2023. Adapun data pelatihan diambil 18 mahasiswa dan data pengujian diambil 9 mahasiswa. Tujuan penelitian ini adalah memprediksi kelulusan mahasiswa tepat waktu atau terlambat (tidak tepat waktu), yang dimaksudkan sebagai variable dependen. Sedangkan data-data yang bertindak sebagai variable independent dapat dilihat pada table 1, yang merupakan data awal yang belum melalui proses tahapan apapun.

Tabel 1. Data Set Kelulusan

NO	JK	IPS1	IPS2	IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA	KL
1	L	3,05	4,00	3,76	3,90	3,62	121	2	4,3	UMUM	AKTIF	M	VII	TL
2	P	3,38	3,52	3,57	3,35	2,90	120	3	4,9	UMUM	TIDAK	M	VII	TL
3	P	2,95	3,17	3,19	3,30	2,95	121	3	4,9	UMUM	TIDAK	M	VII	TL
4	L	2,81	3,56	3,71	3,40	3,63	124	2	4,3	UMUM	TIDAK	M	VII	TL
5	P	3,24	3,24	3,19	3,25	3,28	118	3	4,9	UMUM	TIDAK	M	VII	TL
6	P	3,48	3,76	3,38	3,30	3,17	118	2	4,3	UMUM	TIDAK	M	VII	TL
7	P	3,76	3,70	3,57	3,50	3,91	124	1	3,8	UMUM	AKTIF	B	VI	TP
8	P	3,24	3,90	3,62	3,40	3,56	118	2	4,3	UMUM	TIDAK	M	VII	TL
9	p	2,86	3,72	3,48	3,57	3,83	118	2	4,3	UMUM	TIDAK	M	VII	TL
10	P	2,52	3,28	3,24	3,65	3,56	118	3	4,9	UMUM	TIDAK	M	VII	TL
11	P	2,67	3,39	3,33	3,65	3,00	118	3	4,9	UMUM	TIDAK	M	VII	TL
12	L	2,81	3,28	3,24	3,55	3,56	118	3	4,3	UMUM	TIDAK	M	VII	TL
13	P	2,95	3,56	3,33	3,50	3,38	124	2	4,3	UMUM	TIDAK	M	VII	TL
14	P	3,05	3,14	3,05	2,35	3,00	117	3	4,8	UMUM	TIDAK	M	VII	TL
15	L	2,81	3,39	3,62	3,43	3,31	119	2	4,3	UMUM	TIDAK	M	VII	TL
16	L	2,90	3,53	3,48	3,20	1,44	113	3	4,3	UMUM	TIDAK	M	VII	TL
17	P	3,10	3,48	3,81	3,40	3,44	118	3	4,3	UMUM	TIDAK	M	VII	TL
18	L	3,67	3,88	3,65	3,52	3,64	124	1	3,9	UMUM	TIDAK	M	VI	TP
19	P	3,19	3,48	3,76	3,70	3,72	118	3	4,3	UMUM	TIDAK	M	VII	??
20	P	2,81	3,39	3,38	3,55	3,71	121	2	4,6	UMUM	AKTIF	M	VII	??
21	P	2,71	3,11	3,28	3,45	3,00	118	2	4,6	UMUM	AKTIF	M	VII	??
22	P	3,00	3,24	3,38	3,45	3,44	118	3	4,9	UMUM	TIDAK	M	VII	??
23	P	3,00	3,38	3,29	3,55	3,62	121	3	4,7	UMUM	AKTIF	M	VII	??
24	P	3,62	3,79	3,13	3,55	3,80	121	1	3,9	UMUM	TIDAK	M	VI	??
25	L	3,19	3,29	3,33	3,55	3,82	124	1	3,8	UMUM	TIDAK	B	VI	??
26	P	3,52	3,04	3,24	3,05	3,71	123	1	3,8	UMUM	TIDAK	M	VI	??
27	L	3,43	3,57	3,35	3,65	3,82	124	1	3,9	UMUM	AKTIF	B	VI	??

Untuk memperjelas data pada table di atas, berikut ini disampaikan uraian penjelasannya sebagai berikut:

- 1) JK, merupakan Jenis Kelamin dari setiap mahasiswa yang digunakan dalam data pelatihan dan data pengujian.
- 2) IPS, singkatan dari Indeks Prestasi Semester, yang dalam data ini diperoleh dari semester 1 sampai semester 5, dan diwakili dengan singkatan IPS1, IPS2, IPS3, IPS4 dan IPS5.
- 3) JKK, Jumlah Kredit Kumulatif yang telah ditempuh oleh mahasiswa pada semester 5.
- 4) LTA, mewakili informasi lamanya masa tugas akhir yang ditempuh mahasiswa, yang dihitung dalam satuan semester
- 5) STUDI, masa studi yang dilalui mahasiswa sampai memperoleh gelar sarjana Pendidikan (S. Pd) yang dihitung dalam satuan tahun.
- 6) JALUR, merupakan atribut yang menerangkan mahasiswa masuk kuliah melalui seleksi mandiri atau jalur SSO.
- 7) ORG, Atribut yang menerangkan tentang keaktifan mahasiswa dalam organisasi kampus.
- 8) Biaya, menerangkan atribut yang merangkan pembiayaan dapat berupa mandiri atau beasiswa.
- 9) WPTA, merupakan atribut yang menerangkan waktu pengusulan judul tugas akhir dilihat pada semester VI dan VII.
- 10) KL, merupakan singkatan dari kelulusan yang merupakan kelas yang akan dicari dari model prediksi. Dimana kelulusan tepat waktu akan diwakili dengan kelas TP, dan kelulusan terlambat (tidak tepat waktu) diwakili dengan kelas TL.

