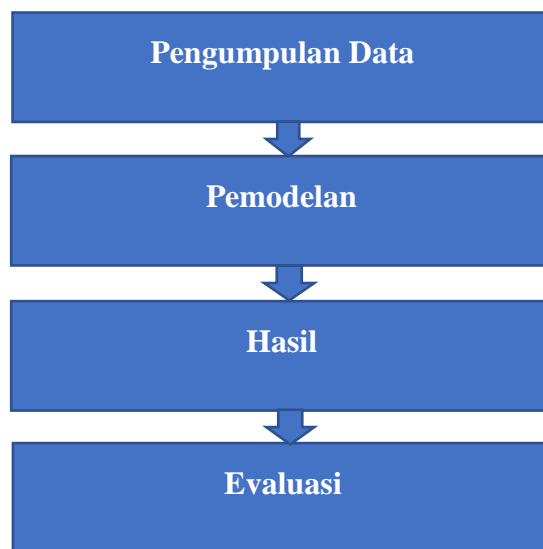


BAB III

METODOLOGI PENELITIAN

3.1. Metode Penelitian

Dalam melakukan analisa dan mencari pola data untuk dijadikan sebuah dataset dalam memudahkan penelitian dan dapat berjalan dengan sistematis dan memenuhi tujuan yang diinginkan maka dibuat alur dalam tahapan atau alur penelitian pada gambar 3.1



Gambar 3. 1. Alur dalam Tahapan Penelitian

Tahapan pada gambar 3.1 . Adalah proses data mining pada penelitian ini :

1. Pengumpulan data adalah mempersiapkan data mentah awal kumpulan data akhir yang akan digunakan untuk semua tahapan berikutnya. Pada tahapan ini memilih kasus dan variabel yang diinginkan, menganalisis yang sesuai untuk analisis. Melakukan transformasi pada variabel tertentu, jika diperlukan, membersihkan data mentah sehingga siap untuk alat pemodelan. Data pada penelitian ini terdiri dari 5434 record, 20 atribut dan

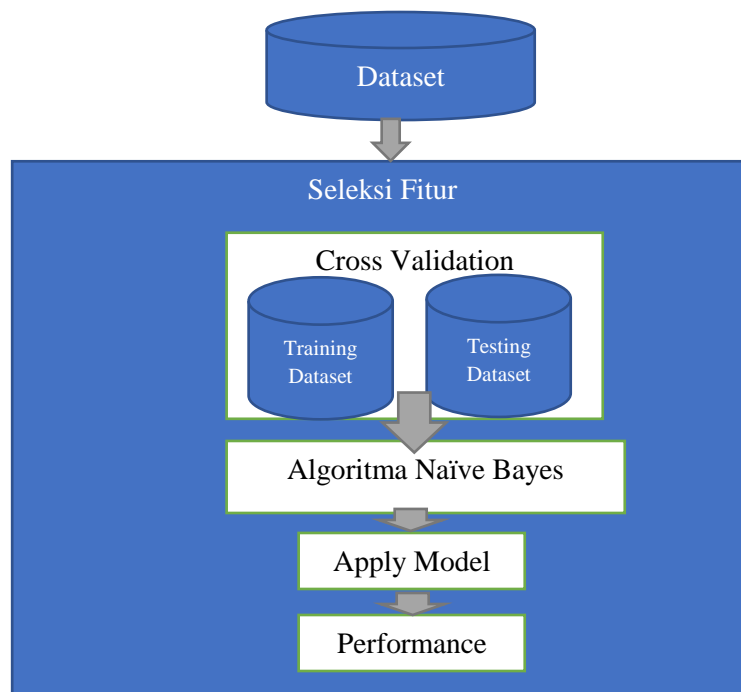
target *class* adalah Covid 19 dengan 2 label kelas Yes/No Covid 19. Tabel 3.1 menjelaskan atribut-atribut yang digunakan pada penelitian ini:

Tabel 3.1 Deskripsi Atribut Data

ATRIBUT	DESKRIPSI	NILAI
Breathing Problem	Mengalami masalah pernapasan	YES/NO
Fever	Mengalami Demam	YES/NO
Dry Cough	Mengalami Batuk Kering	YES/NO
Sore throat	Mengalami sakit tenggorokan	YES/NO
Running Nose	Hidung Meler	YES/NO
Asthma	Memiliki penyakit Astma	YES/NO
Chronic Lung Disease	Memiliki penyakit paru-paru kronis	YES/NO
Headache	Mengalami sakit kepala	YES/NO
Heart Disease	Memiliki penyakit Jantung	YES/NO
Diabetes	Memiliki penyakit diabetes	YES/NO
Hyper Tension	Memiliki penyakit Darah Tinggi	YES/NO
Fatigue	Mengalami kelelahan	YES/NO
Gastrointestinal	Memiliki masalah pencernaan	YES/NO
Abroad travel	Melakukan perjalanan ke luar negeri	YES/NO
Contact with COVID Patient	Kontak dengan pasien covid	YES/NO
Attended Large Gathering	Menghadiri pertemuan besar	YES/NO
Visited Public Exposed Places	Mengunjungi tempat publik	YES/NO
Family working in Public Exposed Places	Memiliki keluarga yang bekerja di tempat umum	YES/NO
Wearing Masks	Menggunakan masker	YES/NO
Sanitization from Market	Mengunjungi pasar	YES/NO
COVID-19	Target <i>Class</i>	YES/NO

Tabel diatas menjelaskan masing-masing atribut dari data pada penelitian ini. Tujuan dari dataset ini adalah untuk memberikan gejala sebagai input dan harus dapat memprediksi apakah COVID mungkin iya atau tidak.

2. Tahap pemodelan, Tahapan ini dilakukan dengan memilih dan menerapkan teknik pemodelan yang sesuai. kalibrasi pengaturan model untuk mengoptimalkan hasil, seringkali, beberapa teknik berbeda dapat digunakan untuk masalah data mining yang sama, apabila diperlukan mengulangi kembali ke tahapan persiapan data untuk membawa bentuk data sesuai dengan persyaratan khusus dari teknik data mining tertentu. Pemodelan pada penelitian ini menggunakan teknik data mining klasifikasi menggunakan seleksi fitur *Particle Swarm Optimization* (PSO) dan naïve bayes. Gambar 3.2 adalah bagan pemodelan dari penelitian ini.



