

Prediksi Nilai Akhir Mahasiswa Dengan Metode Naïve Bayes

Prediction of Final Student Grades Using Naive Bayes Method

Brama Rizky Setia Dwinanda^{*1}, Azriel Noer Ilhami², Hana Choerunisya³, Indra Jiwana Thira⁴

^{1,2,3}Program Studi Teknologi Informasi, Fakultas Komunikasi dan Informasi, Universitas Garut

⁴Program Studi Rekayasa Perangkat Lunak, Fakultas Komunikasi dan Informasi, Universitas Garut

Email: ¹24072121005@fkominfo.uniga.ac.id, ²24072121013@fkominfo.uniga.ac.id,

³24072121026@fkominfo.uniga.ac.id, ⁴indrajiwana@uniga.ac.id

^{*}Penulis Koresponden

Received: 2 November 2023

Accepted: 20 Januari 2024

Published: 05 Februari 2024



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
Copyright (c) 2024 JUSTINDO

ABSTRAK

Memprediksi prestasi belajar atau nilai akhir mahasiswa menjadi suatu hal yang sangat berguna karena dapat mencegah mahasiswa yang beresiko gagal dalam perkuliahan. Prediksi merupakan upaya untuk memperkirakan kemungkinan suatu kejadian, dan hasil dari prediksi tidak dapat dijamin akan terjadi secara pasti. Meskipun begitu kita bisa mengusahakan untuk mencari jawaban sedekat mungkin dengan apa yang akan terjadi. Penelitian ini bertujuan untuk memprediksi nilai akhir mahasiswa berdasarkan IPK yang didapat pada semester 2, karena nilai tersebut bisa digunakan sebagai penilaian awal untuk memperkirakan adanya peningkatan atau penurunan nilai yang akan didapat pada semester berikutnya. *Output* dari penelitian ini adalah rentang nilai yang disesuaikan dengan predikat kelulusan. Data yang digunakan dalam penelitian ini berjumlah 379 yang kemudian dibagi menjadi dua bagian, yaitu data *training* dengan jumlah 303 data dan data *testing* dengan jumlah 76 data. Penelitian ini menggunakan metode *naïve bayes* untuk menghitung probabilitas berdasarkan peristiwa yang telah terjadi sebelumnya. Keterbatasan dataset yang dimiliki, menyebabkan penyebaran data dalam penelitian ini tidaklah merata, sehingga memengaruhi hasil dari penelitian. Hasil penelitian menunjukkan tingkat akurasi dari model algoritma *naïve bayes* yang digunakan sebesar 78.95%.

Kata kunci: *Prediksi, IPK, Naïve Bayes, Probabilitas*

ABSTRACT

Predicting student performance or final grades is a very useful thing because it can prevent students who are at risk of failing in college. Prediction is an attempt to estimate the likelihood of an event, and the results of the prediction cannot be guaranteed to happen for sure. However, we can try to find answers as close as possible to what will happen. This study aims to predict student final grades based on the GPA obtained in semester 2, because this value can be used as an initial assessment to estimate any increase or decrease in grades that will be obtained in the next semester. The output of this study is a range of values that are adjusted to the graduation predicate. The data used in this study amounted to 379 which were then divided into two parts, namely the training data with 303 data and the testing data with 76 data. This study used the naïve bayes method to calculate probabilities based on events that have occurred before. The limitations of the dataset owned, caused the distribution of data in this study was not even, thus affecting the results of the study. The results of the study showed that the accuracy rate of the naïve bayes algorithm model used was 78.95%

Keywords: *Prediction, GPA, Naïve Bayes, Probability*

1. Pendahuluan

Keberhasilan mahasiswa dalam proses belajar dapat diukur dengan nilai akhir yang diperoleh. Nilai akhir yang baik mencerminkan keberhasilan dalam kegiatan belajar mengajar. Keberhasilan ini dapat diartikan sebagai keberhasilan dalam hal pengetahuan, keahlian, dan kemampuan mahasiswa. Masalah yang timbul dalam proses perkuliahan adalah tidak semua proses menghasilkan hasil yang baik. Hal ini dapat disebabkan oleh beberapa faktor, salah satunya adalah tidak adanya pantauan prestasi akademik yang baik dari institusi. Kondisi ini dapat menyulitkan mahasiswa untuk lulus tepat waktu dan dapat memengaruhi kualitas program studi. Prestasi akademik menjadi salah satu indikator keberhasilan pendidikan. Salah satu indikator keberhasilan tersebut adalah indeks prestasi semester (IPS) dan indeks prestasi kumulatif (IPK). (Hartatik, 2020).