Disamping data-data tersebut, terdapat data-data berupa nama, nim, IPK lulus, Tempat tanggal lahir, data-data tersebut tidak termasuk pada factor pemicu kelulusan mahasiswa tepat waktu, Dengan demikian data-data tersebut tidak disajikan dalam table data sampel. Selanjutnya data lulusan di atas akan dikelompokkan dalam dua bagian menurut kegunaannya sebagai data pelatihan dan data pengujian yang disajikan pada table 2. dan 3.

Tabel 2. Data Pelatihan

NO	JK	IPS1	IPS2	IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA	KL
1	L	3,05	4,00	3,76	3,90	3,62	121	2	4,3	UMUM	AKTIF	M	VII	TL
2	P	3,38	3,52	3,57	3,35	2,90	120	3	4,9	UMUM	TIDAK	M	VII	TL
3	P	2,95	3,17	3,19	3,30	2,95	121	3	4,9	UMUM	TIDAK	M	VII	TL
4	L	2,81	3,56	3,71	3,40	3,63	124	2	4,3	UMUM	TIDAK	M	VII	TL
5	P	3,24	3,24	3,19	3,25	3,28	118	3	4,9	UMUM	TIDAK	M	VII	TL
6	P	3,48	3,76	3,38	3,30	3,17	118	2	4,3	UMUM	TIDAK	M	VII	TL
7	P	3,76	3,70	3,57	3,50	3,91	124	1	3,8	UMUM	AKTIF	B	VI	TP
8	P	3,24	3,90	3,62	3,40	3,56	118	2	4,3	UMUM	TIDAK	M	VII	TL
9	p	2,86	3,72	3,48	3,57	3,83	118	2	4,3	UMUM	TIDAK	M	VII	TL
10	P	2,52	3,28	3,24	3,65	3,56	118	3	4,9	UMUM	TIDAK	M	VII	TL
11	P	2,67	3,39	3,33	3,65	3,00	118	3	4,9	UMUM	TIDAK	M	VII	TL
12	L	2,81	3,28	3,24	3,55	3,56	118	3	4,3	UMUM	TIDAK	M	VII	TL
13	P	2,95	3,56	3,33	3,50	3,38	124	2	4,3	UMUM	TIDAK	M	VII	TL
14	P	3,05	3,14	3,05	2,35	3,00	117	3	4,8	UMUM	TIDAK	M	VII	TL
15	L	2,81	3,39	3,62	3,43	3,31	119	2	4,3	UMUM	TIDAK	M	VII	TL
16	L	2,90	3,53	3,48	3,20	1,44	113	3	4,3	UMUM	TIDAK	M	VII	TL
17	P	3,10	3,48	3,81	3,40	3,44	118	3	4,3	UMUM	TIDAK	M	VII	TL
18	L	3,67	3,88	3,65	3,52	3,64	124	1	3,9	UMUM	TIDAK	M	VI	TP

Tabel 3. Data Pengujian

NO	JK	IPS1	IPS2	IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA	KL
19	P	2,71	3,11	3,28	3,45	3,00	118	2	4,6	UMUM	AKTIF	M	VII	??
20	P	3,00	3,24	3,38	3,45	3,44	118	3	4,9	UMUM	TIDAK	M	VII	??
21	P	3,00	3,38	3,29	3,55	3,62	121	3	4,7	UMUM	AKTIF	M	VII	??
17	P	3,19	3,48	3,76	3,70	3,72	118	3	4,3	UMUM	TIDAK	M	VII	??
23	P	3,62	3,79	3,13	3,55	3,80	121	1	3,9	UMUM	TIDAK	M	VI	??
24	L	3,19	3,29	3,33	3,55	3,82	124	1	3,8	UMUM	TIDAK	B	VI	??
25	P	3,52	3,04	3,24	3,05	3,71	123	1	3,8	UMUM	TIDAK	M	VI	??
18	P	2,81	3,39	3,38	3,55	3,71	121	2	4,6	UMUM	AKTIF	M	VII	??
27	L	3,43	3,57	3,35	3,65	3,82	124	1	3,9	UMUM	AKTIF	B	VI	??

