

Analisis Perbandingan Indeks Prestasi (IP) Mahasiswa Sebelum dan Saat Pandemi Menggunakan Metode Naïve Bayes

Widia Putri Septiani¹⁾ & Nunik Pratiwi²⁾

^{1,2)}Teknik Informatika Fakultas Teknologi Industri dan Informatika Universitas Muhammadiyah Prof. Dr. Hamka
E-mail: npratiwi@uhamka.ac.id

Abstrak

Terjadinya pandemi covid-19 yang menyebabkan perubahan dalam berbagai aspek termasuk cara belajar yang berubah menjadi pembelajaran secara daring. Naïve bayes merupakan metode klasifikasi secara statistik dapat membuat prediksi di masa depan. Tujuan penelitian ini adalah mengetahui prediksi nilai yang didapatkan naik atau turun dari dataset, serta mengetahui hasil akurasi pada proses pengujian algoritma dan pengujian evaluasi. Algoritma Naïve Bayes dapat melakukan prediksi apakah IP akan naik atau turun pada semester berikutnya. Nilai akurasi pengujian Algoritma Naïve Bayes didapatkan hasil untuk sebelum pandemi 80% dan saat pandemi 83% dan pada proses evaluasi yang dilakukan untuk sebelum pandemi nilai naik mendapatkan nilai precision sebanyak 91%, recall sebanyak 64%, dan f1-score sebanyak 75%. Sedangkan untuk nilai turun mendapatkan nilai precision sebanyak 76%, recall sebanyak 95% dan f1-score sebanyak 85%. Untuk saat pandemi nilai naik mendapatkan nilai precision sebanyak 83%, recall sebanyak 99%, dan f1-score sebanyak 90%. Sedangkan untuk nilai turun mendapatkan nilai precision sebanyak 88%, recall sebanyak 30% dan f1-score sebanyak 45%. Hasil prediksi sebelum pandemi untuk nilai naik berjumlah 34 dan turun berjumlah 71, dan saat pandemi untuk prediksi nilai naik berjumlah 95 dan turun berjumlah 9.

Keyword: indeks prestasi, naïve bayes, prediksi

Abstract

The occurrence of the COVID-19 pandemic has caused changes in various aspects, including the way of learning that has turned into online learning. Naïve Bayes is statistical classification method can make predictions in the future. The purpose of this study is to determine the predicted value obtained up or down from the dataset and to find out the accuracy results in the algorithm testing process and evaluation testing. Naïve Bayes Algorithm can predict whether the GPA will increase or decrease in the in the following semester. The accuracy value of the Naïve Bayes Algorithm test obtained results of 80% before the pandemic and 83% during the pandemic and in the evaluation, process carried out before the pandemic the value rose to 91% precision, 64% recall, and 75% f1-score. As for the decreasing value, the precision value is 76%, the recall is 95% and the f1-score is 85%. During the pandemic, the value went up to get a precision value of 83%, recall of 99%, and an f1 score of 90%. As for the decreasing value, the precision value is 88%, the recall is 30% and the f1-score is 45%. The prediction results before the pandemic for the value to increase amounted to 34 and decreased to 71, and during the pandemic, the predicted value to increase was 95 and decreased by 9.

Kata kunci: gpa, naïve bayes, prediction

1 PENDAHULUAN

Pada perguruan tinggi keberhasilan mahasiswa pada pembelajaran dapat dilihat dari hasil yang diperolehnya yaitu pada nilai Indeks Prestasi (IP). Indeks Prestasi terbagi

dua yaitu Indeks Prestasi Semester (IPS) dan Indeks Prestasi Kumulatif (IPK). Indeks Prestasi Semester adalah nilai akhir yang didapatkan mahasiswa selama satu semester [1], sedangkan Indeks Prestasi Kumulatif (IPK) yaitu penilaian

berdasarkan seluruh mata kuliah dari semester satu sampai semester yang sedang ditempuh [2].