Prediksi prestasi akademik atau nilai akhir mahasiswa dapat menjadi suatu hal yang sangat berguna apabila mendapat hasil prediksi yang mendekati hasil aktual atau hasil yang didapat secara nyata. Informasi tentang prestasi akademik atau nilai akhir mahasiswa yang diketahui sejak dini dapat mencegah mahasiswa yang diprediksi memiliki prestasi atau nilai yang rendah sehingga mengurangi resiko atau bahkan mencegah adanya mahasiswa yang gagal dalam perkuliahan (Hasudungan & Pranoto, 2021).

Prediksi merupakan suatu proses meramalkan suatu kejadian di masa depan atau masa mendatang berdasarkan data yang sudah ada, atau berdasarkan kejadian yang telah terjadi sebelumnya. Hasil dari prediksi tidak dapat dijamin akan terjadi secara pasti. Meskipun begitu kita bisa mengusahakan untuk mencari jawaban sedekat mungkin dengan apa yang akan terjadi (Hafiz, 2019). Prediksi nilai akhir mahasiswa dapat dilakukan dengan pendekatan *machine learning*. *Machine learning* adalah teknik yang memungkinkan mesin atau aplikasi untuk belajar dari data dengan tujuan mengelompokkan, mengklasifikasikan, atau memprediksi masa depan. Salah satu metode klasifikasi yang dapat digunakan dalam *machine learning* untuk tujuan prediksi adalah *Naïve Bayes* (Armansyah & Ramli, 2022).

Naïve Bayes Classifier adalah metode pengklasifikasian yang berdasarkan kepada *Teorema Bayes*. Proses kerja algoritma ini didasarkan kepada prinsip probabilitas bersyarat, *Teorema Bayes* menemukan probabilitas peluang atau kemungkinan suatu peristiwa akan terjadi berdasarkan peristiwa yang telah terjadi sebelumnya. Dengan kata lain, *Teorema Bayes* adalah metode untuk menemukan probabilitas atau menghitung kemungkinan ketika kita mengetahui probabilitas tertentu yang sudah ada (Muin, Rahman, & Abdillah, 2022).

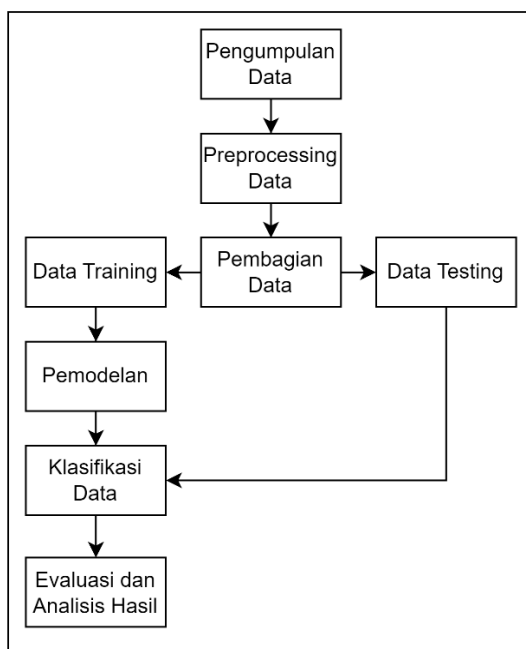
Metode *Naïve Bayes* digunakan pada beberapa penelitian sebelumnya untuk memprediksi kelulusan mahasiswa, diantaranya adalah (Wibowo & Rohman, 2022) melakukan penelitian untuk memprediksi predikat kelulusan mahasiswa dan mendapatkan akurasi sebesar 86%. Selain itu, metode ini juga digunakan oleh (Hendrawan, Saputra, Dewi, Pranata, & Wedasari, 2021) untuk mengklasifikasikan lama studi dan predikat kelulusan mahasiswa dengan data dari Institut Teknologi dan Bisnis STIKOM Bali. Hasil penelitian tersebut menunjukkan tingkat akurasi untuk lama studi 74% dan predikat kelulusan 61% pada Program Studi Sistem Komputer. Kemudian pada Program Studi Sistem Informasi mendapat tingkat akurasi untuk lama studi 73% dan predikat kelulusan 67%.

