

## Analisis Prediksi Tingkat Penyebaran Covid-19 di Sulawesi Selatan menggunakan Teknik Data Mining Naïve Bayes

Muhammad Nur Yusri<sup>1</sup>, Andi Akram Nur Risal<sup>2</sup>, Muhammad Fajar B<sup>3</sup>,  
Dewi Fatmarani Surianto<sup>4</sup>, Fhatiah Adiba<sup>5</sup>

<sup>1</sup>1829141023@student.unm.ac.id, <sup>2</sup>akramandi@unm.ac.id, <sup>3</sup>fajarb@unm.ac.id,

<sup>4</sup>dewifatmaranis@unm.ac.id, <sup>5</sup>adibafhatiah@unm.ac.id

<sup>1,2,3,4,5</sup>Universitas Negeri Makassar

Received : 27 Aug 22  
Accepted : 18 Nov 22  
Published : 27 Nov 22

### Abstract

**Abstract:** The pandemic or corona virus outbreak or commonly referred as COVID-19 which started in Wuhan, Hubei Province, China, continues to spread to various countries, including Indonesia. The number of positive cases of COVID-19 continues to increase significantly every day and spread rapidly to various provinces in Indonesia, including in the province of South Sulawesi. To date, 18,683 positive cases of corona have been recorded in South Sulawesi and 470 people have died. This significant increase in cases has resulted in the reading of data related to positive cases of COVID-19 in South Sulawesi being considered inaccurate. Therefore, this research was conducted as a precautionary measure against the COVID-19 pandemic by predicting the level of spread of COVID-19, especially in South Sulawesi in order to obtain better data accuracy. The research method applied in this research is problem analysis and literature study, data collection and implementation.

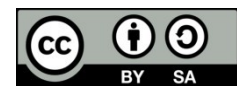
**Keywords:** COVID-19, South Sulawesi, Naïve Bayes

### Abstrak

Pandemi atau wabah virus corona atau biasa disebut juga dengan COVID-19 yang bermula dari Wuhan, Provinsi Hubei, China, terus menyebar ke berbagai negara, termasuk Indonesia. Jumlah kasus positif COVID-19 terus meningkat tiap harinya secara signifikan dan menyebar secara cepat ke berbagai provinsi di Indonesia, termasuk di provinsi Sulawesi Selatan. Hingga saat ini, telah tercatat kasus positif corona di Sulawesi Selatan sebanyak 18.683 dan 470 orang meninggal dunia. Peningkatan kasus yang signifikan ini, mengakibatkan pembacaan data terkait kasus positif COVID-19 di Sulawesi Selatan dinilai kurang akurat. Oleh karena itu, penelitian ini dilakukan sebagai langkah antisipasi terhadap pandemi COVID-19 dengan memprediksi tingkat penyebaran COVID-19 terutama di Sulawesi Selatan agar mendapatkan keakurasian data yang lebih baik. Metode penelitian yang di terapkan pada penelitian ini ialah analisis masalah dan studi literatur, mengumpulkan data dan implementasi.

**Kata Kunci:** COVID-19, Sulawesi Selatan, Naïve Bayes

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license*



## 1. Pendahuluan

Saat ini, dunia tengah dilanda oleh wabah Virus Corona atau yang biasa disebut dengan COVID-19. Virus corona merupakan bagian dari keluarga virus yang menyebabkan penyakit mulai dari flu hingga penyakit yang lebih berat seperti Middle East Respiratory Syndrome (MERS-CoV) dan Severe Acute Respiratory Syndrome (SARS-CoV). Penyakit yang disebabkan virus corona, atau dikenal dengan COVID-19, adalah jenis baru yang ditemukan pada tahun 2019 dan belum pernah diidentifikasi menyerang manusia sebelumnya. Kasus virus corona muncul dan menyerang manusia pertama kali di provinsi Wuhan, China. Awal kemunculannya diduga merupakan penyakit pneumonia, dengan gejala serupa dengan sakit flu pada umumnya

Karena penularan virus corona yang sangat cepat inilah, Organisasi Kesehatan Dunia atau World Health Organization (WHO) menetapkan virus corona sebagai pandemi global pada tanggal 11 Maret 2020. Dari status pandemik atau epidemik global ini, menandakan bahwa penyebaran virus ini sangatlah cepat, yang berdampak hampir keseluruhan negara di Dunia. Dalam artian, penyebaran virus corona melalui antar negara sudah tidak dapat dihindari atau dicegah lagi.