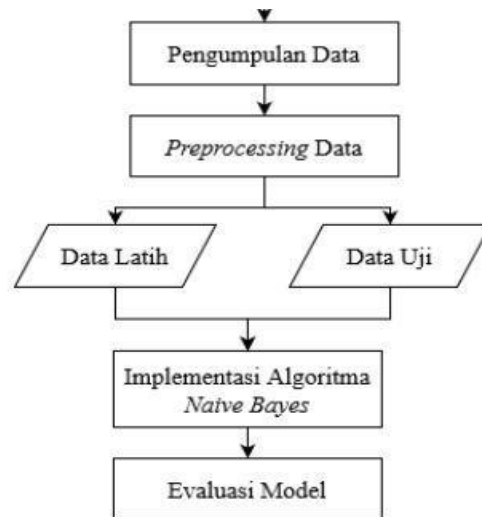
Pandemi *Coronavirus Disease* (Covid-19) yang melanda di belahan dunia, banyak memberikan perubahan dan juga memberikan tantangan terlebih pada pendidikan [3]. Pemerintah memberlakukan banyak aturan untuk mengantisipasi penularan covid-19, salah satunya yaitu *social distancing*, dengan diberlakukannya aturan tersebut berdampak terhadap aktivitas pembelajaran di perguruan tinggi. Pembelajaran daring merupakan kegiatan belajar yang menjadi jalan keluar pada era pandemi covid-19. Pembelajaran daring yaitu kegiatan belajar yang dilakukan secara *online*, menggunakan platform tertentu. Platform yang sering digunakan *Zoom Meeting*, *Google Meet*, *Google Classroom*, *WhatsApp Youtube* dan lainnya atau melalui platform yang sudah disediakan oleh universitas masing-masing.

Pembelajaran di masa pandemi ini yang mempunyai beberapa faktor kendala apakah dapat membuat nilai IP menjadi lebih baik atau sebaliknya. Menganalisis nilai IP Mahasiswa TI UHAMKA ini data yang didapatkan dari kesekretariatan, data yang digunakan berjumlah 1.041 data dari angkatan 2017-2019, untuk sebelum pandemi angkatan 2017 digunakan 136 data mahasiswa, angkatan 2018 digunakan 198 data mahasiswa, dan untuk angkatan 2019 digunakan 187 data mahasiswa, dan saat pandemi angkatan 2017 digunakan 136 data mahasiswa, angkatan 2018 digunakan 198 data mahasiswa, dan untuk Angkatan 2019 digunakan 1876 data mahasiswa serta data yang digunakan yaitu mahasiswa yang mengikuti semua semester yang telah ditentukan tanpa mengambil cuti atau berhenti.

2 METODOLOGI PENELITIAN

Metodologi dalam penelitian ini terdapat 4 (empat) tahapan, yaitu: pengumpulan data, *pre-processing* data, implementasi Naïve Bayes dan evaluasi model. Gambar 1

merepresentasikan tahapan-tahapan dalam penelitian ini.



Gambar 1 Tahapan-tahapan penelitian

1. Pengumpulan Data

Tahap pertama yaitu pengumpulan data yang akan dipakai dalam penelitian. Data didapatkan dari sekretariat FT UHAMKA. Semua data mahasiswa TI angkatan 2017, 2018 dan 2019. Dengan data untuk angkatan 2017 yaitu semester 4-7, angkatan 2018 yaitu semester 2-5 dan angkatan 2019 yaitu semester 1-4. Data tersebut berisi Nama, Nomor Induk Mahasiswa (NIM), SKS Kumulatif, SKS Semester, IPS serta IPK dari angkatan 2017-2019 dengan semester yang berbeda-beda.

2. Pre-processing Data

Data yang sudah didapat selanjutnya terlebih dahulu akan dilangsungkan pengolahan data awal atau data *preprocessing*. Tahap ini memiliki tujuan untuk mendapatkan bentuk data yang sesuai dengan keinginan sebelum masuk ke tahap implementasi. Proses ini akan dilakukan data *reduction*, *cleaning*, *transformation*, data *exploration*. Pada data *reduction* data akan direduksi atau menghapus data yang tidak perlu atau tidak berkaitan, pada data *cleaning* data kosong akan dibersihkan atau dihilangkan, sedangkan pada data *transformation* akan menginisialisasi data dengan angka, pada data *exploration* akan dilakukan mendeklarasikan *library* dan

lainnya, kemudian dilakukan pembagian data atau split data menjadi 80:20.