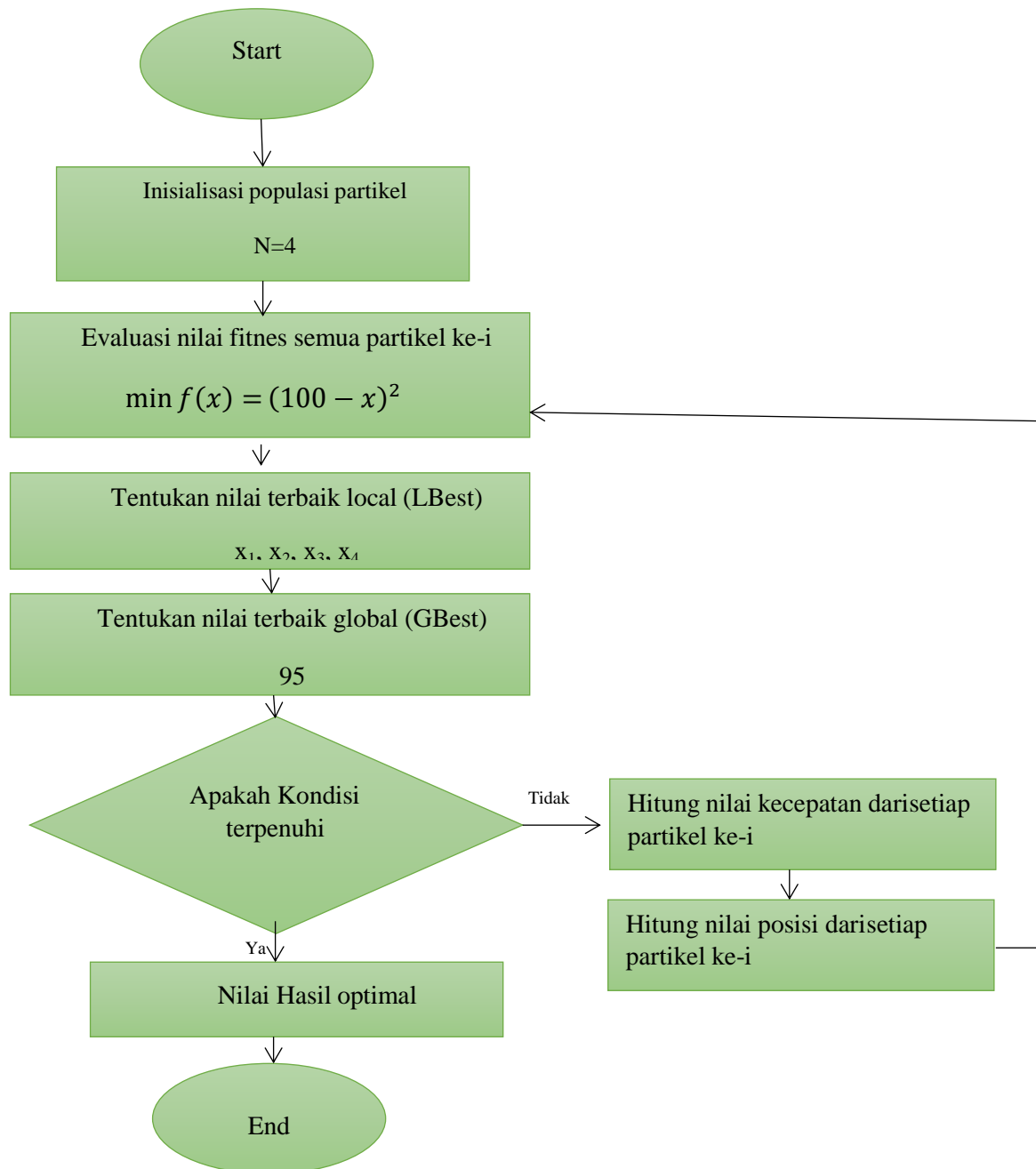
Gambar 3.2. Bagan Pemodelan

Gambar 3.2 adalah bagan dari pemodelan yang akan dilakukan pada penelitian ini pada tahap awal data yang ada akan dilakukan seleksi fitur menggunakan *Particle Swarm Optimization* dan *Genetic Algorithm* (GA). Selanjutnya akan dilakukan proses klasifikasi menggunakan algoritma *Naïve Bayes* dengan *Cross Validation* untuk membagi data testing dan training, *Apply Model* digunakan untuk melakukan pengujian

terhadap prediksi label, terakhir performance untuk mengetahui hasil dari akurasi. Berikut :

a. Seleksi fitur *Particle Swarm Optimization* (PSO)

Langkah-langkah atau cara kerja dari Algoritma PSO pada gambar 3.3



Gambar 3.3. Flowchart Algoritma PSO

Berikut ini adalah pengolahan data penyakit covid 19 secara manual menggunakan Particle Swarm Optimization, perhitungan dilakukan menggunakan sampel random:

1. Tentukan jumlah partikel $N = 4$
Tentukan populasi awal secara random contoh

$$\mathbf{X1(0) = 75}$$

$$\mathbf{X2(0) = 80}$$

$$\mathbf{X3(0) = 90}$$

$$\mathbf{X4(0) = 95}$$

Langkah selanjut nya memasukan angka angkanya kedalam persamaan untuk menentukan nilai fitness

Rumus

$$\min f(x) = (100 - x)^2$$

Partikel	X (nilai random)	f(x)
x1	75	625
x2	80	400
x3	90	100
x4	95	25

2. Evaluasi nilai fungsi tujuan untuk setiap partikel $X_j(0)$ untuk $j=1,2,3,4$ dan nyatakan dengan

$$\mathbf{F1 = f(75) = 625}$$

$$\mathbf{F2 = f(80) = 400}$$

$$\mathbf{F3 = f(90) = 100}$$

$$\mathbf{F4 = f(95) = 25}$$

3. Tentukan kecepatan awal $V1(0) = V2(0)$

= V3 (0) = V4 (0) = 0(kecepatan Awalnya)

(Velositi)	V
v1	0
v2	0
v3	0
v4	0

4. Tentukan nilai Pbest dan Gbest

$$P_{best\ 1} = 75$$

$$P_{best\ 2} = 80$$

$$P_{best\ 3} = 90$$

$$P_{best\ 4} = 95$$

$$G_{best} = 95$$

Mencari nilai minimum

$$\min f(x) = (100 - x)^2$$

Partikel	X (Nilai Random	f(x)
x1	75	625
x2	80	400
x3	90	100
x4	95	25
Ini termasuk lokal best	Kalau Pbest semuanya	