COVID-19 telah menyebar ke lebih dari 210 Negara di seluruh dunia, termasuk Indonesia. Hingga saat ini, total kasus positif corona di Indonesia telah mencapai 433.836 kasus dan sebanyak 18.440 meninggal dunia, dan masih terus bertambah tiap harinya. Di dalam itu, Provinsi Sulawesi Selatan telah tercatat dengan total kasus positif corona sebanyak 18.683 dan 470 orang meninggal dunia.

Berdasarkan kondisi tersebut, maka Indonesia, termasuk Sulawesi Selatan telah berada dalam status waspada terhadap ancaman virus corona. Cara pencegahan virus corona yang paling efektif adalah dengan menggunakan vaksin. Namun, vaksin corona hingga saat ini masih dalam tahap penelitian dan pengembangan. Hingga vaksin ditemukan, maka bentuk pencegahan dari meluasnya penyebaran virus dapat dilakukan dengan cara memutuskan rantai penularannya. Metode yang dapat dilakukan sesuai anjuran pemerintah saat ini adalah dengan menerapkan protokol kesehatan, yang diantaranya adalah memakai masker, mencuci tangan, dan menjaga jarak minimal 1 meter.

Mengantisipasi dan mengurangi jumlah penderita virus corona di Indonesia sudah dilakukan di seluruh daerah. Diantaranya dengan memberikan kebijakan membatasi aktifitas keluar rumah, kegiatan sekolah dirumahkan, bekerja dari rumah (work from home), bahkan kegiatan beribadah pun dirumahkan. Hal ini sudah menjadi kebijakan pemerintah berdasarkan pertimbangan-pertimbangan yang sudah dianalisa dengan maksimal tentunya. Namun, kasus positif yang tercatat masih terus bertambah sampai sekarang. Masih kurangnya penelitian terkait data kasus COVID-19, terutama dalam memprediksi data menjadi salah satu penyebabnya yang dapat berpengaruh pada keakurasian data.

Pada penelitian ini, dilakukan penerapan data mining untuk memprediksi tingkat penyebaran COVID-19 di Sulawesi Selatan dengan mengimplementasikan Algoritma Naïve Bayes, yang berguna sebagai langkah antisipasi pandemik COVID-19 di Sulawesi Selatan.

Penelitian ini dilakukan sebagai langkah antisipasi terhadap pandemi COVID-19 dengan memprediksi tingkat penyebaran COVID-19 terutama di Sulawesi Selatan. Metode penelitian yang diterapkan pada penelitian ini ialah analisis masalah dan studi literatur, mengumpulkan data dan implementasi.

Diharapkan, penelitian ini dapat membantu memprediksi data kasus positif COVID-19 di provinsi Sulawesi Selatan dengan akurat dan tepat sehingga dapat menghasilkan keakurasian data yang lebih tinggi.