Penelitian lainnya dilakukan oleh (Novianto, Hermawan, & Avianto, 2023) dengan menggunakan metode klasifikasi *k-nearest neighbor*, *naïve bayes*, dan *decision tree* untuk memprediksi status kelulusan mahasiswa Program Studi Ilmu Hukum, Fakultas Hukum, Universitas Sebelas Maret menggunakan aplikasi RapidMiner. Dari penelitian itu, masing-masing metode menunjukkan tingkat akurasi sebesar 96.67%, 77.33%, dan 94%. Kemudian penelitian lain yang memprediksi status kelulusan mahasiswa menggunakan metode *naïve bayes* yaitu (Aji, Atina, & Sudiby, 2023) dengan jumlah data sebanyak 302 data yang terdiri dari atribut jenis kelamin, status mahasiswa, status menikah, umur, indeks prestasi dari semester 1 sampai dengan semester 8, dan indeks prestasi kumulatif (IPK). Hasil akurasi prediksi dari metode tersebut adalah 85%.

Dalam penelitian ini dilakukan prediksi nilai akhir atau IPK mahasiswa, berdasarkan IP yang didapat pada semester 1, IP semester 2, dan akumulasi rata-rata dari kedua nilai tersebut atau bisa diasumsikan sebagai IPK semester 2 menggunakan algoritma *naïve bayes*, sehingga diperoleh prediksi nilai akhir berdasar kelas predikat yang dibagi menjadi tidak normal, normal, memuaskan, sangat memuaskan, dan pujian.

2. Metode Penelitian

Penelitian ini akan memprediksi nilai akhir mahasiswa berdasarkan kelas predikat yang didapatkan. Rancangan penelitian yang dimulai dari tahap pengumpulan data sampai dengan evaluasi dan analisis hasil akhir dapat dilihat pada Gambar 1.



Gambar 1. Rancangan Penelitian

Penelitian ini akan didasarkan kepada dataset yang berisi nilai Indeks Prestasi setiap semester sampai dengan Indeks Prestasi Kumulatif atau IPK.

Dataset yang digunakan dalam penelitian ini adalah data yang terdiri dari nama, jenis kelamin, status mahasiswa, umur, status menikah, indeks prestasi dari semester 1 sampai dengan semester 8, indeks prestasi kumulatif, dan juga status kelulusan. Dataset tersebut diperoleh dari <https://www.kaggle.com/datasets/permanaraga/student-graduation/data>. Total jumlah nilai dalam dataset tersebut adalah 379 nilai.

Sebelum langsung digunakan, dataset melalui tahapan berikut ini terlebih dahulu:

1. *Data cleaning* (menghapus dan atau memperbaiki data yang memiliki nilai kosong, kesalahan penulisan, juga duplikasi data).
2. *Data selection* (memilih data yang relevan dengan tugas analisis).
3. *Data transformation* (mengubah atau menyesuaikan data menjadi bentuk yang sesuai untuk dilakukan analisis) (Nurviana, Ati, & Hanifah, 2019).

Variabel yang digunakan dalam penelitian ini hanyalah IPS 1, IPS 2, dan IPK, karena penelitian ini berfokus untuk menentukan nilai akhir berdasarkan IPK yang didapat pada semester 2. Lalu, ada penambahan variabel, yaitu IPK 2 yang merupakan nilai akumulasi tadi. Jadi, total variabel yang digunakan dalam penelitian ini yaitu IPS 1, IPS 2, IPK 2, dan IPK.

Variabel IPK atau nilai akhir yang akan diprediksi telah diubah menjadi variabel Predikat dan nilai yang ada dalam variabel tersebut pun telah diubah, awalnya berupa nilai angka dan diubah menjadi predikat kelulusan. Daftar kelas pada dataset yang digunakan yaitu tidak normal, normal, memuaskan, sangat memuaskan, dan pujian. Dataset yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1.

Tabel 1. Dataset

IPS 1	IPS 2	IPK 2	Predikat
2.29	0.5	1.395	Tidak Normal

2.76	1.09	1.925	Tidak Normal
2.45	1.43	1.94	Tidak Normal
2.64	1.61	2.125	Tidak Normal
2.21	2.03	2.12	Tidak Normal
...
3.71	3.6	3.655	Pujian
3.71	3.68	3.695	Pujian
3.71	3.79	3.75	Pujian
3.76	3.81	3.785	Pujian
3.64	3.96	3.8	Pujian

Proses pembagian dataset menjadi data *training* dan data *testing* membagi data dengan rasio 80:20. Pembagian dataset ini didasarkan kepada masing-masing kelas pada variabel Predikat, untuk lebih jelasnya dapat dilihat pada Tabel 2.

Data *training* merupakan suatu proses untuk melatih algoritma menggunakan sebagian data yang terdapat dalam dataset untuk mempelajari dataset itu sendiri. Proses ini bertujuan untuk menghasilkan sebuah model yang nantinya akan digunakan pada saat melakukan proses pengujian yang dilakukan terhadap data *testing* (Fremmuzar & Baita, 2023).