1 kawanan 4 partikel dalam 1 kawanan maka kita pilih mana yang paling the best maka didapatlah nilai Pbest

Partikel	X (nilai random)	f(x)	Pbest
x1	75	625	75
x2	80	400	80
x3	90	100	90
x4	95	25	95
Nilai pbes kita ambil dari nilai lokal(Lokasi) bestnya			

Partikel	X (nilai random)	f(x)	Pbest	Gbest
x1	75	625	75	95
x2	80	400	80	
x3	90	100	90	
x4	95	25	95	

Nilai Gbestnya dari mana dari nilai 95 yang terbaik untuk lokalbestnya, Artinya dari 1 kawanan partikel ini maka Gbest adalah 95

Hitung V(j) dengan $c_1=c_2=1$, misalkan nilai random yang didapat $r_1=0,4$, $r_2=0,5$ dengan rumus:

$$\begin{aligned}
 V_j(i) &= V_j(i-1) \\
 &+ c_1 r_1 [Pbest_{.j} - x_j(i-1)] \\
 &+ c_2 r_2 [Gbest_{.j} - x_i(i-1)]
 \end{aligned}$$

5. Hitung V(j) dengan $C_1, c_2= 1$ misalkan nilai random yang didapat $r_1=0.4$, $r_2=0.5$

c1	1
c2	1
r1	0,4
r2	0,5

Gunakan rumus

$$\begin{aligned}
 V_j(i) &= V_j(i-1) + c_1 r_1 [Pbest_{.j} - x_j(i-1)] \\
 &+ c_2 r_2 [Gbest_{.j} - x_i(i-1)]
 \end{aligned}$$

Tetapkan iterasi $i=1$;

	V
v1	0
v2	0
v3	0
v4	0

Diperoleh

$$V_1(1) = 0 + 0.4(75-75) + 0.5(95-75) = 10$$

$$V_2(1) = 0 + 0.4(80-80) + 0.5(95-80) = 7,5$$

$$V_3(1) = 0 + 0.4(90-90) + 0.5(95-90) = 2,5$$

$$V_4(1) = 0 + 0.4(95-95) + 0.5(95-95) = 0$$

	V	V _j (velositi)
v1	0	10
v2	0	7,5
v3	0	2,5
v4	0	0

Selanjutnya menghitung nilai x_1, x_2, x_3, x_4

$$x_i(t) = v_i(t) + x_i(t-1)$$

$$X_1(1) = 75 + 10 = 85$$

$$X_2(1) = 80 + 7,5 = 87,5$$

$$X_3(1) = 90 + 2,5 = 92,5$$

$$X_4(1) = 95 + 0 = 95$$

	V	v _j (velosopi)	x _i (lokasi yang baru)
v1	0	10	85
v2	0	7,5	87,5
v3	0	2,5	92,5
v4	0	0	95

6. Evaluasi nilai fungsi tujuan sekarang pada partikel $X_j(1)$

$$f_1(1) = f(85) = 225$$

$$f_2(1) = f(87,5) = 156,25$$

$$f_3(1) = f(92,5) = 56,25$$

$$f_4(1) = f(95) = 25$$

Sedangkan pada partikel iterasi sebelumnya kita dapatkan

$$f_1(0) = f(75) = 625$$

$$f_2(0) = f(80) = 400$$

$$f_3(0) = f(90) = 100$$

$$f_4(0) = f(95) = 25$$

$$\min f(x) = (100 - x)^2$$

	V	vj(velosofi)	xi (Lokasi yang baru)	f(x)
v1	0	10	85	225
v2	0	7,5	87,5	156,25
v3	0	2,5	92,5	56,25
v4	0	0	95	25

7. Setelah di dapat kita hitung lagi Pbest dan Gbestnya
 Nilai f dari iterasi sebelumnya tidak ada yang lebih baik sehingga Pbest untuk maing-masing partikel sama dengan nilai x-nya nilai Gbest **95**

$$f_1(1) = f(85) = 225$$

$$f_2(1) = f(87,5) = 156,25$$

$$f_3(1) = f(92,5) = 56,25$$

$$f_4(1) = f(95) = 25$$

Sedangkan pada partikel iterasi sebelumnya kita dapatkan

$$f_1(0) = f(75) = 625$$

$$f_2(0) = f(80) = 400$$

$$f_3(0) = f(90) = 100$$

$$f_4(0) = f(95) = 25$$

Maka tidak ada yang lebih baik dari pbest
 maka masing-masing partikel itu nilainya
 sama dengan Gbestnya (**95**)