## 2. Metode

### 2.1. Analisis Masalah dan Studi Literatur

Pada tahap ini merupakan langkah awal untuk menentukan rumusan masalah dari penelitian. Dalam hal ini mengamati permasalahan yang berhubungan dengan tingkat penyebaran COVID-19 yang terjadi khususnya di Sulawesi Selatan. Permasalahan-permasalahan yang ada, selanjutnya dianalisa untuk mengetahui bagaimana cara penyelesaian terhadap masalah tersebut dan menentukan ruang lingkup permasalahan yang akan diteliti. Mempelajari dan memahami dasar-dasar teori dari berbagai literatur mengenai tingkat penyebaran COVID-19 di Sulawesi Selatan, konsep dan tahapan data mining, dan penerapan algoritma Naïve Bayes. Literatur didapatkan melalui berbagai jurnal yang tersedia, agar mendapatkan dasar pengetahuan dari penelitian

sebelumnya, yang selanjutnya akan ditingkatkan pada penelitian ini.

## 2.2. Pengumpulan Data

Proses pengumpulan data dilakukan dengan menggunakan metode penelitian kuantitatif. Pelaksanaan metode penelitian kuantitatif fokus pada penggunaan angka, tabel, grafik dan diagram untuk menampilkan hasil data yang diperoleh dan akan digunakan untuk bahan penganalisisan data terhadap metode Naïve Bayes. Pengumpulan data didapatkan dari situs resmi penanggulangan COVID-19 di Sulawesi Selatan (<https://covid19.sulselprov.go.id/data>). Setelah data dikumpulkan maka dilakukan analisa data untuk menyesuaikan proses data yang akan diolah ke dalam metode Naïve Bayes.

## 2.3. Implementasi dan Pengujian

Tahap implementasi adalah tentang bagaimana pengolahan datanya diterapkan dalam sebuah tools atau aplikasi. Pengujian dilakukan menggunakan bahasa pemrograman Python dan tools Microsoft Excel 2019. Pengujian dilakukan untuk mengetahui apakah penelitian yang dilakukan telah sesuai dengan tujuan yang diharapkan yaitu untuk memprediksi tingkat penyebaran COVID-19 di Sulawesi Selatan.

## 3. Hasil dan Pembahasan

### 3.1. Perhitungan Metode Naïve Bayes

Perhitungan dilakukan menggunakan tools Microsoft Excel 2019 dengan mengikuti alur atau tahapan dari metode Naïve Bayes, yaitu membaca data training (data latih), menghitung nilai mean dan standar deviasi, menghitung nilai probabilitas, dan menghitung nilai gaussian (data uji).

#### a. Data Training (Data Latih)

Tabel 1. Data Latih

N o	Kabupaten/K ota	Diraw at	Sembu h	Meningg al	Zona
1	Kota Makassar	606	8958	303	Orang e
2	Kabupaten Gowa	119	1438	30	Orang e

N o	Kabupaten/K ota	Diraw at	Sembu h	Meningg al	Zona
3	Kabupaten Luwu Timur	62	1519	4	Orang e
4	Kabupaten Jeneponto	56	530	3	Orang e
5	Kabupaten Maros	45	597	8	Orang e
6	Kabupaten Pinrang	13	183	6	Kunin g
7	Kabupaten Bulukumba	20	324	6	Orang e
8	Kabupaten Sinjai	29	435	0	Orang e
9	Kabupaten Luwu Utara	56	340	13	Kunin g
10	Kabupaten Bone	117	229	4	Kunin g
11	Kota Parepare	33	303	8	Orang e
12	Kota Palopo	76	294	9	Kunin g
13	Kabupaten Enrekang	28	145	10	Orang e
14	Kabupaten Takalar	21	306	1	Orang e
15	Kabupaten Pangkajene dan Kepulauan	20	194	7	Orang e
16	Kabupaten Bantaeng	32	233	4	Orang e
17	Kabupaten Tana Toraja	1	113	2	Kunin g
18	Kabupaten Luwu	21	106	5	Orang e
19	Kabupaten Soppeng	21	210	6	Orang e
20	Kabupaten Sidenreng Rappang	15	161	3	Orang e
21	Kabupaten Wajo	56	151	4	Kunin g
22	Kabupaten Kepulauan Selayar	3	168	0	Kunin g
23	Kabupaten Toraja Utara	12	33	2	Orang e
24	Kabupaten Barru	6	101	1	Orang e

Keterangan :

- Zona “Orange” = Kasus positif lebih tinggi  
Zona “Kuning” = Kasus positif lebih rendah

### b. Menghitung nilai Mean dan Standar Deviasi

Setelah membaca data *training* atau data latih, Langkah selanjutnya adalah mencari nilai *mean* dan *standar deviasi* dari masing – masing parameter. Hal ini dilakukan karena data latih berisi data numerik.