Tabel 2. Pembagian Dataset

Kelas	Data Training	Data Testing
Tidak Normal	12	3
Normal	66	17
Memuaskan	72	18
Sangat Memuaskan	133	33
Pujian	20	5

Model yang digunakan dalam penelitian ini adalah algoritma *Naïve Bayes*. Prediksi Bayes didasarkan pada teorema Bayes dengan formula umum sebagai berikut:

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

$P(H|E)$: *Posterior*: Probabilitas akhir bersyarat suatu hipotesis H terjadi jika diberikan bukti E terjadi.

$P(E|H)$: *Likelihood*: Probabilitas sebuah bukti E akan memengaruhi hipotesis H.

$P(H)$: *Prior Probability*: Probabilitas awal hipotesis H terjadi tanpa memandang bukti apapun.

$P(E)$: *Evidence*: Probabilitas awal bukti E terjadi tanpa memandang hipotesis/bukti yang lain (Wibowo & Djafar, 2023).

3. Hasil dan Pembahasan

Semua Pembagian kelas pada variabel Predikat disesuaikan dengan kategori predikat kelulusan mahasiswa yang diatur dalam SNDikti, dengan ketentuan lebih jelas sebagai berikut.

SNDIKTI-PEND-4-01-01 Mahasiswa dinyatakan lulus dengan predikat memuaskan apabila mencapai Indeks Prestasi Kumulatif (IPK) 2,76 (dua koma tujuh enam) sampai dengan 3,00 (tiga koma nol nol)

SNDIKTI-PEND-4-01-02 Mahasiswa dinyatakan lulus dengan predikat sangat memuaskan apabila mencapai Indeks Prestasi Kumulatif (IPK) 3,01 (tiga koma nol satu) sampai dengan 3,50 (tiga koma lima nol)

SNDIKTI-PEND-4-01-03 Mahasiswa dinyatakan lulus dengan predikat pujian apabila mencapai Indeks Prestasi Kumulatif (IPK) lebih dari 3,50 (tiga koma nol) (Hidayat, Iskandar, Afuan, Nugroho, & Arwiansyah, 2022).

Kemudian, karena dalam dataset terdapat Indeks Prestasi Kumulatif diluar ketentuan tersebut, maka penulis menambahkan 2 predikat baru untuk penelitian ini, yaitu normal dan tidak normal.

Predikat normal digunakan apabila mahasiswa lulus dengan Indeks Prestasi Kumulatif (IPK) 2,00 (dua koma nol nol) sampai dengan 2,75 (dua koma tujuh lima)

Predikat tidak normal digunakan apabila mahasiswa lulus dengan Indeks Prestasi Kumulatif (IPK) kurang dari 2,00 (dua koma nol nol).

Masing-masing variabel yang digunakan dalam penelitian ini dibagi menjadi 5 kelas yang disesuaikan dengan ketentuan dari variabel Predikat.

Probabilitas Predikat dari data training yang digunakan adalah sebagai berikut.

Tabel 3. Probabilitas Predikat

Probabilitas Predikat		
Status	Banyak	Nilai
Tidak Normal	12	0.04
Normal	66	0.22
Memuaskan	72	0.24
Sangat Memuaskan	133	0.44
Pujian	20	0.07
Total	303	1.00

Setelah memetakan probabilitas dari Predikat, penelitian ini dilanjutkan dengan memetakan masing-masing variabel IPS 1, IPS 2, dan IPK 2 terhadap Predikat, untuk menentukan nilai yang nantinya bisa dihitung dan membuat prediksi. Maka bisa didapatkan probabilitas dari atribut-atribut tersebut sebagai berikut (Wahyuni & Adinda, 2021).

Tabel 4. Sebaran data dan probabilitas IPS 1

Nilai	Predikat				
	Tidak Normal	Normal	Memuaskan	Sangat Memuaskan	Pujian
<2	1	3	0	1	0
2 – 2.75	7	43	40	15	3
2.76 – 3	3	11	25	46	1
3 – 3.5	1	9	7	63	10
>3.5	0	0	0	8	6
Total	12	66	72	133	20
Probabilitas					
<2	0.08	0.05	0.00	0.01	0.00
2 – 2.75	0.58	0.65	0.56	0.11	0.15
2.76 – 3	0.25	0.17	0.35	0.35	0.05
3 – 3.5	0.08	0.14	0.10	0.47	0.50
>3.5	0.00	0.00	0.00	0.06	0.30
Total	1	1	1	1	1

Tabel tersebut menunjukkan probabilitas dari variabel Indeks Prestasi Semester 1 terhadap masing-masing kelas dalam variabel Predikat.