3. Implementasi Naïve Bayes

Salah satu metode yang dipakai untuk memprediksi probabilitas yaitu metode Naïve Bayes. Algoritma Naïve Bayes adalah metode mendasar tentang teorema Bayesian dan probabilitas dengan cara berpikir bahwa setiap variabel bersifat *independent* dan keberadaan suatu fitur (*variabel*) tidak ada kaitannya dengan fitur lainnya pada data yang sama [4]. Kaitan antara klasifikasi dengan Naïve Bayes, bukti dan korelasi hipotesis dengan klasifikasi yaitu sebuah hipotesis dalam Teorema Bayes adalah sasaran pemetaan dalam klasifikasi dari label kelas, sementara itu bukti adalah beberapa fitur dalam model klasifikasi yang menjadi masukan. Klasifikasi Bayes adalah klasifikasi secara statistik yang dapat memperkirakan peluang [5].

Gaussian *Naïve Bayes* digunakan untuk menghitung fitur yang mempunyai nilai kontinu. Model ini merupakan model klasifikasi dengan anggapan bahwa kedua fitur tidak saling bergantung. Persamaan umum Gaussian *Naïve Bayes* sebagai berikut:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right) \quad (1)$$

Tahap ini akan melakukan implementasi Algoritma *Naïve Bayes* menggunakan model klasifikasi Gaussian *Naïve Bayes*. Pada model ini data akan dibagi menjadi kedalam dua proses yaitu proses latihan (*training*) yang bertujuan untuk melatih algoritma dan proses uji (*testing*) yang bertujuan untuk mengetahui akurasi dari data yang telah dilatih dengan algoritma tersebut.

4. Evaluasi Model

Setelah semua proses selesai, dimulai dari pengumpulan data sampai mendapatkan hasil dari tahap implementasi Algoritma, tahap terakhir yang akan dilakukan yaitu

melakukan evaluasi model. Evaluasi ini akan dibagi menjadi 4 (empat) proses, yaitu: *accuracy*, *precision*, *recall*, dan *f-measure*. Persamaan umum ke empat metode evaluasi sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

3 HASIL DAN PEMBAHASAN

Bagian ini akan menjelaskan hasil eksperimen yang dilakukan dalam penelitian ini. Sesuai dengan tahapan penelitian maka pembahasan hasil penelitian sebagai berikut:

1. Hasil Pre-processing Data

Preprocessing data merupakan tahap yang digunakan agar mendapatkan bentuk data sesuai yang diinginkan. Data yang telah didapatkan masih dalam keadaan mentah dan mengandung *noise* sehingga data tersebut masih belum dapat digunakan. Pada tahap ini, yaitu merekap data mahasiswa prodi Informatika angkatan 2017-2019 dan *dataset* yang digunakan yaitu data mahasiswa yang mengikuti perkuliahan pada semester yang telah ditentukan (tidak sedang cuti dan keluar atau pindah) dan data tersebut disimpan pada file *excel* yang menghimpun nama, jenis kelamin, Nomor Induk Mahasiswa, angkatan, SKS semester, SKS kumulatif, IPS, keterangan (sebelum pandemi dan saat pandemi) dan status (naik atau turun).

a. *Reduction* Data

Tahap *preprocessing* pertama yaitu *reduction* data, data tersebut akan mengalami proses pengurangan atau mereduksi sejumlah data yang tidak dibutuhkan atau tidak berhubungan. Dalam penelitian ini atribut yang akan digunakan yaitu jenis kelamin, angkatan, SKS_S1, SKS_S2, SKS_Kum_1, SKS_Kum2, IPS1, IPS2, keterangan dan status

b. *Cleaning Data*

Tahap *cleaning* data, data akan di filter setiap baris serta kolom dan akan mencari data yang kosong, data kosong tersebut akan dihapus karena data yang diambil data mahasiswa yang mengikuti perkuliahan pada semua semester yang telah ditentukan. Data sebelum dilakukan proses *cleaning* dapat dilihat pada Gambar 2.