2. Transformasi Data

Untuk memudahkan proses analisis data pelatihan dan pengujian perlu ditrasformasikan kedalam bentuk yang sesuai dengan metode analisisnya. Adapun trasformasi data pelatihan dan pengujian yang dimaksud dapat dilihat pada table 4. dan Tabel 5

Tabel 4. Trasformasi Data Pelatihan

JK	IPS1	IPS3	IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA	KL
L	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	AKTIF	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5<3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5<3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
L	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA<=1	STUDI<4	UMUM	AKTIF	B	VI	TP
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
p	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
L	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4<3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
L	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
L	IPS1<3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5<3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
L	IPS1>=3,00	IPS3>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	JKK>=100	LTA<=1	STUDI<4	UMUM	TIDAK	M	VI	TP

Tabel 5. Trasformasi Data Pengujian

JK	IPS1	IPS2	IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA	KL
P	IPS1<3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA>1	STUDI>4	UMUM	AKTIF	M	VII	TL
P	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA>1	STUDI>4	UMUM	AKTIF	M	VII	TL
L	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA>1	STUDI>4	UMUM	TIDAK	M	VII	TL
P	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA<=1	STUDI<4	UMUM	TIDAK	M	VI	TP
L	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA<=1	STUDI<4	UMUM	TIDAK	B	VI	TP
P	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA<=1	STUDI<4	UMUM	TIDAK	M	VI	TP
P	IPS1<3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA>1	STUDI>4	UMUM	AKTIF	M	VII	TL
L	IPS1>=3,00	IPS2>=3,00	IPS3>=3,00	IPS4>=3,00	IPS5>=3,00	KK>=100	TA<=1	STUDI<4	UMUM	AKTIF	B	VI	TP

3. Penyajian Data Analisis

Sebelum analisis data dilakukan, data yang telah dikoleksi akan dipetakan berdasarkan kriteria status kelulusan kedalam bentuk yang mudah dibaca. Data tersebut adalah Jenis Kelamin (JK), Indek Prestasi mulai semester 1 sampai 5 (IPS1-IPS5), Jumlah Kredit Kumulatif (JKK), Lama Tugas Akhir (LTA), Masa Studi (STUDI), Jalur masuk (JALUR), Aktif organisasi (ORG), Biaya dan Waktu Pengajuan judul tugas akhir (WPTA). Adapun pemetaan datanya disajikan seperti pada table-table berikut ini.

Tabel 6. Kelulusan berdasarkan Jenis Kelamin

No.	JK	KL		
		TP	TL	JLH L/P
1	L	1	5	6
2	P	1	11	12
JUMLAH		2	16	18

Tabel 7. Kelulusan IP Semester 1

IPS	MAHASISWA	TP	TL
$\geq 3,00$	9	2	7
$< 3,00$	9	0	9
JLH	18	2	16

Tabel 8. Kelulusan IP Semester 2

IPS	MAHASISWA	TP	TL
$\geq 3,00$	18	2	16
$< 3,00$	0	0	0
JLH	18	2	16

4. Analisis dan Pemodelan Naïve Bayes

Pada tahap ini dilakukan analisis data dengan membuat model naïve bayes dalam melakukan prediksi kelulusan mahasiswa tepat waktu dan terlambat uantuk dapat diterapkan pada mahasiswa dimasa mendatang. Tahapan-tahapan analisis dengan metode naïve bayes diuraikan sebagai berikut:

a) Menghitung Probabilitas Kelas

Untuk melihat kelayakan suatu metode naïve bayes, probabalitas kelayakan setiap kelasnya harus ditentukan terlebih dahulu. Dalam penelitian ini kelas yang dicari adalah kelas kelulusan tepat waktu dan kelas kelulusan tidak tepat waktu atau terlambat. Sebelum menerapkan teori naïve bayes dicari terlebih dahulu probabalitas dari kedua kelas tersebut dengan persamaan (1)

Jika melihat data kelulusan pada table kelulusan, maka ada 2 mahasiswa yang keluar atau lulus tepat waktu dari 18 mahasiswa keluar lulus. Kelulusan disini diwakili dengan variable Y, maka probabilitas kelas diperoleh sebagai berikut:

$$P(Y="Tepat Waktu") = 2/18 = 0,11$$

$$P(Y="Tidak Tepat Waktu") = 16/18 = 0,88$$

b) Menghitung Probabilitas Kejadian

Kejadian yang dimaksud adalah fakta dari factor-faktor yang mempengaruhi kelulusan tepat dan tidak tepat waktu, dengan menggunakan persamaan yang sama dengan persamaan untuk probabilitas kelas yang dicari. Kejadian disini diwakili dengan variable X, sehingga dapat dilihat sebagai berikut:

1) Probabilitas Kejadian Atribut X untuk Kelas Tepat Waktu

Untuk X1 adalah Jenis Kelamin, probabilitasnya adalah

$$P(X="L" | Y = "Tepat Waktu") = 1/2 = 0,5$$

$$P(X="P" | Y = "Tepat Waktu") = 1/2 = 0,5$$

Untuk X2 adalah IP Semester 1, Probabilitasnya adalah:

$$P(X="IPS1 \geq 3" | Y = "Tepat Waktu") = 2/2 = 1$$

$$P(X="IPS1 < 3" | Y = "Tepat Waktu") = 0/2 = 0$$

Untuk X3 adalah IP Semester 2, Probabilitasnya adalah:

$$P(X="IPS2 \geq 3" | Y = "Tepat Waktu") = 2/2 = 1$$

$$P(X="IPS2 < 3" | Y = "Tepat Waktu") = 0/2 = 0$$

Untuk X4 adalah IP Semester 3, Probabilitasnya adalah:

$$P(X="IPS3 \geq 3" | Y = "Tepat Waktu") = 2/2 = 1$$

$$P(X=\text{"IPS3}<3" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X5 adalah IP Semester 4, Probabilitasnya adalah:

$$P(X=\text{"IPS4}\geq 3" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"IPS4}<3" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X6 adalah IP Semester 5, Probabilitasnya adalah:

$$P(X=\text{"IPS5}\geq 3" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"IPS5}<3" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X7 adalah Jumlah Kredit Kumulatif Semester 5, Probabilitasnya adalah:

$$P(X=\text{"JKK5}\geq 100" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"JKK5}<100" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X8 adalah lama tugas Akhir, Probabilitasnya adalah:

$$P(X=\text{"LTA}\leq 1" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"LTA}>1" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X9 adalah Masa Studi, Probabilitasnya adalah:

$$P(X=\text{"MSTUDI}\leq 4" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"MSTUDI}>4" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X10 adalah Jalur Masuk, Probabilitasnya adalah:

$$P(X=\text{"JALUR=UMUM" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"JALUR=SSO" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

Untuk X11 adalah aktif organisasi kampus, Probabilitasnya adalah:

$$P(X=\text{"ORG=Aktif" \mid Y = \text{"Tepat Waktu"}) = 1/2 = 0,5$$

$$P(X=\text{"ORG= Tidak" \mid Y = \text{"Tepat Waktu"}) = 1/2 = 0,5$$

Untuk X12 adalah pembiayaan, Probabilitasnya adalah:

$$P(X=\text{"BIAYA=Beasiswa" \mid Y = \text{"Tepat Waktu"}) = 1/2 = 0,5$$

$$P(X=\text{"BIAYA= Mandiri" \mid Y = \text{"Tepat Waktu"}) = 1/2 = 0,5$$

Untuk X13 adalah waktu pengusulan judul TA, Probabilitasnya adalah:

$$P(X=\text{"WP TA=VI" \mid Y = \text{"Tepat Waktu"}) = 2/2 = 1$$

$$P(X=\text{"WP TA= VII" \mid Y = \text{"Tepat Waktu"}) = 0/2 = 0$$

2) Probabilitas Kejadian Atribut X untuk Kelas Tidak Tepat Waktu

Untuk X1 adalah Jenis Kelamin, probabilitasnya adalah

$$P(X=\text{"L" \mid Y = \text{"Tidak Tepat Waktu"}) = 5/16 = 0,3125$$

$$P(X=\text{"P" \mid Y = \text{"Tidak Tepat Waktu"}) = 11/16 = 0,6875$$

Untuk X2 adalah IP Semester 1, Probabilitasnya adalah:

$$P(X=\text{"IPS1}\geq 3" \mid Y = \text{"Tidak Tepat Waktu"}) = 7/16 = 0,4375$$

$$P(X=\text{"IPS1}<3" \mid Y = \text{"Tidak Tepat Waktu"}) = 9/16 = 0,5625$$

Untuk X3 adalah IP Semester 2, Probabilitasnya adalah:

$$P(X=\text{"IPS2}\geq 3" \mid Y = \text{"Tidak Tepat Waktu"}) = 16/16 = 1$$

$$P(X=\text{"IPS2}<3" \mid Y = \text{"Tidak Tepat Waktu"}) = 0/16 = 0$$

Untuk X4 adalah IP Semester 3, Probabilitasnya adalah:

$$P(X=\text{"IPS3}\geq 3" \mid Y = \text{"Tidak Tepat Waktu"}) = 16/16 = 1$$

$$P(X=\text{"IPS3}<3" \mid Y = \text{"Tidak Tepat Waktu"}) = 0/16 = 0$$

Untuk X5 adalah IP Semester 4, Probabilitasnya adalah:

$$P(X=\text{"IPS4}\geq 3" \mid Y = \text{"Tidak Tepat Waktu"}) = 15/16 = 0,9375$$

$$P(X=\text{"IPS4}<3" \mid Y = \text{"Tidak Tepat Waktu"}) = 1/16 = 0,0625$$

Untuk X6 adalah IP Semester 5, Probabilitasnya adalah:

$$P(X=\text{"IPS5}\geq 3" \mid Y = \text{"Tidak Tepat Waktu"}) = 13/16 = 0,8125$$