Tabel 5. Sebaran data dan probabilitas IPS 2

Nilai	Predikat				
	Tidak Normal	Normal	Memuaskan	Sangat Memuaskan	Pujian
<2	4	8	0	0	0
2 – 2.75	5	46	41	25	0
2.76 – 3	1	8	18	43	0
3 – 3.5	2	4	13	56	12
>3.5	0	0	0	9	8
Total	12	66	72	133	20
Probabilitas					
<2	0.33	0.12	0.00	0.00	0.00
2 – 2.75	0.42	0.70	0.57	0.19	0.00
2.76 – 3	0.08	0.12	0.25	0.32	0.00
3 – 3.5	0.17	0.06	0.18	0.42	0.60
>3.5	0.00	0.00	0.00	0.07	0.40
Total	1	1	1	1	1

Tabel tersebut menunjukkan probabilitas dari variabel Indeks Prestasi Semester 2 terhadap masing-masing kelas dalam variabel Predikat.

Tabel 6. Sebaran data dan probabilitas IPK 2

Nilai	Predikat				
	Tidak Normal	Normal	Memuaskan	Sangat Memuaskan	Pujian
<2	2	5	0	1	0
2 – 2.75	7	49	42	11	0
2.76 – 3	2	9	23	52	3
3 – 3.5	1	3	7	65	13
>3.5	0	0	0	4	4
Total	12	66	72	133	20
Probabilitas					
<2	0.17	0.08	0.00	0.01	0.00
2 – 2.75	0.58	0.74	0.58	0.08	0.00
2.76 – 3	0.17	0.14	0.32	0.39	0.15
3 – 3.5	0.08	0.05	0.10	0.49	0.65
>3.5	0.00	0.00	0.00	0.03	0.20
Total	1	1	1	1	1

Tabel tersebut menunjukkan probabilitas dari variabel Indeks Prestasi Kumulatif semester 2 terhadap masing-masing kelas dalam variabel Predikat.

Setelah mengetahui nilai probabilitas masing-masing kelas pada variabel Predikat, dan juga nilai probabilitas variabel lainnya terhadap kelas Predikat, selanjutnya bisa dilakukan prediksi nilai akhir atau predikat kelulusan mahasiswa.

Seorang mahasiswa memiliki nilai IPS 1 yaitu 2.57, dan nilai IPS 2 yaitu 3.02. Berapakah prediksi nilai akhir mahasiswa tersebut?

Tabel 7. Membuat Prediksi Nilai Akhir

Probabilitas	Tidak Normal	Normal	Memuaskan	Sangat Memuaskan	Pujian
Kelas	0.04	0.22	0.24	0.44	0.07
IPS 1 2.57	0.58	0.65	0.56	0.11	0.15
IPS 2 3.02	0.17	0.06	0.18	0.42	0.60
IPK 2 2.795	0.17	0.14	0.32	0.39	0.15
Total	0.0006	0.0012	0.0077	0.0079	0.0009

Pembuatan prediksi tersebut didasarkan kepada perhitungan nilai probabilitas Predikat, dan juga nilai masing-masing variabel terhadap Predikat. Untuk mendapatkan nilai pada kolom Prediksi, dapat dilakukan dengan mengalikan semua probabilitas berdasarkan Predikat yang sama. Jadi, nilai terbesar berdasarkan perkalian tersebut yang dijadikan prediksi nilai akhir ini (Fadrial, 2021), (Gerhana, Fallah, Zulfikar, Maylawati, & Ramdhani, 2019).

Dikarenakan total nilai yang dihasilkan oleh kelas Sangat Memuaskan lebih besar daripada nilai lainnya, maka prediksi nilai akhir atau IPK mahasiswa tersebut ada di rentang 3–3.5, atau mendapat predikat Sangat Memuaskan.

Dari 303 data mahasiswa yang digunakan sebagai *data training* dan membangun suatu model, didapat hasil sebaran prediksi seperti yang diperlihatkan pada Tabel 8.