J	A	B	C	D	E	F	G	H	I	J
129	Perempuan	2017	14	18	102	73	2.36	2.33	Turun	Sebelum Pandemi
130	Perempuan	2017	18	21	119	76	2.56	3.43	Turun	Sebelum Pandemi
131	Laki-Laki	2017	21	17	113	80	2.67	2.65	Turun	Sebelum Pandemi
132	Laki-Laki	2017	18	21	107	71	2.50	2.14	Turun	Sebelum Pandemi
133	Laki-Laki	2017	24	18	112	79	2.00	2.56	Naik	Sebelum Pandemi
134	Perempuan	2017	21	18	112	80	2.38	2.28	Turun	Sebelum Pandemi
135	Perempuan	2017	21	16	110	79	2.14	2.81	Naik	Sebelum Pandemi
136	Perempuan	2017	18	15	73	58		2.73		Sebelum Pandemi
137	Laki-Laki	2017	22	17	124	86	2.41	2.71	Naik	Sebelum Pandemi
138	Laki-Laki	2018	20	21	72	72	2.65	2.85	Naik	Sebelum Pandemi
139	Laki-Laki	2018	20	18	81		2.80			Sebelum Pandemi
140	Laki-Laki	2018	24	24	85	85	3.45	3.25	Turun	Sebelum Pandemi
141	Laki-Laki	2018	24	24	91	91	3.50	3.88	Naik	Sebelum Pandemi
142	Laki-Laki	2018	24	20	91	91	3.67	2.50	Turun	Sebelum Pandemi
143	Laki-Laki	2018	18	21	73	73	2.06	2.33	Naik	Sebelum Pandemi
144	Laki-Laki	2018	18	18	70	70	2.39	2.11	Turun	Sebelum Pandemi
145	Perempuan	2018	24	22	93	93	3.25	3.75	Naik	Sebelum Pandemi
146	Laki-Laki	2018	24	19	89	89	3.00	3.13	Naik	Sebelum Pandemi
147	Laki-Laki	2018	11	15	53	53	0.27	1.64	Naik	Sebelum Pandemi
148	Perempuan	2018	24	22	91	91	3.46	3.92	Naik	Sebelum Pandemi
149	Laki-Laki	2018	22	23	89	89	4.00	3.77	Turun	Sebelum Pandemi
150	Laki-Laki	2018	21	23	89	80	2.86	3.29	Naik	Sebelum Pandemi
151	Laki-Laki	2018	23	24	86	86	3.10	3.52	Naik	Sebelum Pandemi

Gambar 2 Data sebelum proses cleaning

Data setelah melalui proses *cleaning* berkurang dari setiap angkataannya, data sebelum pandemi sebelumnya berjumlah 531 menjadi 521 dan saat pandemi sebelumnya berjumlah 530 menjadi 520. Data setelah mengalami proses *cleaning* dapat dilihat pada Gambar 3.