8. Cek apakah solusi X sudah konvergen,
 dimana nilai X saling dekat, jika tidak,
 tingkatkan ke iterasi berikutnya $i=2$

$$P_{best\ 1} = 85$$

$$P_{best\ 2} = 87,5$$

$$P_{best\ 3} = 92,5$$

$$P_{best\ 4} = 95$$

$$G_{best} = 95$$

Pbest	Gbest
85	95
87,5	
92,5	
95	

9. Hitung kecepatan baru dengan $r_1=0,3$ dan $r_2=0,6$ secara random

Pbest	Gbest	r_1	0,3
85	95	r_2	0,6
87,5		C1	1
92,5		C2	1
95			

$$V_1(2) = 10 + 0.3(85-85) + 0.6(95-85) = 16$$

$$V_2(2) = 7,5 + 0.3(87,5-87,5) + 0.6(95-87,5) = 12$$

$$V_3(2) = 2,5 + 0.3(92,5-92,5) + 0.6(95-92,5) = 4$$

$$V_4(2) = 0 + 0.3(95-95) + 0.6(95-95) = 0$$

	V_j (pilositi yg baru)
v1	16
v2	12
v3	4
v4	0

Sedangkan untuk nilai X adalah

$$X_1(2) = 85 + 16 = 101$$

$$X_2(2) = 87,5 + 12 = 99,5$$

$$X_3(2) = 92,5 + 4 = 96,5$$

$$X_4(2) = 95 + 0 = 95$$

	V_j	x_i
v1	16	101
v2	12	99,5
v3	4	96,5
v4	0	95

10. Evaluasi nilai fungsi tujuan sekarang pada partikel $X_j(2)$

$$f_1(1) = f(101) = 1$$

$$f_2(1) = f(99,5) = 0,25$$

$$f_3(1) = f(96,5) = 12,25$$

$$f_4(1) = f(95) = 25$$

$$\min f(x) = (100 - x)^2$$

	Vj	xi	f(x)
v1	16	101	1
v2	12	99,5	0,25
v3	4	96,5	12,25
v4	0	95	25

Jika dibandingkan dengan nilai f dari iterasi sebelumnya, ada nilai yang lebih baik dari nilai f sekarang yaitu $f_2(1) = 0,25$, sehingga P_{best} untuk partikel 3 sama dengan 99,5, dan G_{best} dicari dari $\min(1, 0,25, 12,25, 25) = 0$ yang dicapai pada $X_2(1) = 99,5$, sehingga untuk iterasi berikutnya $P_{best} = (101, 99,5, 96,5, 95)$ dan $G_{best} = 95$

11. Evaluasi nilai fungsi tujuan sekarang pada partikel $x_j(2)$

$$f_1(2) = f(75) = 1$$

$$f_2(2) = f(80) = 0,25$$

$$f_3(2) = f(90) = 12,25$$

$$f_4(2) = f(95) = 25$$

Sedangkan pada partikel iterasi sebelumnya kita dapatkan

$$f_1(1) = f(85) = 225$$

$$f_2(1) = f(87,5) = 156,25$$

$$f_3(1) = f(92,5) = 56,25$$

$$f_4(1) = f(95) = 25$$

Sedangkan pada partikel iterasi sebelumnya kita dapatkan

$$f_1(0) = f(80) = 625$$

$$f_2(0) = f(90) = 400$$

$$f_3(0) = f(110) = 100$$

$$f_4(0) = f(75) = 25$$

Cek apakah solusi sudah konvergen, dimana nilai x saling dekat. Jika belum konvergen, set $i=3$, masuk ke iterasi berikutnya. Lanjutkan ke langkah berikutnya dengan menghitung kecepatan v dan ulangi langkah2 selanjutnya sampai mencapai konvergen

