

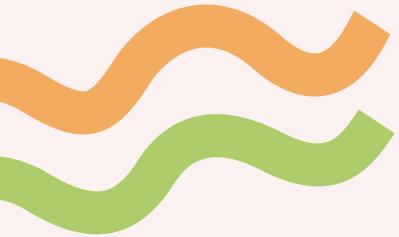
Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia



Latar Belakang

Latar belakang penelitian ini muncul dari perlunya langkah-langkah antisipatif dalam menghadapi pandemi COVID-19 yang melanda dunia. Sebagai krisis kesehatan global, COVID-19 telah mengubah tatanan kehidupan sehari-hari dan menimbulkan dampak kesehatan, sosial, dan ekonomi yang signifikan.





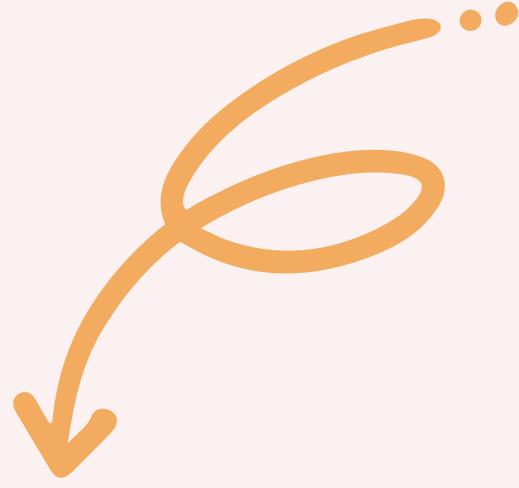
Rumusan Masalah

Dunia saat ini tengah waspada dengan penyebaran sebuah virus yang dikenal dengan virus corona. Corona viruses (CoV) merupakan bagian dari keluarga virus yang menyebabkan penyakit mulai dari flu hingga penyakit yang lebih berat seperti Middle East Respiratory Syndrome (MERS-CoV) and 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 (World Health Organization, 2019).