J	A	B	C	D	E	F	G	H	I	J
129	Perempuan	2017	14	18	102	73	2.36	2.33	Turun	Sebelum Pandemi
130	Perempuan	2017	18	21	119	76	2.56	3.43	Turun	Sebelum Pandemi
131	Laki-Laki	2017	21	17	113	80	2.67	2.65	Turun	Sebelum Pandemi
132	Laki-Laki	2017	18	21	107	71	2.50	2.14	Turun	Sebelum Pandemi
133	Laki-Laki	2017	24	18	112	79	2.00	2.56	Naik	Sebelum Pandemi
134	Perempuan	2017	21	18	112	80	2.38	2.28	Turun	Sebelum Pandemi
135	Perempuan	2017	21	16	110	79	2.14	2.81	Naik	Sebelum Pandemi
136	Laki-Laki	2017	22	17	124	86	2.41	2.71	Naik	Sebelum Pandemi
137	Laki-Laki	2018	20	21	72	72	2.65	2.85	Naik	Sebelum Pandemi
138	Laki-Laki	2018	24	24	85	85	3.45	3.25	Turun	Sebelum Pandemi
139	Laki-Laki	2018	24	24	91	91	3.50	3.88	Naik	Sebelum Pandemi
140	Laki-Laki	2018	24	20	91	91	3.67	2.50	Turun	Sebelum Pandemi
141	Laki-Laki	2018	18	21	73	73	2.06	2.33	Naik	Sebelum Pandemi
142	Laki-Laki	2018	18	18	70	70	2.39	2.11	Turun	Sebelum Pandemi
143	Perempuan	2018	24	22	93	93	3.25	3.75	Naik	Sebelum Pandemi
144	Laki-Laki	2018	24	19	89	89	3.00	3.13	Naik	Sebelum Pandemi
145	Laki-Laki	2018	11	15	53	53	0.27	1.64	Naik	Sebelum Pandemi
146	Perempuan	2018	24	22	91	91	3.46	3.92	Naik	Sebelum Pandemi
147	Laki-Laki	2018	22	23	89	89	4.00	3.77	Turun	Sebelum Pandemi
148	Laki-Laki	2018	21	23	89	80	2.86	3.29	Naik	Sebelum Pandemi
149	Laki-Laki	2018	23	24	86	86	3.10	3.52	Naik	Sebelum Pandemi
150	Laki-Laki	2018	24	22	91	91	3.25	3.50	Naik	Sebelum Pandemi
151	Laki-Laki	2018	17	20	76	76	2.10	2.94	Naik	Sebelum Pandemi

Gambar 3 Data setelah proses cleaning

c. *Transformation Data*

Tahap *transformation* data, atribut jenis kelamin dan keterangan adalah data yang

berupa text, maka data tersebut akan diubah kedalam bentuk angka. Jenis kelamin laki-laki diberi inisial (1) dan perempuan diberi inisial (2). Keterangan sebelum pandemi diberi inisial (1) dan pandemi diberi inisial (2). Data setelah dilakukan proses *transformation* terlihat pada Gambar 4 dan 5.

A	B	C	D	E	F	G	H	I	J
Jenis_Kelamin	Angkatan	SKS_S1	SKS_S2	SKS_Kumul	SKS_Kum2	IPS1	IPS2	Status	Keterangan
1	2017	11	15	70	100	1.91	1.93	Naik	1
2	2017	21	19	82	123	3.00	3.53	Naik	1
1	2017	20	16	70	96	2.00	2.13	Naik	1
1	2017	18	17	60	90	2.39	1.53	Turun	1
1	2017	15	12	65	85	2.53	1.25	Turun	1
1	2017	21	18	79	121	2.38	3.67	Naik	1
2	2017	18	11	74	90	0.50	0.64	Naik	1
1	2017	20	15	77	109	1.70	2.20	Naik	1
1	2017	21	18	74	112	2.29	2.06	Turun	1
1	2017	18	12	80	108	0.83	2.17	Naik	1
2	2017	21	21	71	112	2.62	2.43	Turun	1
1	2017	21	21	81	123	3.05	2.67	Turun	1
1	2017	21	21	75	117	2.76	3.29	Naik	1
2	2017	24	20	85	119	2.75	2.60	Turun	1
1	2017	24	23	85	126	3.46	3.48	Naik	1
1	2017	23	19	91	131	3.78	3.84	Naik	1
2	2017	22	19	87	126	3.18	3.47	Naik	1
1	2017	21	23	85	127	3.38	3.43	Naik	1
1	2017	18	17	60	89	2.06	1.76	Turun	1
1	2017	20	21	86	122	2.55	2.52	Turun	1
1	2017	21	18	83	120	2.43	3.06	Naik	1
1	2017	20	18	83	121	3.05	2.89	Turun	1