$$P(X=\text{"IPS5}<3" \mid Y = \text{"Tidak Tepat Waktu"}) = 3/16 = 0,1875$$

Untuk X7 adalah Jumlah Kredit Kumulatif Semester 5, Probabilitasnya adalah:

$$P(X=\text{"JKK5}\geq 100" \mid Y = \text{"Tidak Tepat Waktu"}) = 16/16 = 1$$

$$P(X=\text{"JKK5}<100" \mid Y = \text{"Tidak Tepat Waktu"}) = 0/16 = 0$$

Untuk X8 adalah lama tugas Akhir, Probabilitasnya adalah:

$$P(X="LTA \leq 1" \mid Y = "Tidak Tepat Waktu") = 0/16 = 0$$

$$P(X="LTA > 1" \mid Y = "Tidak Tepat Waktu") = 16/16 = 1$$

Untuk X9 adalah Masa Studi, Probabilitasnya adalah:

$$P(X="MSTUDI \leq 4" \mid Y = "Tidak Tepat Waktu") = 0/16 = 0$$

$$P(X="MSTUDI > 4" \mid Y = "Tidak Tepat Waktu") = 16/16 = 1$$

Untuk X10 adalah Jalur Masuk, Probabilitasnya adalah:

$$P(X="JALUR=SSO" \mid Y = "Tidak Tepat Waktu") = 0/2 = 0$$

$$P(X="JALUR=UMUM" \mid Y = "Tidak Tepat Waktu") = 16/16 = 1$$

Untuk X11 adalah aktif organisasi kampus, Probabilitasnya adalah:

$$P(X="ORG=Aktif" \mid Y = "Tidak Tepat Waktu") = 1/16 = 0,0625$$

$$P(X="ORG= Tidak" \mid Y = "Tidak Tepat Waktu") = 15/16 = 0,9375$$

Untuk X12 adalah pembiayaan, Probabilitasnya adalah:

$$P(X="BIAYA=Beasiswa" \mid Y = "Tidak Tepat Waktu") = 0/16 = 0$$

$$P(X="BIAYA= Mandiri" \mid Y = "Tidak Tepat Waktu") = 16/16 = 1$$

Untuk X13 adalah waktu pengusulan judul TA, Probabilitasnya adalah:

$$P(X="WP TA=VI" \mid Y = "Tidak Tepat Waktu") = 0/16 = 0$$

$$P(X="WP TA= VII" \mid Y = "Tidak Tepat Waktu") = 16/16 = 1$$

3) Menghitung probabilitas label "Tepat Waktu"

Setelah menghitung probabilitas semua kejadian yang memiliki keluaran Tepat Waktu dari data pelatihan diatas, maka selanjutnya akan ditentukan probabilitas dari label "Tepat Waktu" sebagai berikut:

$$P(Y="Tepat Waktu") = P(X1(L) \mid ("Tepat Waktu")) * P(X2(>=3) \mid Y("Tepat Waktu")) * P(X3(>=3) \mid Y("Tepat Waktu")) * P(X4(>=3) \mid Y("Tepat Waktu")) * P(X5(>=3) \mid Y("Tepat Waktu")) * P(X6(>=3) \mid Y("Tepat Waktu")) * P(X7(>=100) \mid Y("Tepat Waktu")) * P(X8(<=1) \mid Y("Tepat Waktu")) * P(X9(<=4) \mid Y("Tepat Waktu")) * P(X10(UMUM) \mid Y("Tepat Waktu")) * P(X11(Aktif) \mid Y("Tepat Waktu")) * P(X12(Mandiri) \mid Y("Tepat Waktu")) * P(X13(VI) \mid Y("Tepat Waktu")) * P(Y| "Tepat Waktu")$$

$$P(Y="Tidak Tepat Waktu") = 0,11 * (0,5 * 1 * 1 * 1 * 1 * 1 * 1 * 1 * 0,5 * 0,5 * 1) = 0,01375$$

4) Menghitung probabilitas label "Tidak Tepat Waktu"

$$P(Y="Tidak Tepat Waktu") = P(X1(L) \mid ("Tidak Tepat Waktu")) * P(X2(>=3) \mid Y("Tidak Tepat Waktu")) * P(X3(>=3) \mid Y("Tidak Tepat Waktu")) * P(X4(>=3) \mid Y("Tidak Tepat Waktu")) * P(X5(>=3) \mid Y("Tidak Tepat Waktu")) * P(X6(>=3) \mid Y("Tidak Tepat Waktu")) * P(X7(>=100) \mid Y("Tidak Tepat Waktu")) * P(X8(>1) \mid Y("Tidak Tepat Waktu")) * P(X9(>) \mid Y("Tidak Tepat Waktu")) * P(X10(UMUM) \mid Y("Tidak Tepat Waktu")) * P(X11(Aktif) \mid Y("Tidak Tepat Waktu")) * P(X12(Mandiri) \mid Y("Tidak Tepat Waktu")) * P(X13(VII) \mid Y("Tidak Tepat Waktu")) * P(Y| "Tidak Tepat Waktu")$$