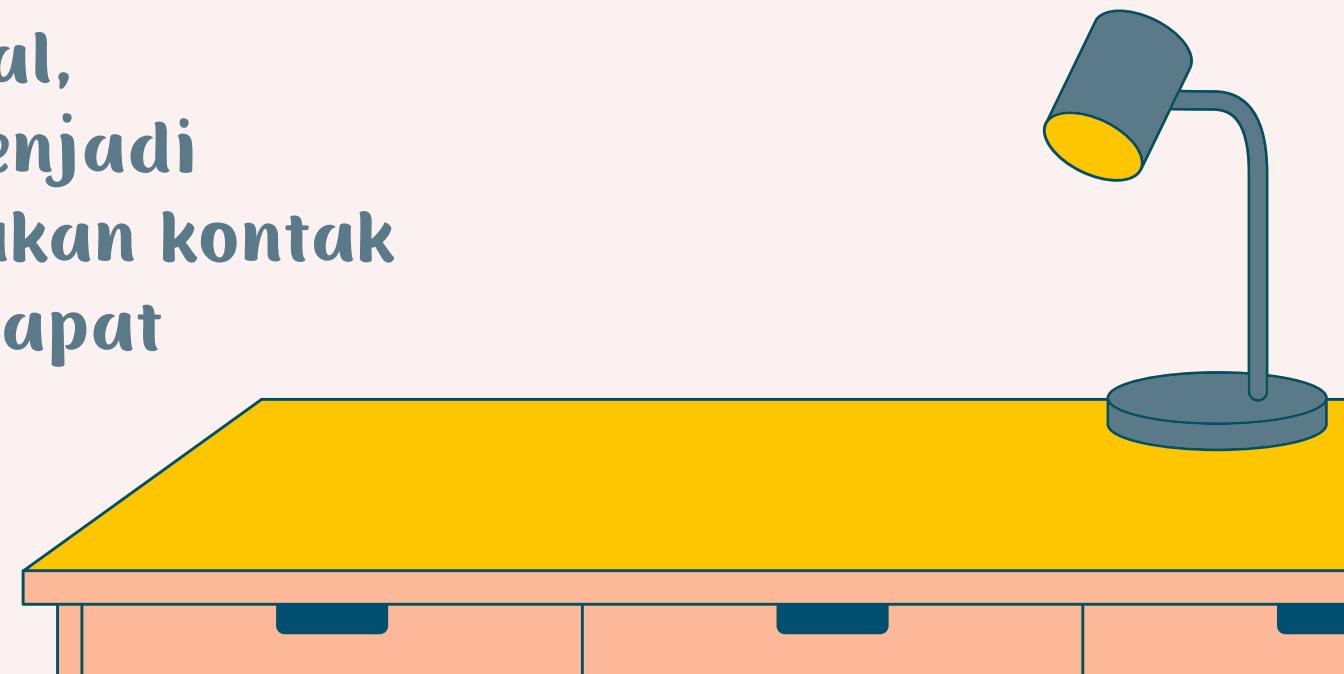




Respon Di Indonesia



Indonesia dalam status waspada terhadap ancaman virus corona tersebut dan sampai sekarang belum ditemukan vaksin COVID-19, maka bentuk pencegahan dari meluasnya penyebaran virus dapat dilakukan dengan cara memutuskan rantai penularannya. Salah satu metode untuk memutuskan rantai penularan tersebut adalah dengan melakukan pembatasan sosial (social distancing). Dengan adanya pembatasan sosial, harapannya setiap masyarakat tidak akan menjadi penular maupun tertular karena tidak melakukan kontak dengan siapapun sehingga laju penyebaran dapat menurun.



Tujuan Penelitian

Pada penelitian ini kami akan melakukan penerapan algoritma Naive Bayes untuk memprediksi tingkat penyebaran COVID-19 di Indonesia dengan mengimplementasikan Algoritma Naive Bayes, yang berguna sebagai langkah antisipasi pandemic COVID-19.



Metode Penelitian

1. Analisis Masalah dan Studi Literatur
Tahap ini adalah langkah awal untuk menentukan rumusan masalah dari penelitian. Dalam hal ini mengamati permasalahan yang berhubungan dengan tingkat penyebaran COVID-19 yang terjadi khususnya di Indonesia.

2. Mengumpulkan Data
Prosedur sistematis yang digunakan untuk mengumpulkan data yaitu dengan metode penelitian kuantitatif. Pelaksanaan metode penelitian kuantitatif focus pada penggunaan angka, tabel, grafik dan diagram untuk menampilkan hasil data yang diperoleh dan akan digunakan untuk bahan penganalisaan data terhadap metode Naive Bayes.

3. Implementasi
Sesuai dengan pengolahan data maka pada tahap implementasi adalah tentang bagaimana pengolahan datanya diterapkan dalam sebuah tools. Tools yang akan digunakan dalam implementasi penelitian ini adalah dengan menggunakan Software Weka.

Implementasi Metode Naive Bayes

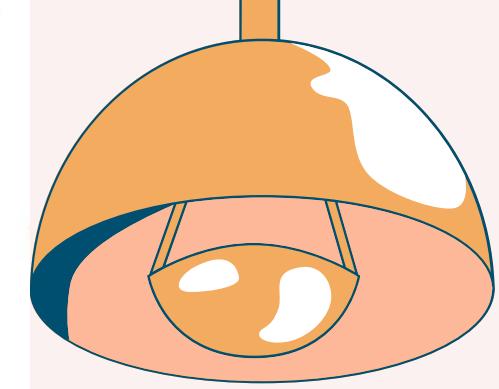
Tabel 1. Data Training (Data Latih)

No	Provinsi	Dalam Perawatan	Sembuh	Meninggal	Kasus Terbesar Per Provinsi				
1	DKI Jakarta	4038	1276	460	Positif	13	Kalimantan Timur	187	61
2	Jawa Timur	1449	294	178	Positif	14	Kalimantan Tengah	129	87
3	Jawa Barat	1237	259	100	Positif	15	Sumatera Utara	125	53
4	Jawa Tengah	805	234	70	Positif	16	Daerah Istimewa Yogyakarta	110	71
5	Sulawesi Selatan	527	293	51	Positif	17	Sulawesi Tenggara	154	24
6	Banten	404	158	60	Positif	18	Kalimantan Utara	117	23
7	Sumatera Selatan	376	73	9	Positif	19	Kalimantan Barat	104	22
8	Sumatera Barat	284	88	21	Positif	20	Kepulauan Riau	27	77
9	Kalimantan Selatan	330	24	9	Positif	21	Sulawesi Tengah	87	21
10	Nusa Tenggara Barat (NTB)	163	188	7	Negatif	22	Sulawesi Utara	70	30
11	Bali	107	232	4	Negatif	23	Riau	36	53
12	Papua	281	48	6	Positif	24	Papua Barat	85	2
						25	Maluku Utara	70	12
									3
									Positif
						26	Sulawesi Barat	50	22
						27	Jambi	66	3
						28	Lampung	39	22
						29	Maluku	41	17
						30	Bengkulu	50	1
						31	Kepulauan Bangka Belitung	21	2
						32	Gorontalo	7	14
						33	Nusa Tenggara Timur (NTT)	17	1
									1