Gambar 4 Data setelah proses transformation (sebelum pandemi)

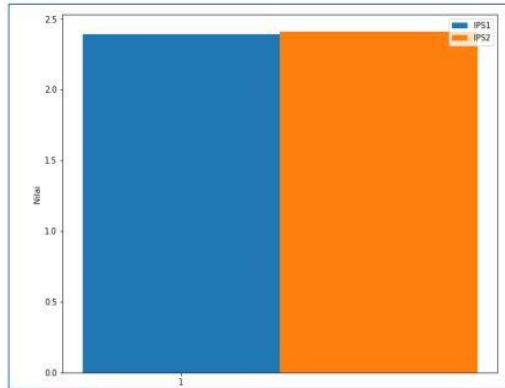
A	B	C	D	E	F	G	H	I	J
Jenis_Kelamin	Angkatan	SKS_S1	SKS_S2	SKS_Kumul	SKS_Kum2	IPS1	IPS2	Status	Keterangan
1	2017	15	21	100	138	2.53	3.14	Naik	2
2	2017	22	16	123	153	3.86	3.31	Turun	2
1	2017	18	20	96	131	2.67	3.20	Naik	2
1	2017	15	15	90	90	0.53	0.53	Turun	2
1	2017	8	10	85	85	1.00	1.00	Turun	2
1	2017	23	22	121	154	3.57	3.64	Naik	2
2	2017	11	15	90	112	2.00	2.67	Naik	2
1	2017	18	19	109	147	2.67	3.53	Naik	2
1	2017	19	21	112	146	2.58	3.52	Naik	2
1	2017	18	24	108	144	3.17	2.67	Turun	2
2	2017	18	21	112	142	2.94	2.90	Turun	2
1	2017	21	21	123	151	2.52	2.71	Naik	2
1	2017	21	23	117	148	3.14	3.52	Naik	2
2	2017	20	24	119	147	3.30	1.83	Turun	2
1	2017	18	13	126	153	3.33	3.77	Naik	2
1	2017	21	11	131	153	3.67	4.00	Naik	2
2	2017	20	18	126	153	3.45	3.61	Naik	2
1	2017	21	15	127	153	3.38	3.67	Naik	2
1	2017	14	18	89	114	2.29	2.61	Naik	2
1	2017	21	20	122	139	3.14	3.60	Naik	2
1	2017	24	21	120	154	3.50	3.62	Naik	2
1	2017	20	18	121	153	3.45	3.61	Naik	2

Gambar 5 Data setelah proses transformation (pandemi)

d. *Exploration Data*

Dataset terdiri jenis kelamin, angkatan, SKS 1, SKS 2, SKS Kumulatif 1, SKS Kumulatif 2, IPS1, IPS2, status dan keterangan. Data yang digunakan berjumlah 521 data sebelum pandemi dan 520 data saat pandemi dari angkatan 2017-2019. Untuk sebelum pandemi angkatan 2017 digunakan 136 data mahasiswa, angkatan 2018 digunakan 198 data mahasiswa, dan untuk Angkatan 2019 digunakan 187 data mahasiswa serta saat pandemi angkatan 2017 digunakan 136 data mahasiswa, angkatan 2018 digunakan 198 data mahasiswa, dan untuk angkatan 2019 digunakan 186 data mahasiswa.

Tahapan selanjutnya yaitu membandingkan nilai Indeks Prestasi (IP), dengan membandingkan dua nilai IPS di sebelum pandemi dan dua nilai IPS di saat pandemi, yaitu dengan melihat rata-rata nilai yang dikelompokkan menurut keterangan. Visualisasi perbandingan nilai IPS terdapat pada Gambar 6 dan 7.