Tabel 8. Sebaran Prediksi Data Training

Sebaran Nilai IPS 1: <2					
Data Training			Prediksi		
IPS 1	IPS 2	IPK 2	Predikat	IPK	
<2	<2	<2	Tidak Normal	<2	
<2	2 - 2.75	<2	Normal	2 - 2.75	
<2	2 - 2.75	2 - 2.75	Normal	2 - 2.75	
<2	2.76 - 3	<2	Normal	2 - 2.75	
<2	2.76 - 3	2 - 2.75	Normal	2 - 2.75	
Sebaran Nilai IPS 1: 2 - 2.75					
Data Training			Prediksi		
IPS 1	IPS 2	IPK 2	Predikat	IPK	
2 - 2.75	<2	<2	Normal	2 - 2.75	
2 - 2.75	<2	2 - 2.75	Normal	2 - 2.75	
2 - 2.75	2 - 2.75	2 - 2.75	Normal	2 - 2.75	
2 - 2.75	2.76 - 3	2 - 2.75	Memuaskan	2.76 - 3	

2 - 2.75	2.76 - 3	2.76 - 3	Memuaskan	2.76 - 3
2 - 2.75	3	2 - 2.75	Memuaskan	2.76 - 3
2 - 2.75	3 - 3.5	2.76 - 3	Sangat Memuaskan	3 - 3.5
2 - 2.75	3 - 3.5	3 - 3.5	Sangat Memuaskan	3 - 3.5
2 - 2.75	>3.5	2.76 - 3	Sangat Memuaskan	3 - 3.5
2 - 2.75	>3.5	3 - 3.5	Pujian	>3.5
Sebaran Nilai IPS 1: 2.76 - 3				
Data Training			Prediksi	
IPS 1	IPS 2	IPK 2	Predikat	IPK
2.76 - 3	<2	<2	Tidak Normal	<2
2.76 - 3	<2	2 - 2.75	Normal	2 - 2.75
2.76 - 3	2 - 2.75	2 - 2.75	Memuaskan	2.76 - 3
2.76 - 3	2 - 2.75	2.76 - 3	Memuaskan	2.76 - 3
2.76 - 3	2.76 - 3	2.76 - 3	Sangat Memuaskan	3 - 3.5
2.76 - 3	3 - 3.5	2.76 - 3	Sangat Memuaskan	3 - 3.5
2.76 - 3	3 - 3.5	3 - 3.5	Sangat Memuaskan	3 - 3.5
2.76 - 3	>3.5	3 - 3.5	Sangat Memuaskan	3 - 3.5
Sebaran Nilai IPS 1: 3 - 3.5				
Data Training			Prediksi	
IPS 1	IPS 2	IPK 2	Predikat	IPK
3 - 3.5	<2	<2	Normal	2 - 2.75
3 - 3.5	<2	2 - 2.75	Normal	2 - 2.75
3 - 3.5	2 - 2.75	2 - 2.75	Normal	2 - 2.75
3 - 3.5	2 - 2.75	2.76 - 3	Sangat Memuaskan	3 - 3.5
3 - 3.5	2 - 2.75	3 - 3.5	Sangat Memuaskan	3 - 3.5
3 - 3.5	2.76 - 3	2.76 - 3	Sangat Memuaskan	3 - 3.5
3 - 3.5	2.76 - 3	3 - 3.5	Sangat Memuaskan	3 - 3.5
3 - 3.5	3 - 3.5	3 - 3.5	Sangat Memuaskan	3 - 3.5
3 - 3.5	>3.5	3 - 3.5	Sangat Memuaskan	3 - 3.5
3 - 3.5	>3.5	>3.5	Pujian	>3.5
Sebaran Nilai IPS 1: >3.5				
Data Training			Prediksi	
IPS 1	IPS 2	IPK 2	Predikat	IPK
>3.5	2 - 2.75	2.76 - 3	Sangat Memuaskan	3 - 3.5
>3.5	2 - 2.75	3 - 3.5	Sangat Memuaskan	3 - 3.5
>3.5	2.76 - 3	3 - 3.5	Sangat Memuaskan	3 - 3.5
>3.5	3 - 3.5	3 - 3.5	Pujian	>3.5
>3.5	3 - 3.5	>3.5	Pujian	>3.5
>3.5	>3.5	>3.5	Pujian	>3.5

Berdasarkan tabel tersebut, bisa dikatakan bahwa IPK yang didapat pada semester 2 bisa dijadikan sebagai acuan dasar untuk memprediksi prestasi akademik atau nilai akhir mahasiswa. Jika mengacu kepada model yang dibangun oleh data training, nilai akhir atau IPK akhir yang didapat oleh mahasiswa tidak akan kurang dari IPK yang didapat pada saat semester 2. Kemudian dari 76 data mahasiswa yang digunakan sebagai data *testing* dan diuji terhadap model yang dibangun, didapat hasil seperti yang diperlihatkan pada Tabel 9.