#### 1) Menghitung nilai Mean

*Mean* merupakan nilai rata-rata yang didapatkan dari hasil penjumlahan seluruh nilai dari masing-masing data, lalu dibagi dengan banyaknya data yang ada. Hasil dari perhitungan nilai *mean* dari data latih adalah sebagai berikut :

Tabel 2. Nilai Mean

Mean			
Zona	Dirawat	Sembuh	Meninggal
Orange	61,16666667	711,291667	18,29166667
Kuning	33,05882353	229,117647	5,176470588

Berdasarkan tabel 2 diatas, maka proses untuk menghitung nilai *mean* pada data adalah:

1. Pada zona “orange” terdapat sebanyak 17 data (per provinsi). Dengan menjumlahkan seluruh isi parameter pada masing-masing data, maka hasil dari nilai *mean* yang didapatkan pada parameter “dirawat” adalah 61.1667, parameter “sembuh” adalah 711.29, dan parameter “meninggal” adalah 18.29167.
2. Pada zona “kuning” terdapat sebanyak 7 data (per provinsi). Dengan menjumlahkan seluruh isi parameter pada masing-masing data, maka hasil dari nilai *mean* yang didapatkan pada parameter “dirawat” adalah 33,0588, parameter “sembuh” adalah 229.12, dan parameter “meninggal” adalah 5.176471.

#### 2) Menghitung standar deviasi

*Standar deviasi* atau simpangan adalah ukuran yang digunakan untuk mengukur jumlah variasi atau sebaran sejumlah nilai data. Semakin rendah nilai standar deviasi, maka semakin mendekati rata-rata, sedangkan jika nilai standar deviasi semakin tinggi maka semakin lebar rentang variasi datanya. Sehingga standar deviasi merupakan besar perbedaan dari nilai sampel terhadap rata-rata.

Tabel 3. Standar Deviasi

Standar Deviasi			
Zona	Dirawat	Sembuh	Meninggal
Orange	120,2879396	1795,55878	60,94579143
Kuning	28,96435436	91,2413574	3,557304002

Berdasarkan tabel 3 diatas, maka proses untuk menghitung *standar deviasi* pada data adalah:

1. Pada zona “orange” terdapat *standar deviasi* dengan parameter “dirawat” sekitar 120.288, parameter “sembuh” sekitar 1795.6, dan parameter “meninggal” sekitar 60.94579.
2. Pada zona “kuning” terdapat *standar deviasi* dengan parameter “dirawat” sekitar 28.9644, parameter “sembuh” sekitar 91.241, dan parameter “meninggal” sekitar 3.557304.

### c. Nilai Probabilitas

Probabilitas merupakan suatu nilai yang digunakan untuk mengukur tingkat terjadinya suatu kejadian yang acak. Kata probabilitas itu sendiri sering disebut dengan peluang, kemungkinan, maupun prediksi. Probabilitas secara umum merupakan peluang bahwa sesuatu akan terjadi. Sesuai dengan judul dari penelitian ini, probabilitas digunakan untuk mencari prediksi tingkat penyebaran COVID-19 di Sulawesi Selatan.