Kesimpulannya : solusi sudah konvergen dengan jarak terdekat pada $f(2) = f(80) = 0,25$

b. Klasifikasi menggunakan Naïve Bayes

Perhitungan klasifikasi akan menggunakan data pada dataset penelitian ini:

$$P(x|y) = \frac{P(y|x)xP(x)}{P(y)}$$

$$P(y) = \sum_{n=1}^n P(y|x)P(x)$$

Total Record	=5434
Class Yes Covid	=4383
Class No Covid	=1051
P C "Yes"	=0.806588149
P C "No"	=0.193411851

1. Menghitung Probabilitas Atribut Breathing Problem

<u>Jumlah Yes Breathing Problem</u>	Yes Covid	3369=0.768652
Jumlah Class Yes Covid		4383
<u>Jumlah Yes Breathing Problem</u>	No Covid	251=0.23882
Jumlah Class NO Covid		1051
<u>Jumlah No Breathing Problem</u>	Yes Covid	1014=0.231348
Jumlah Class Yes Covid		4383
<u>Jumlah No Breathing Problem</u>	No Covid	1014=0.76118
Jumlah Class No Covid		1051

Breathing Problem	YES	NO	P(YES)	P(NO)
YES	3369	251	0.768652	0.2388
NO	1014	800	0.231348	0.7611

2. Menghitung Probabilitas Atribut Fever

$$\frac{\text{Jumlah Yes Fever Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{3757}{4383} = 0.857175$$

$$\frac{\text{Jumlah Yes Fever No Covid}}{\text{Jumlah Class NO Covid}} = \frac{516}{1051} = 0.490961$$

$$\frac{\text{Jumlah No Fever Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{626}{4383} = 0.142825$$

$$\frac{\text{Jumlah No Fever No Covid}}{\text{Jumlah Class No Covid}} = \frac{535}{1051} = 0.509039$$

Fever	YES	NO	P(YES)	P(NO)
YES	3757	516	0.857175	0.490961
NO	626	535	0.142825	0.509039

3. Menghitung Probabilitas Atribut Dry Cough

$$\frac{\text{Jumlah Yes Dry Cough Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{3878}{4383} = 0.884782$$

$$\frac{\text{Jumlah Yes Dry Cough No Covid}}{\text{Jumlah Class NO Covid}} = \frac{429}{1051} = 0.408183$$

$$\frac{\text{Jumlah No Dry Cough Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{505}{4383} = 0.115218$$

$$\frac{\text{Jumlah No Dry Cough No Covid}}{\text{Jumlah Class No Covid}} = \frac{622}{1051} = 0.591817$$

Dry Cough	YES	NO	P(YES)	P(NO)
YES	3878	429	0.884782	0.408183
NO	505	622	0.115218	0.591817

4. Menghitung Probabilitas Atribut Sore Throat

$$\frac{\text{Jumlah Yes Sore Throat Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{3669}{4383} = 0.837097878$$

$$\frac{\text{Jumlah Yes Sore Throat No Covid}}{\text{Jumlah Class NO Covid}} = \frac{284}{1051} = 0.270219$$

$$\frac{\text{Jumlah No Sore Throat Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{714}{4383} = 0.162902122$$

$$\frac{\text{Jumlah No Sore Throat No Covid}}{\text{Jumlah Class No Covid}} = \frac{767}{1051} = 0.729781$$

Sore throat	YES	NO	P(YES)	P(NO)
YES	3669	284	0.837097878	0.270219
NO	714	767	0.162902122	0.729781

5. Menghitung Probabilitas Atribut Running Nose

$$\frac{\text{Jumlah Yes Running Nose Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2375}{4383} = 0.541866$$

$$\frac{\text{Jumlah Yes Running Nose No Covid}}{\text{Jumlah Class NO Covid}} = \frac{577}{1051} = 0.549001$$

$$\frac{\text{Jumlah No Running Nose Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2008}{4383} = 0.458134$$

$$\frac{\text{Jumlah No Running Nose No Covid}}{\text{Jumlah Class No Covid}} = \frac{474}{1051} = 0.450999$$

Running Nose	YES	NO	P(YES)	P(NO)
YES	2375	577	0.541866	0.549001
NO	2008	474	0.458134	0.450999

6. Menghitung Probabilitas Atribut Asthma

$$\frac{\text{Jumlah Yes Asthma Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2124}{4383} = 0.4846$$

$$\frac{\text{Jumlah Yes Asthma No Covid}}{\text{Jumlah Class NO Covid}} = \frac{390}{1051} = 0.371075$$

$$\frac{\text{Jumlah No Asthma Yes Covid } 2259}{\text{Jumlah Class Yes Covid } 4383} = 0.5154$$

$$\frac{\text{Jumlah No Asthma No Covid } 661}{\text{Jumlah Class No Covid } 1051} = 0.628925$$