$$P(Y="Tepat Waktu") = 0,88 * (0,3125 * 0,4375 * 1 * 1 * 0,9375 * 0,8125 * 1 * 1 * 1 * 0,0625 * 1 * 1) = 0,00573$$

c) Membandingkan Probabilitas Label dan Kelas

Setelah melihat uraian perhitungan dari semua atribut di atas (X1-X13 atau probabilitas JK, IPS1, IPS2, IPS3, IPS4, IPS5, JKK, LTA, MSTUDI, ORG, BIAYA, dan WPTA, yang diakhiri dengan mengakumulasi nilai-nilai probabilitas atribut tersebut untuk probabilitas kelas yang dicari. Adapun probabilitas kelas yang dicari adalah kelas kelulusan tepat waktu dan kelas kelulusan tidak tepat waktu. Maka selanjutnya membandingkan nilai probabilitas dari kedua kelas tersebut. Jika dilihat dari akumulasi nilai probabilitas kelas tersebut, yang masing-masing yaitu 0,01375 untuk kelas tepat waktu, dan 0,00573 untuk kelas tidak tepat waktu. Untuk itu perbandingan kedua kelas adalah :

$P(Y= \text{“Tepat Waktu”}) < P(Y= \text{“Tidak Tepat Waktu”})$
 ATAU
 $P(Y= \text{“Tepat Waktu”}) > P(Y= \text{“Tidak Tepat Waktu”})$

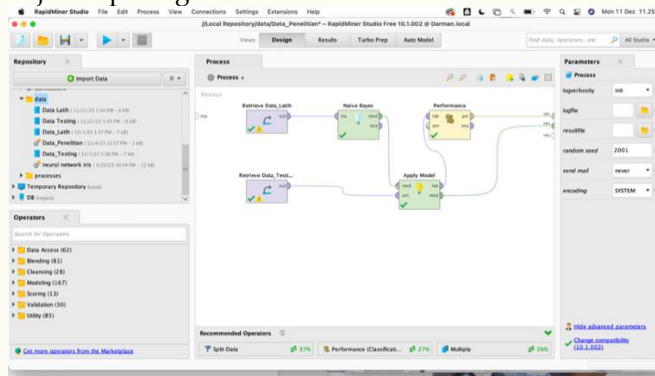
Dengan Demikian, Didapatkan Kesimpulan Bahwa Dari Data Pelatihan Di Atas, mahasiswa dominan masuk dalam kategori kelas “ Tidak Tepat Waktu” dengan rincian 16 mahasiswa keluar lulus tidak tepat waktu, dan 2 mahasiswa keluar lulus tepat waktu.

Model prediksi kelulusan menggunakan metode klasifikasi naïve bayes di atas, jika diterapkan pada data pengujian sebanyak 9 mahasiswa yang keluar lulus pada periode semester ganjil dan genap 2022/2023 dapat disimpulkan bahwa ada 5 mahasiswa keluar lulus tidak tepat waktu dan 4 mahasiswa keluar lulus tepat waktu dengan akurasi 100% . Sehingga dapat dikatakan Sebagian besar mahasiswa keluar lulus masuk dalam kategori kelas “ Tidak Tepat Waktu”.

5. Penerapan Model dengan RapidMiner

a) Pengujian data dengan Model

Setelah merancang model dibentuk dan diterapkan dalam melakukan pelatihan data, selanjutnya pengujian data sudah dapat dilakukan. Namun dalam proses pembelajaran mesin dengan RapiMiner, perlu mengkonfigurasi data dengan model yang digunakan. Adapun konfigurasi data dan model pelatihan akan tampak seperti pada interfaces model ditunjukkan pada gambar 2



Gambar 2. Interfaces Model Untuk Pengujian Data

Setelah rancangan model dikonfigurasi selanjutnya pengujian data dilakukan. Adapun pengujian data menggunakan data testing yang terdiri dari 9 orang mahasiswa. Setelah dilakukan, diperoleh data prediksi kelulusan tepat waktu dan tidak tepat waktu seperti yang ditampilkan pada gambar 3.

Row No.	KL	predictionL	confidenceL	confidenceL	JK	IPS1	IPS2	IPS3	IPS4
1	TL	TL	1.000	0.000	P	IPS1 < 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
2	TL	TL	1.000	0.000	P	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
3	TL	TL	1.000	0.000	P	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
4	TL	TL	1.000	0.000	P	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
5	TP	TP	0.000	1.000	P	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
6	TP	TP	0.000	1.000	L	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
7	TP	TP	0.000	1.000	P	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
8	TL	TL	1.000	0.000	P	IPS1 < 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00
9	TP	TP	0.000	1.000	L	IPS1 > = 3,00	IPS2 > = 3,00	IPS3 > = 3,00	IPS4 > = 3,00

Implementasi Metode Naïve Bayes Dalam Memprediksi Kelulusan Mahasiswa Tepat Waktu Pada Program Studi Pendidikan Teknologi Informasi