Tabel 4. Probabilitas

Probabilitas	
Zona	Nilai
Orange	0,708333333
Kuning	0,291666667
Jumlah	1

Berdasarkan hasil pada tabel 4 diatas, dapat diketahui bahwa probabilitas tertinggi pada kategori zona penyebaran COVID-19 di Sulawesi Selatan adalah zona orange, atau kasus positif per provinsi lebih besar sekitar 0.70833, dan probabilitas rendah pada kategori zona penyebaran COVID-19 di Sulawesi Selatan adalah zona kuning dengan nilai probabilitas sebanyak 0.29167. Probabilitas jika dijumlahkan harus bernilai 1 karena nilai probabilitas yang tetap adalah 1.

#### d. Menghitung Nilai Gaussian (Data Uji)

*Gaussian* atau Distribusi *Gaussian* merupakan langkah terakhir untuk mengetahui hasil dari data *training*, atau sebuah model uji data dengan mengambil nilai peluang dari data *training*.

Tabel 5. Data Uji

Data Uji (Kabupaten /Kota Lain)	Data Uji			
	Dirawat	Sembuh	Meninggal	Zona
	8	80	3	???
Orange	0,0329 97979	0,0088 5275	0,04953 1111	0.00001 0249
Kuning	0,0509 97966	0,0109 8811	0,17545 804	0,00002 86772
Max			0,000028677	2

Berdasarkan hasil data uji pada tabel 5 diatas, dapat disimpulkan bahwa data uji di kabupaten/kota lain dengan parameter “dirawat” adalah 8 memiliki nilai *Gaussian* 0.033 dan 0.051, parameter “sembuh” adalah 80 memiliki nilai *Gaussian* 0,0089 dan 0.011, dan parameter “meninggal” adalah 3 memiliki nilai *Gaussian* 0.049531 dan 0.175458. Dengan menggunakan algoritma *Naïve Bayes*, maka hasil prediksi kasus positif COVID-19 di kabupaten/kota lain masuk ke dalam zona “kuning” dengan nilai yang dihasilkan sekitar 0.0000287. Jadi, data uji kasus positif COVID-19 di kabupaten/kota lain terbilang lebih rendah.

### 3.2. Pengujian metode Naïve Bayes menggunakan Bahasa Pemrograman Python

Pengujian dilakukan menggunakan *tools Jupyter Notebook*. *Jupyter Notebook* merupakan *tools* yang populer untuk mengolah data di *python*. *Jupyter Notebook* memungkinkan untuk mengintegrasikan antara kode dengan *output* di dalam satu dokumen secara interaktif.

#### 3.2.1 Import Data CSV

Langkah awal yang dilakukan adalah meng-*import library* *sklearn* untuk menghitung prediksi *Naïve Bayes*, dan *library-library* pendukung lainnya. Setelah itu, dataset berupa format “CSV” di *import* dan dideklarasikan ke dalam variabel ‘data’.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score