Tabel 9. Hasil Data *Testing*

		Aktual				
		Tidak Normal	Normal	Memuaskan	Sangat Memuaskan	Pujian
Prediksi	Tidak Normal	2	0	0	0	0
	Normal	0	14	3	2	0
	Memuaskan	0	2	12	1	0
	Sangat Memuaskan	1	1	3	29	2
	Pujian	0	0	0	1	3

Terdapat 60 data valid dan 16 data tidak valid. Data valid merupakan data yang nilai prediksi kelas dan nilai aktual kelasnya menunjukkan nilai yang sama. Model yang diujikan memiliki nilai kevalidan atau nilai akurasi sebesar $60 \div 76 \times 100\% = 78.95\%$ dengan nilai *error* 21.05%. (Azahari, Yulindawati, Rosita, & Mallala, 2020)

Nilai akurasi yang didapat menunjukkan bahwa atribut dan juga metode yang digunakan dalam penelitian ini cukup baik. Selain itu, pembagian variabel ke dalam 5 kelas juga sudah cukup efektif. Meskipun demikian, penyebaran data yang tidak merata menyebabkan adanya kelas mayoritas dan

kelas minoritas. Kelas mayoritas adalah kelas yang memiliki jumlah data lebih banyak, sedangkan kelas minoritas adalah kelas yang memiliki jumlah data lebih sedikit. Ketidakseimbangan ini dapat menyebabkan pengklasifikasian cenderung lebih banyak mempelajari pola data mayoritas daripada pola data minoritas. Hal ini dapat menyebabkan keputusan yang dihasilkan dari pengklasifikasian menjadi bias terhadap kelas mayoritas (Hendrawan, Saputra, Dewi, Pranata, & Wedasari, 2021).

4. Kesimpulan

Hasil penelitian menunjukkan bahwa Indeks Prestasi Semester 1, Indeks Prestasi Semester 2, serta Indeks Prestasi Kumulatif semester 2 yang digunakan untuk memprediksi nilai akhir atau predikat IPK mahasiswa cukup baik dengan mendapat nilai akurasi sebesar 78.95%.

Pembagian masing-masing variabel menjadi 5 kelas yang disesuaikan dengan pembagian kelas Predikat dirasa cukup efektif. Kesamaan dalam mendapat probabilitas tinggi terhadap kelas yang sama dari setiap variabel menjadi nilai yang sangat penting untuk diperhitungkan, terlepas dari probabilitas Predikat itu sendiri. Nilai akhir yang didapat kemungkinan berada pada kelas yang sama atau bahkan bisa jadi berada pada kelas yang lebih tinggi daripada nilai IPK semester 2.

Keterbatasan dalam penelitian ini yaitu penyebaran data dari masing-masing kelas yang tidak merata, sehingga menyebabkan ketidakseimbangan data, dan membuat hasil prediksi menjadi bias terhadap suatu kelas yang memiliki lebih banyak data.

Penelitian kedepannya dapat lebih memerhatikan penyebaran data agar tidak terdapat perbedaan jumlah data dalam suatu kelas dengan kelas lainnya, dan juga dapat menambah variabel lain diluar nilai mahasiswa, karena pastinya banyak faktor yang memengaruhi nilai akhir mahasiswa, yang tidak terdapat pada penelitian ini.

Daftar Pustaka

- Aji, N.M., Atina, V., Sudiby, N.A., 2023. PEMODELAN PREDIKSI KELULUSAN MAHASISWA DENGAN METODE NAÏVE BAYES DI UNIBA. *Jurnal Manajemen Informatika dan Sistem Informasi*, 6(2), 148-158. <https://doi.org/10.36595/misi.v5i2>
- Armansyah, A., Ramli, R.K., 2022. Model Prediksi Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika* 6(1), 1–10. <https://doi.org/10.29408/edumatic.v6i1.4789>
- Azahari., Yulindawati., Rosita, D., Mallala, S., 2020. KOMPARASI DATA MINING NAIVE BAYES DAN NEURAL NETWORK MEMPREDIKSI MASA STUDI MAHASISWA S1. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(3), 443–452. <https://doi.org/10.25126/jtik.202072093>
- Fadrial, Y.E., 2021. ALGORITMA NAIVE BAYES UNTUK MENCARI PERKIRAAN WAKTU STUDI MAHASISWA NAIVE BAYES ALGORITHM FOR FINDING STUDENT ESTIMATED TIME STUDENTS. *Journal of Information Technology and Computer Science (INTECOMS)*, 4(1), 20-29. <https://doi.org/10.31539/intecom.v4i1.2219>
- Fremmuzar, P., Baita, A., 2023. Uji Kernel SVM dalam Analisis Sentimen Terhadap Layanan Telkomsel di Media Sosial Twitter. *Jurnal Sistem Komputer*, 12(2), 57-66. <https://doi.org/10.34010/komputika.v12i2.9460>
- Gerhana, Y.A., Fallah, I., Zulfikar, W.B., Maylawati, D.S., Ramdhani, M.A., 2019. Comparison of naive Bayes classifier and C4.5 algorithms in predicting student study period, in: *Journal of Physics: Conference Series*. Institute of Physics Publishing. <https://doi.org/10.1088/1742-6596/1280/2/022022>
- Hafiz, M.I., 2019. PEMANFAATAN METODE CART UNTUK MEMPREDIKSI OMSET SEPATU PRIA. *Pelita Informatika: Informasi dan Informatika*, 8(2), 227-235. <http://www.ejurnal.stmik-budidarma.ac.id/index.php/pelita/article/view/1831>
- Hartatik, 2020. Optimasi Model Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Naive Bayes. *IJAI (Indonesian Journal of Applied Informatics)*, 5(1), 32-38. <https://dx.doi.org/10.20961/ijai.v5i1.44379>