IPS3	IPS4	IPS5	JKK	LTA	STUDI	JALUR	ORG	BIAYA	WPTA
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA > 1	STUDI > 4	UMUM	AKTIF	M	VII
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA > 1	STUDI > 4	UMUM	TIDAK	M	VII
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA > 1	STUDI > 4	UMUM	AKTIF	M	VII
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA > 1	STUDI > 4	UMUM	TIDAK	M	VII
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA <= 1	STUDI < 4	UMUM	TIDAK	M	VI
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA <= 1	STUDI < 4	UMUM	TIDAK	B	VI
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA <= 1	STUDI < 4	UMUM	TIDAK	M	VI
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA > 1	STUDI > 4	UMUM	AKTIF	M	VII
IPS3 >= 3,00	IPS4 >= 3,00	IPS5 >= 3,00	JKK >= 100	LTA <= 1	STUDI < 4	UMUM	AKTIF	B	VI

Gambar 3. Hasil Prediksi Dari Pengujian Data

Jika dilihat dari gambar 3 di atas tampak bahwa model prediksi berhasil melakukan klasifikasi kelulusan. Dari gambar tersebut diperoleh kesimpulan bahwa 4 mahasiswa masuk dalam kategori kelulusan “Tepat waktu” dan 5 mahasiswa masuk dalam kategori kelulusan “ Tidak Tepat Waktu”.

b) Akurasi Model

Untuk melihat hubungan antara model prediksi dengan pendekatan matematis dan penerapan aplikasin RapidMiner, perlu mempertimbangkan tingkat akurasi model yang digunakan. Akurasi pelatihan model untuk pengujian data dilakukan dengan RapidMiner ditampilkan gambar confusion matrikx dan performance vector dengan tingkat akurasi mencapai 100%, dapat dilihat pada gambar 4 .



Gambar 4. Akurasi Pelatihan Model

Data prediksi yang disajikan dari gambar 4 di atas dapat dilihat bahwa dari pengujian dari 9 data lulusan menunjukkan klasifikasi kelulusan tepat waktu dan tidak tepat waktu sesuai dengan model yang dirancang. Sementara akurasi untuk pengujian menunjukkan akurasi 100%.

Pembahasan

Berdasarkan uraian hasil di atas, dapat dilihat bahwa hasil model prediksi, baik melalui model pendekatan matematis metode klasifikasi naïve bayes, serta penerapan model prediksi naïve bayes dengan RapidMiner didapat bahwa kedua model tersebut menunjukkan sinkronisasi yang kuat. dengan kata lain pendekatan matematis metode klasifikasi naïve bayes dapat diterapkan dalam

memprediksi kelulusan tepat waktu dan tidak tepat waktu mahasiswa. Dengan akurasi 100%, hal ini senada yang dikatakan (Armansyah & Kurniawan, 2021; Damanik et al., 2021) bahwa algoritma naïve bayes dapat digunakan dalam memodelkan prediksi.

Memodelkan prediksi kelulusan mahasiswa tepat waktu dengan. Melibatkan factor-faktor yang mendukung kelulusan mahasiswa tepat waktu, dalam penelitian ini factor yang menjadi menjadi kelulusan mahasiswa yaitu Jenis Kelamin (JK), Indek Prestasi mulai semester 1 sampai 5 (IPS1-IPS5), Jumlah Kredit Kumulatif (JKK), Lama Tugas Akhir (LTA), Masa Studi (STUDI), Jalur masuk (JALUR), Aktif organisasi (ORG), Biaya dan Waktu Pengajuan judul tugas akhir (WPTA).

Namun perlu dipertimbangkan pemilihan atribut lain selain atribut yang telah dijelaskan di atas. Ini dipandang sangat penting mengingat hampir semua kriteria yang dimiliki oleh mahasiswa secara logis memenuhi syarat terpenuhinya kelulusan tepat waktu bahkan lebih awal. Contoh seperti atribut IP Semester 1 sampai IP Semester 5 didominasi oleh lulusan yang memiliki nilai yang baik bahkan beberapa diantaranya mendapat nilai cumlaud, Sementara Jumlah Kredit Kumulatif disemester 5 yang diperoleh mahasiswa dipandang cukup tinggi bahkan melebihi batas normal sks mahasiswa di semester 5. Namun faktanya menunjukkan, mahasiswa dengan kriteria yang tersebut diatas, masuk kedalam klasifikasi kelulusan yang "tidak tepat waktu atau terlambat", artinya ada factor lain yang mempengaruhi keterlambatan mahasiswa Program Studi Pendidikan Teknologi Informasi Universitas Muhammadiyah Kendari, yang tidak digunakan dalam penelitian ini.. Hal ini senada dengan yang dijelaskan oleh (Armansyah & Kurniawan, 2021), dalam penelitiannya menyatakan bahwa untuk memprediksi kelulusan mahasiswa tepat waktu dan tidak tepat waktu dapat dipecahkan dengan algoritma naïve bayes.