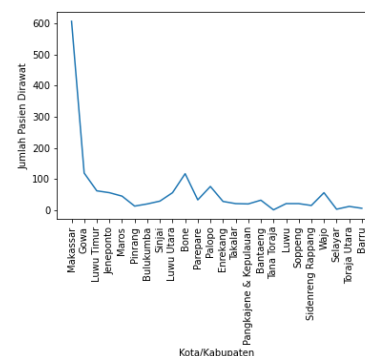
# Import data dari CSV
data = pd.read_csv('dataset_covid.csv')
data.head()
```

	Dirawat	Sembuh	Meninggal	Zona
0	606	8958	303	Orange
1	119	1438	30	Orange
2	62	1519	4	Orange
3	56	530	3	Orange
4	45	597	8	Orange

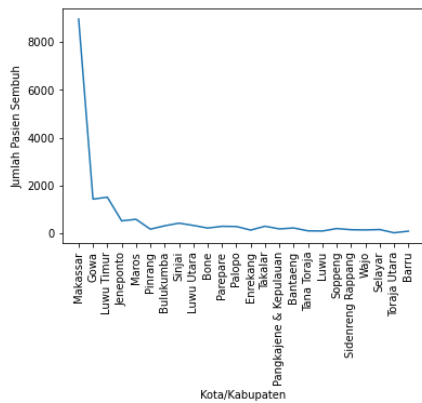
Gambar 1. Import Library dan data csv

#### 3.2.2 Visualisasi Data

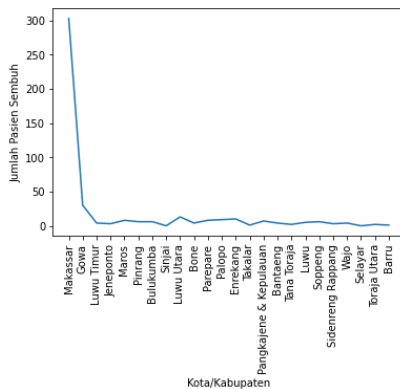
Visualisasi data dilakukan untuk menampilkan grafik perbandingan dari tiap parameter yang ada di dalam data.



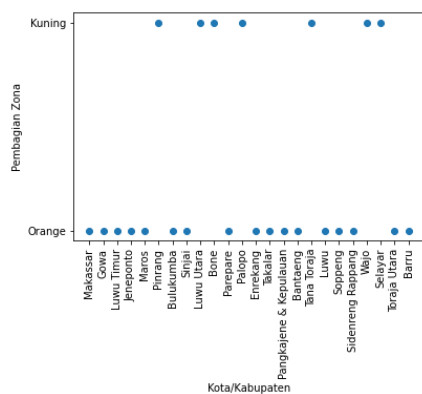
Gambar 2. Grafik pasien dirawat



Gambar 3. Grafik pasien sembuh



Gambar 4. Grafik pasien meninggal



Gambar 5. Grafik pembagian zona

### 3.2.3 Konversi Nilai string ke dalam bentuk numerik

Library *Naïve Bayes* hanya dapat membaca data dalam bentuk numerik, maka nilai parameter “Zona” akan dikonversi ke dalam bentuk numerik, dimana zona “Orange” bernilai 0, dan zona “Kuning” bernilai 1.

```
# Konversi value 'zona' kedalam bentuk angka
data.loc[data['Zona'] == 'Orange', 'Zona'] = 0
data.loc[data['Zona'] == 'Kuning', 'Zona'] = 1

data = data.apply(pd.to_numeric, errors='coerce')
data.head()
```

	Dirawat	Sembuh	Meninggal	Zona
0	606	8958	303	0
1	119	1438	30	0
2	62	1519	4	0
3	56	530	3	0
4	45	597	8	0

Gambar 6. Konversi nilai parameter zona ke dalam bentuk numerik

### 3.2.4 Penentuan Variabel Independen dan Variabel Dependen

Untuk memanggil variabel independen, cukup menghapus saja variabel dependen dengan perintah `drop` diikuti nama kolomnya yaitu kolom “Zona”. Axis = 1 berarti yang dihapus adalah kolom. Sedangkan, untuk memanggil variabel dependen menggunakan nama data diikuti nama kolomnya, yaitu kolom “Zona”.

```
x = data.drop(["Zona"], axis=1)
x.head()
```

	Dirawat	Sembuh	Meninggal
0	606	8958	303
1	119	1438	30
2	62	1519	4
3	56	530	3
4	45	597	8

```
y = data["Zona"]
y.head()
```

```
0 0
1 0
2 0
3 0
4 0
Name: Zona, dtype: int64
```

Gambar 7. Variabel independen dan dependen

### 3.2.5 Pemisahan atribut data latih dan data uji

Data *training* dan data uji perlu dipisahkan agar proses prediksi data dapat dilakukan. Data uji yang digunakan adalah sebanyak 40% dari data *training*. Maka `test_size` yang dimasukkan adalah 0.40.