Keterangan :

Positif = Kasus Pasien yang terkena Positif lebih besar

Negatif = Kasus pasien yang terkena positif lebih rendah



Implementasi Metode Naive Bayes

- a. untuk menghitung nilai rata – rata hitung (*mean*) dapat dilihat pada tabel sebagai berikut :

Tabel 2. Nilai Mean

Mean			
Kasus Terbesar Per Provinsi	Dalam Perawatan	Sembuh	Meninggal
Positif	401.8928571	115.2142857	37.35714286
Negatif	68	112.8	5.8

Mean merupakan nilai Rata-rata yang didapatkan dari hasil penjumlahan seluruh nilai dari masing-masing data, lalu dibagi dengan banyaknya data yang ada.

- b. Dan persamaan untuk menghitung nilai simpangan baku (*standar deviasi*) dapat dilihat sebagai berikut:

Tabel 3. Standar Deviasi

Standar Deviasi			
Kasus Terbesar Per Provinsi	Dalam Perawatan	Sembuh	Meninggal
Positif	795.2873852	244.5758774	91.65968227
Negatif	92.8477248	92.8477248	3.701351105

Standar deviasi merupakan nilai statistik yang dimanfaatkan untuk menentukan bagaimana sebaran data dalam sampel, serta seberapa dekat titik data individu ke mean atau rata-rata nilai sampel.

3.1.3 Probabilitas

Probabilitas merupakan suatu nilai yang digunakan untuk mengukur tingkat terjadinya suatu kejadian yang acak. Kata probabilitas itu sendiri sering disebut dengan peluang atau kemungkinan. Probabilitas secara umum merupakan peluang bahwa sesuatu akan terjadi.

Tabel 4. Probabilitas Kasus Terbesar Per Provinsi

Probabilitas Kasus Terbesar Per Provinsi	
Kasus Terbesar Per Provinsi	Nilai
Positif	0.848484848
Negatif	0.151515152

3.1.4 Menghitung Nilai Gaussian

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

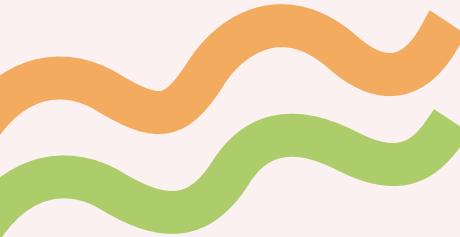
Tabel 5. Data Tes

Data Tes				
Provinsi	Dalam Perawatan	Sembuh	Meninggal	Kasus Terbesar Per Provinsi
Aceh	3	13	1	?
Positif	0.012477561	0.0233822	0.038527102	1.12E-05
Negatif	0.032268233	0.0232408	0.089464815	6.71E-05
	Max adalah			6.71E-05

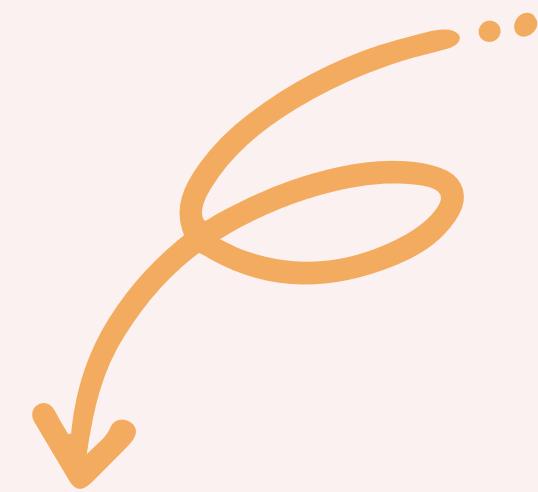
Berdasarkan data uji dari provinsi Aceh dengan data pasien "dalam perawatan" yaitu 3, data pasien "sembuh" yaitu 13, dan data pasien "meninggal" yaitu 1 diprediksi algoritma *naive bayes* dengan hasil kasus terbesar per provinsi yaitu kategori "Negatif" sekitar 6.71E-05. Maka data uji dari provinsi Aceh dikatakan bahwa tingkat penyebaran COVID-19 dengan kasus pasien yang terkena positif lebih rendah .