Gambar 6 Visualisasi perbandingan rata-rata IPS sebelum pandemi



Gambar 7 Visualisasi perbandingan rata-rata IPS saat pandemi

2. Hasil Implementasi Naïve Bayes
 Tahap selanjutnya yaitu implementasi Algoritma terhadap *dataset*. Pada proses ini Algoritma yang digunakan yaitu model Gaussian Naïve Bayes. Penggunaan algoritma yaitu dengan melakukannya pada data *testing* dan data *training* yang telah dideklarasikan pada tahap sebelumnya. *Output* yang dihasilkan setelah melakukan proses pengujian pada data *training* dan data *testing* yaitu hasil prediksi Algoritma. Hasil yang menjadi *output* ini penting karena merupakan indikasi bahwa Algoritma yang digunakan sudah tepat atau belum dengan menguji juga hasil akurasi.

Setelah dilakukan proses mendeklarasikan Algoritma tersebut, selanjutnya yaitu tahapan prediksi data secara keseluruhan dengan metode Gaussian Naïve Bayes. Dari pembagian data 80:20, data *testing* berjumlah 105 (sebelum pandemi) dan 104 (saat pandemi).

```
Prediksi Naive Bayes : ['Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun' 'Turun' 'Turun'
'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun' 'Naik' 'Naik' 'Turun'
'Naik' 'Naik' 'Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik'
'Turun' 'Turun' 'Turun' 'Naik' 'Turun' 'Turun' 'Naik' 'Turun' 'Turun'
'Naik' 'Turun' 'Turun' 'Naik' 'Naik' 'Turun' 'Turun' 'Naik' 'Naik'
'Turun' 'Naik' 'Turun' 'Naik' 'Turun' 'Naik' 'Turun' 'Turun' 'Turun'
'Turun' 'Naik' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun' 'Naik'
'Naik' 'Turun' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Turun' 'Turun'
'Turun' 'Naik' 'Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun'
'Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun'
'Naik' 'Turun' 'Turun' 'Turun' 'Turun' 'Naik' 'Turun' 'Turun'
'Turun' 'Turun' 'Naik' 'Turun' 'Turun' 'Naik']
```

Gambar 8 Hasil prediksi naïve bayes pada data sebelum pandemi

Berdasarkan hasil prediksi metode Gaussian Naïve Bayes pada 105 data testing sebelum pandemi terdapat 34 data yang diprediksi akan naik dan 71 data yang diprediksi akan turun pada semester berikutnya.

```
Prediksi Naive Bayes : ['Naik' 'Naik' 'Naik' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Naik' 'Turun' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik'
'Naik' 'Naik' 'Turun' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik' 'Naik'
'Turun' 'Naik' 'Naik' 'Naik']
```

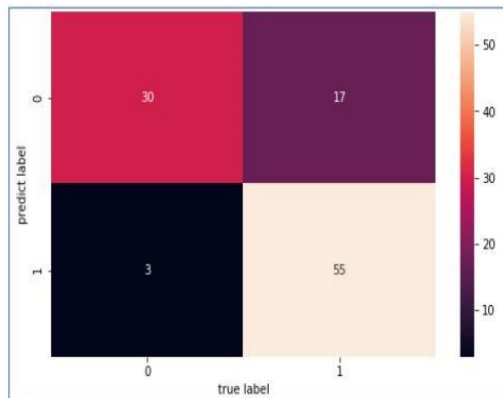
Gambar 9 Hasil prediksi naïve bayes pada data saat pandemi

Berdasarkan hasil prediksi metode Gaussian Naïve Bayes pada 104 data testing saat pandemi terdapat 95 data yang diprediksi akan naik dan 9 data yang diprediksi akan turun pada semester berikutnya.