- Hasudungan, R., Pranoto, W.J., 2021. Implementasi Teorema Naïve Bayes Pada Prediksi Prestasi Mahasiswa. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, 5(1), 10-16. <http://dx.doi.org/10.30872/jurti.v5i1.4996>
- Hendrawan, I.N.R., Saputra, I.M.A.B., Dewi, G.A.P.C., Pranata, I.G.S.A., Wedasari, N.L.N., 2021. Klasifikasi Lama Studi dan Predikat Kelulusan Mahasiswa menggunakan Metode Naïve Bayes. *Jurnal Eksplora Informatika*, 11(1), 50–56. <https://doi.org/10.30864/eksplora.v11i1.606>
- Hidayat, N., Iskandar, D., Afuan, L., Nugroho, A.K., Arwiansyah, N., 2022. ANALYSIS AND DESIGN OF DATA WAREHOUSE BASED ON SNIKTI USING DATA WAREHOUSE LIFE CYCLE METHOD AT UNSOED ENGINEERING FACULTY. *Jurnal Teknik Informatika (JUTIF)*, 3(3), 797-805. <https://doi.org/10.20884/1.jutif.2022.3.3.514>
- Muin, A.A., Rahman., Abdillah, A., 2022. Perancangan Sistem Klasifikasi Mahasiswa untuk Prediksi Performa Mahasiswa Menggunakan Naïve Bayes Classifier. *Jurnal INSYPRO (Information System and Processing)*, 7(2). <https://doi.org/10.24252/insypro.v7i2.35247>
- Khasanah, N., Salim, A., Afni, N., Komarudin, R., Maulana, Y.I., 2022. PREDIKSI KELULUSAN MAHASISWA DENGAN METODE NAIVE BAYES. *Technologia: Jurnal Ilmiah*, 13(3), 207-214. <http://dx.doi.org/10.31602/tji.v13i3.7312>
- Novianto, E., Hermawan, A., Avianto, D., 2023. KLASIFIKASI ALGORITMA K-NEAREST NEIGHBOR, NAIVE BAYES, DECISION TREE UNTUK PREDIKSI STATUS KELULUSAN MAHASISWA S1. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 8(2), 146–154. <https://doi.org/10.36341/rabit.v8i2.3434>
- Nurviana, N., Ati, S.K., Hanifah, H.P., 2019. Predictive Model of Passengers Trans Metro Bandung Encouraging Smart Transportation. *Jurnal Sistem Cerdas*, 2(2), 111-118. <https://doi.org/10.37396/jsc.v2i2.30>
- Wahyuni, S., Adinda., 2021. PENERAPAN ALGORITMA NAÏVE BAYES PADA PREDIKSI KELULUSAN MAHASISWA (STUDI KASUS: INSTITUT MEDIKA Drg. SUHERMAN). *Jurnal Informatika SIMANTIK*, 6(2), 29-34. <https://www.simantik.panca-sakti.ac.id/index.php/simantik/article/view/120>
- Wibowo, A., Rohman, A., 2022. Prediksi Predikat Kelulusan Mahasiswa Menggunakan Naive Bayes dan Decision Tree pada Universitas XYZ. *EXPERT: Jurnal Manajemen Sistem Informasi dan Teknologi*, 12(2), 104-112. <https://doi.org/10.36448/expert.v12i2.2810>
- Wibowo, M., Djafar, M.R.F., 2023. Perbandingan Metode Klasifikasi Untuk Deteksi Stress Pada Mahasiswa di Perguruan Tinggi. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(1), 153-159. <https://doi.org/10.30865/mib.v7i1.5182>