KESIMPULAN

Setelah melakukan pemaparan hasil dan pembahasan penelitian yang diuraikan di atas yang merujuk pada permasalahan dalam penelitian ini dapat disimpulkan sebagai berikut: (1) Memprediksi Kelulusan mahasiswa tepat waktu dan tidak tepat waktu dengan melibatkan faktor-faktor yang mendukung kelulusan mahasiswa tepat waktu dan tidak tepat waktu dapat dipecahkan dengan menggunakan Algoritma Naïve Bayes, dimulai dengan melakukan analisis kebutuhan data pelatihan (training) dan data pengujian (testing), melakukan transformasi data baik data pelatihan maupun data testing dan penyajian data. Data-Data tersebut digunakan dalam memodelkan prediksi kelulusan mahasiswa tepat waktu dan tidak tepat waktu dengan model naïve bayes. (2) Faktor-faktor yang digunakan dalam melakukan klasifikasi kelas lulus tepat waktu dan tidak tepat waktu diantaranya adalah Jenis Kelamin (JK), Indek Prestasi mulai semester 1 sampai 5 (IPS1-IPS5), Jumlah Kredit Kumulatif (JKK), Lama Tugas Akhir (LTA), Masa Studi (STUDI), Jalur masuk (JALUR), Aktif Organisasi (ORG), Biaya dan Waktu Pengajuan judul tugas akhir (WPTA). (3) Metode naïve bayes dapat digunakan dalam memprediksi kelulusan mahasiswa tepat waktu, dengan probabilitas Kelulusan Tepat Waktu Sebesar 11% dan Tidak Tepat Waktu Sebesar 89%. Dengan Akurasi 100%, dengan hasil pelatihan dan pengujian tanpa error. Ini dibuktikan pada data pengujian model melalui RapidMiner yang menampilkan kelulusan tidak tepat waktu untuk semua record data testing yaitu sebanyak 5 data mahasiswa. Data tersebut sama dengan hasil perhitungan model matematika naïve bayes dengan penerapan *tools RapidMiner*. Temuan ini dapat dijadikan rujukan bagi kalangan peneliti untuk melakukan penelitian lanjutan dengan menerapkan atribut-atribut lain, dengan juga memperhatikan fenomena yang terjadi atau yang sedang berlaku.

DAFTAR PUSTAKA

- Armansyah, & Kurniawan, R. (2021). *Model Prediksi Kelulusan Mahasiswa Tepat Waktu*. Universitas Islam Negeri Sumatera Utara. <http://repository.uinsu.ac.id/13307/>
- Damanik, A. R., Sumijan, S., & Nurcahyo, G. W. (2021). Prediksi Tingkat Kepuasan dalam Pembelajaran Daring Menggunakan Algoritma Naïve Bayes. *Jurnal Sistim Informasi Dan Teknologi*, 88–94. <https://doi.org/10.37034/jsisfotek.v3i3.49>
- Fitriani, A. S. (2019). Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 3(2), 98–104. <https://doi.org/10.31764/jtam.v3i2.995>
- Herlambang, A. D., & Wijoyo, S. H. (2019). Algoritma Naive Bayes untuk Klasifikasi Sumber Belajar Berbasis Teks pada Mata Pelajaran Produktif di SMK Rumpun Teknologi Informasi dan Komunikasi. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 6(4), 430-436. <https://doi.org/10.25126/jtiik.201961323>
- Irmayani, W. (2021). Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi Naive Bayes. *Jurnal Khatulistiwa Informatika*, 9(1), 68-72.
- Mustofa, H., & Mahfudh, A. A. (2019). Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes. *Walisongo Journal of Information Technology*, 1(1), 1. <https://doi.org/10.21580/wjit.2019.1.1.3915>
- Putu, N. L. P. M., Ahmad Zuli Amrullah, & Ismarmiaty. (2021). Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 123–131. <https://doi.org/10.29207/resti.v5i1.2587>
- Saleh, A. (2015). Implementasi metode klasifikasi naive bayes dalam memprediksi besarnya penggunaan listrik rumah tangga. *Creative Information Technology Journal*, 2(3), 207-217.
- Sihombing, I. A., Hartama, D., Parlina, I., Gunawan, I., & Kirana, I. O. (2021). Analisis keberhasilan pembelajaran daring pada masa pandemi Covid-19 menggunakan algoritma C4. 5 dan Naive Bayes. *JUKI: Jurnal Komputer dan Informatika*, 3(2), 89-96. <https://doi.org/10.53842/juki.v3i2.68>
- Watratan, A. F., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7-14. <https://doi.org/10.52158/jacost.v1i1.9>
- Yunus, M., Biddinika, M. K., & Fadlil, A. (2023). Optimasi Algoritma Naïve Bayes Menggunakan Fitur Seleksi Backward Elimination untuk Klasifikasi Prevalensi Stunting. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 278-285. <https://doi.org/10.51454/decode.v3i2.188>