```
# Pisahkan atribut untuk training & test
x_train, x_test, y_train, y_test =
train_test_split(x, y, test_size = 0.40, random_state = 0)
```

Gambar 8. Pemisahan atribut data training dan data uji

### 3.2.6 Pembuatan model Naïve Bayes pada data latih

```
# Membuat model Naïve Bayes terhadap Training set
model = GaussianNB()
model.fit(x_train, y_train)

GaussianNB()
```

Gambar 10. Membuat model Naïve Bayes pada data training

### 3.2.7 Perhitungan prediksi Naïve Bayes

Proses perhitungan prediksi *Naïve Bayes* menggunakan model *Gaussian* yang telah dibuat sebelumnya. Data yang diprediksi merupakan data uji.

```
y_predict = model.predict(x_test)
print("Prediksi Naïve Bayes : ",y_predict)

Prediksi Naïve Bayes : [0 0 0 0 1 0 0 1 1 0]
```

Gambar 11. Hasil prediksi Naïve Bayes

Berdasarkan gambar diatas, hasil prediksi dari *Naïve Bayes* yang didapatkan yaitu [0 0 0 0 1 0 0 1 1 0]. Jika dideklarasikan kembali ke dalam parameter “zona”, maka hasilnya adalah [Orange, Orange, Orange, Orange, Kuning, Orange, Orange, Kuning, Kuning, Orange].

### 3.2.8 Perhitungan Probabilitas

Sama halnya dengan proses perhitungan prediksi *Naïve Bayes*, perhitungan probabilitas menggunakan model *Gaussian*, yang kemudian menggunakan data uji untuk mencari nilai probabilitasnya.

```
model.predict_proba(x_test)

array([[9.02145887e-01, 9.78541135e-02],
 [9.29724785e-01, 7.02752146e-02],
 [9.99986173e-01, 1.38272555e-05],
 [5.82314581e-01, 4.17685419e-01],
 [1.19241861e-03, 9.98807581e-01],
 [1.00000000e+00, 0.00000000e+00],
 [8.57325162e-01, 1.42674838e-01],
 [2.31590698e-01, 7.68409302e-01],
 [4.90966583e-02, 9.50903342e-01],
 [9.99999802e-01, 1.98472700e-07]])
```

Gambar 12. Nilai Probabilitas

### 3.2.8 Perhitungan tingkat akurasi

Setelah mendapatkan hasil prediksi *Naïve Bayes* dan probabilitasnya, langkah terakhir yang dilakukan adalah menghitung tingkat akurasi prediksi *Naïve Bayes* yang dihasilkan dari proses perhitungan sebelumnya.