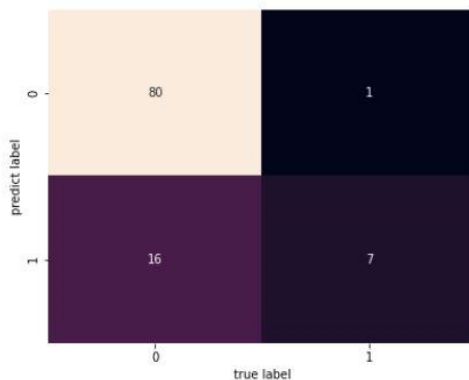
3. Hasil Evaluasi Model

Setelah semua proses telah dilakukan maka bagian ini terakhir dalam rangkaian implementasi algoritma, yaitu evaluasi. Pada proses ini akan didapatkan hasil *accuracy*, *precision*, *recall*, dan *f-measure* dari implementasi algoritma Naïve Bayes. Berdasarkan hasil training dan testing pada dataset dengan komposisi pembagian data 80:20 yang mempunyai nilai akurasi yang paling baik yaitu sebesar 80,95% (sebelum pandemi) dan 83,65% (pandemi). Evaluasi yang lain untuk data sebelum pandemi prediksi IPS “Naik” mendapatkan nilai *precision* sebanyak 91%, *recall* sebanyak 64%, dan *f1-score* sebanyak 75%. Sedangkan untuk prediksi IPS “Turun” mendapatkan nilai *precision* sebanyak 76%,

recall sebanyak 95% dan *f1-score* sebanyak 85%. Untuk saat pandemi prediksi IPS “Naik” mendapatkan nilai *precision* sebanyak 83%, *recall* sebanyak 99%, dan *f1-score* sebanyak 90%. Sedangkan untuk prediksi IPS “Turun” mendapatkan nilai *precision* sebanyak 88%, *recall* sebanyak 30% dan *f1-score* sebanyak 45%. *Confusion matrix* untuk kedua dataset dapat dilihat pada Gambar 10 dan 11.



Gambar 10 Visualisasi confusion matrix data sebelum pandemi



Gambar 11 Visualisasi confusion matrix data saat pandemi

4 SIMPULAN

Berdasarkan hasil yang didapatkan dalam penelitian ini dapat diambil simpulan sebagai berikut:

1. Setelah dilakukan proses *preprocessing*, pada tahap *exploration* data, didapatkan bahwa Indeks Prestasi Semester (IPS) di dua semester pada saat pandemi lebih tinggi dibandingkan Indeks Prestasi Semester (IPS) sebelum pandemi. Hasil prediksi sebelum pandemi untuk IPS

“Naik” berjumlah 34 dan IPS “Turun” berjumlah 71, dan saat pandemi untuk prediksi IPS “Naik” berjumlah 95 dan IPS “Turun” berjumlah 9.

2. Prediksi IPS naik atau turun pada semester berikutnya menggunakan metode *Naïve Bayes* dapat dikatakan melakukan prediksi dengan baik. Hasil akurasi pada pengujian Algoritma dari metode *Naïve Bayes* didapatkan hasil untuk sebelum pandemi 80,95% dan saat pandemi 83,65%.

KEPUSTAKAAN

- [1] Anna Maria; Hartini H. (2020). Pengaruh Kompetensi Pedagogik dan Profesional Dosen Terhadap Indeks Prestasi Mahasiswa Semester II AKJP II Pekanbaru. *Jurnal HUMMANSI* 3(1), 15–22. <https://journal.stikomys.ac.id/index.php/jurnalhummani/article/download/244/136>.
- [2] Susilo, U., & Arifin, M. (2020). Analisis Hubungan Indeks Prestasi Semester Dan Indeks Prestasi Kumulatif Dengan Prestasi Mahasiswa Fakultas Ekonomi Universitas Kadiri. *Jurnal Riset Bisnis Dan Ekonomi*, 1(1), 12–22. <http://ojs.unik-kediri.ac.id/index.php/jimek>
- [3] Jamaluddin, D., Ratnasih, T., Gunawan, H., & Paujiah, E. (2020). Pembelajaran Daring Masa Pandemi Covid-19 Pada Calon Guru : Hambatan, Solusi dan Proyeksi. *Karya Tulis Ilmiah UIN Sunan Gunung Djati Bandung*, 1–10. <http://digilib.uinsgd.ac.id/30518/>.
- [4] Sari, Z., & Sarosa, M. (2018). “Si Tole” *Chatterbot untuk Melatih Rasa Percaya Diri Menggunakan Naive Bayes Classification*. 7(1), 64–71.
- [5] Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>