KESIMPULAN

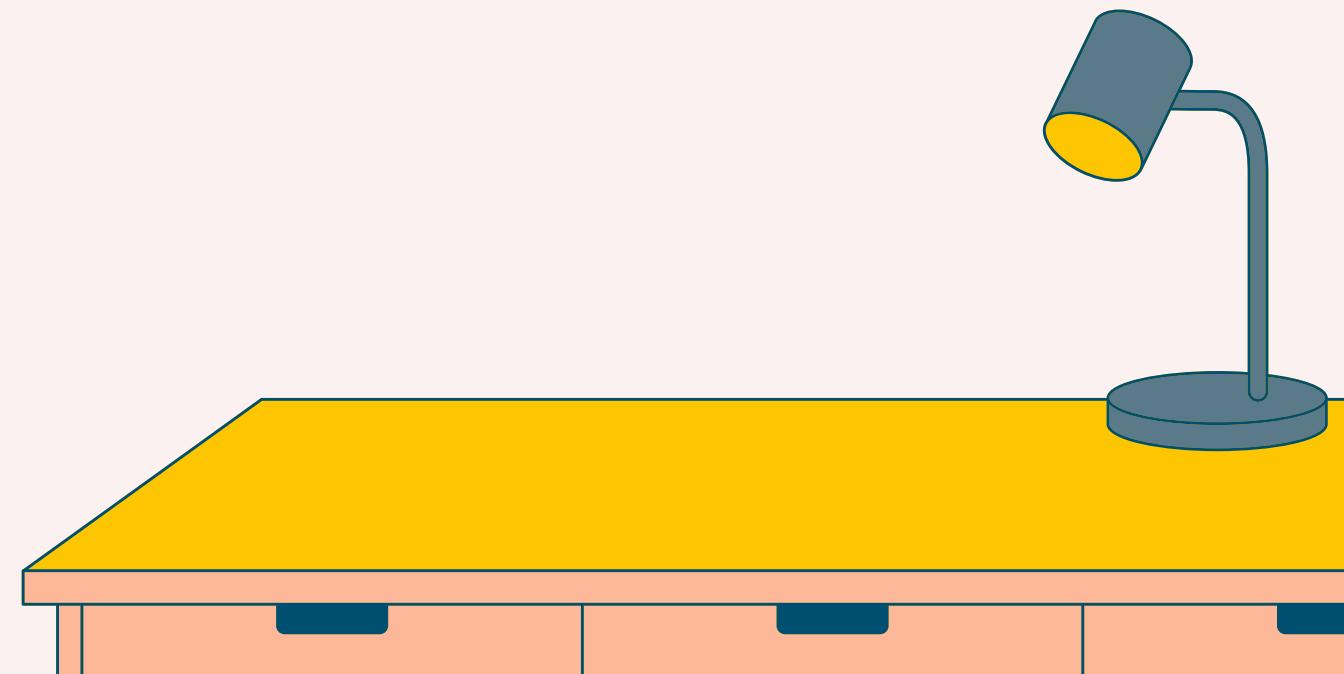
1. Metode Naive Bayes memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk class yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi Tingkat Penyebaran Covid 19 di Indonesia berdasarkan proses klasifikasi yang dilakukan oleh metode Naive Bayes itu sendiri.
2. Berdasarkan data Kasus COVID-19 Per Provinsi yang dijadikan data training, metode Naive Bayes berhasil mengklasifikasikan 16 data dari 33 data yang diuji. Sehingga metode Naive Bayes berhasil memprediksi besarnya Kasus COVID-19 Per Provinsi dengan persentase keakuratan sebesar 48,4848%.
3. Algoritma Naive Bayes di dukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Pada algoritma Naive Bayes, semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain



Saran



Adapun Saran yang diberikan mengenai penelitian tentang Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid 19 di Indonesia yaitu pengujian, sebaiknya dilakukan dengan menggunakan metode lain untuk melihat metode mana yang lebih akurat dalam memprediksi Tingkat Penyebaran Covid 19 di Indonesia.



Terima Kasih