```
# Menghitung tingkat akurasi
accuracy_score(y_test, y_predict)*100

70.0
```

Gambar 9. Nilai Akurasi

Berdasarkan gambar 12 diatas, nilai akurasi yang dihasilkan adalah 70.0. Dalam artian, persentase *Correctly Classified Instance* adalah 70%, dan *Incorrectly Classified Instance* adalah 30%. Dimana, dari total 24 kabupaten/kota di Sulawesi Selatan dengan kasus positif COVID-19, ada 18 kabupaten/kota yang berhasil diklasifikasikan dengan benar, dan ada 6 kabupaten/kota yang tidak berhasil diklasifikasikan dengan benar.

## 4. Kesimpulan dan Saran

Berdasarkan pembahasan yang telah diuraikan, yaitu tentang Penerapan Data Mining Menggunakan Metode Naïve Bayes Dalam Memprediksi Tingkat Penyebaran COVID-19 Di Sulawesi Selatan, maka dapat disimpulkan beberapa hal sebagai berikut: 1) Metode Naïve Bayes memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk tiap kelas atau parameter yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi Tingkat Penyebaran COVID-19 di Provinsi Sulawesi Selatan berdasarkan proses klasifikasi yang dilakukan oleh metode Naïve Bayes itu sendiri. 2) Hasil pengujian pada perhitungan menggunakan tools Microsoft Excel 2019 pada data uji kabupaten/kota lain menghasilkan nilai sekitar

0.0000287, dalam artian kasus positif COVID-19 di kabupaten/kota lain terbilang lebih rendah. 3) Hasil pengujian menggunakan bahasa pemrograman Python di tools Jupyter Notebook dengan jumlah data training adalah 24, dan data uji-nya sebanyak 0.40, menghasilkan nilai akurasi sebesar 70%.

Untuk saran, beberapa pengembangan yang dapat dilakukan kedepannya adalah: 1) Dataset yang digunakan pada penelitian ini dinilai kurang efektif, dikarenakan data kasus positif COVID-19 di Sulawesi Selatan terus bertambah dan berubah setiap harinya. Oleh karena itu dibutuhkan dataset yang lebih efektif dan akurat, 2) Proses pengujian sebaiknya dilakukan menggunakan metode-metode lain, agar dapat menghasilkan perbandingan metode yang lebih akurat dalam memprediksi tingkat penyebaran COVID-19.

### Daftar Pustaka

- [1] Felicia, Alvina., Puspita. B, Arwini., Moeis, Dikwan. 2020. Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Journal Of Applied Computer Science And Technology (JACOST)*. 1(1). 7-14.
- [2] Arunala.com. (2020, 29 April). PSBB Sebagai Antisipasi Penyebaran Corona Virus Covid-19 di Sumatera Barat. Diakses pada 23 November 2020. <https://www.arunala.com/artikel/5/psbb-sebagai-antisipasi-penyebaran-corona-virus-covid-19-di-sumatera-barat.html>.
- [3] Alodokter.com. (2020, 20 November). Virus Corona. Diakses pada 19 November 2020. <https://www.alodokter.com/virus-corona>.
- [4] World Health Organization. (2019). Coronavirus. Diakses pada 23 November 2020. <https://www.who.int/health-topics/coronavirus>.
- [5] Pakpahan, Roida. Fitriani, Yuni. 2020. Analisa Pemanfaatan Teknologi Informasi Dalam Pembelajaran Jarak Jauh Di Tengah Pandemi Virus Corona Covid-19. *Journal of Information System, Applied, Management, Accounting and Research*. 4(2). 30-36.
- [6] Hadumaon, Tiodora. 2020. Mencari Kelompok Berisiko Tinggi Terinfeksi Virus Corona Dengan Discourse Network Analysis. *Jurnal Kebijakan Dan Kesehatan Indonesia (JKKI)*. 9(2). 98-106.
- [7] Inforcorona.makassar.go.id. (2020, 19 Maret). Kasus Pertama Di Sulsel, 2 Orang Positif Corona. Diakses pada 23 November 2020. <https://infocorona.makassar.go.id/blog/kasus-pertama-di-sulsel-2-orang-positif-corona/>.
- [8] Travel.detik.com. (2020, 7 Agustus). Ini Daftar Lengkap Zona Kuning dan Oranye Virus Corona di Indonesia. Diakses pada 19 November 2020. <https://travel.detik.com/travel-news/d-5124352/ini-daftar-lengkap-zona-kuning-dan-oranye-virus-corona-di-indonesia>
- [9] Bootup.ai. (2019, 17 Januari). Data Mining Adalah? Pengertian Hingga Belajar Clustering Lengkap!. Diakses pada 23 November 2020. <https://bootup.ai/blog/data-mining-adalah/>.
- [10] Hasan, Maryam. 2017. Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Menggunakan Algoritma Naive Bayes Berbasis Forward Selection. *ILKOM Jurnal Limiah*. 9(3). 317-324.
- [11] Rahmatullah, Sidik. Utmai, Ema. 2019. Prediksi Tingkat Kelulusan Tepat Waktu Dengan Metode Naive Bayes Dan K-Nearest Neighbor. *Jurnal Informasi dan Komputer*. 7(1). 7-16.
- [12] Covid19.sulselprov.go.id. (2020). Data Pantuan, COVID-19 di Sulawesi Selatan. Diakses pada 19 November 2020. <https://covid19.sulselprov.go.id/data>