Asthma	YES	NO	P(YES)	P(NO)
YES	2124	390	0.4846	0.371075
NO	2259	661	0.5154	0.628925

7. Menghitung Probabilitas Atribut Chronic Lung Disease

$$\frac{\text{Jumlah Yes Chronic Lung Disease Yes Covid } 2008}{\text{Jumlah Class Yes Covid } 4383} = 0.458134$$

$$\frac{\text{Jumlah Yes Chronic Lung Disease No Covid } 557}{\text{Jumlah Class NO Covid } 1051} = 0.529971$$

$$\frac{\text{Jumlah No Chronic Lung Disease Yes Covid } 2375}{\text{Jumlah Class Yes Covid } 4383} = 0.541866$$

$$\frac{\text{Jumlah No Chronic Lung Disease No Covid } 494}{\text{Jumlah Class No Covid } 1051} = 0.470029$$

Chronic Lung Disease	YES	NO	P(YES)	P(NO)
YES	2008	557	0.458134	0.529971
NO	2375	494	0.541866	0.470029

8. Menghitung Probabilitas Atribut Heart Disease

$$\frac{\text{Jumlah Yes Heart Disease Yes Covid } 2064}{\text{Jumlah Class Yes Covid } 4383} = 0.47091$$

$$\frac{\text{Jumlah Yes Heart Disease No Covid } 459}{\text{Jumlah Class NO Covid } 1051} = 0.436727$$

$$\frac{\text{Jumlah No Heart Disease Yes Covid } 2319}{\text{Jumlah Class Yes Covid } 4383} = 0.52909$$

$$\frac{\text{Jumlah No Heart Disease No Covid } 592}{\text{Jumlah Class No Covid } 1051} = 0.563273$$

Heart Disease	YES	NO	P(YES)	P(NO)
YES	2064	459	0.47091	0.436727
NO	2319	592	0.52909	0.563273

9. Menghitung Probabilitas Atribut Diabetes

$$\frac{\text{Jumlah Yes Diabetes Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2131}{4383} = 0.486197$$

$$\frac{\text{Jumlah Yes Diabetes No Covid}}{\text{Jumlah Class No Covid}} = \frac{457}{1051} = 0.434824$$

$$\frac{\text{Jumlah No Diabetes Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2252}{4383} = 0.513803$$

$$\frac{\text{Jumlah No Diabetes No Covid}}{\text{Jumlah Class No Covid}} = \frac{594}{1051} = 0.565176$$

Diabetes	YES	NO	P(YES)	P(NO)
YES	2131	457	0.486197	0.434824
NO	2252	594	0.513803	0.565176

10. Menghitung Probabilitas Atribut Abroad travel

$$\frac{\text{Jumlah Yes Abroad travel Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2451}{4383} = 0.559206$$

$$\frac{\text{Jumlah Yes Abroad travel No Covid}}{\text{Jumlah Class NO Covid}} = \frac{0}{1051} = 0$$

$$\frac{\text{Jumlah No Abroad travel Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{1932}{4383} = 0.440794$$

$$\frac{\text{Jumlah No Abroad travel No Covid}}{\text{Jumlah Class No Covid}} = \frac{1051}{1051} = 1$$

Abroad travel	YES	NO	P(YES)	P(NO)
YES	2451	0	0.559206	0
NO	1932	1051	0.440794	1

11. Menghitung Probabilitas Atribut **Contact with COVID Patient**

$$\frac{\text{Jumlah Yes **Contact with COVID Patient** Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2582}{4383} = 0.589094$$

$$\frac{\text{Jumlah Yes **Contact with COVID Patient** No Covid}}{\text{Jumlah Class No Covid}} = \frac{144}{1051} = 0.137012$$

Jumlah Class NO Covid 1051

$\frac{\text{Jumlah No Contact with COVID Patient Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{1801}{4383} = 0.137012$

$\frac{\text{Jumlah No Contact with COVID Patient No Covid}}{\text{Jumlah Class No Covid}} = \frac{907}{1051} = 0.862988$

Contact with COVID Patient	YES	NO	P(YES)	P(NO)
YES	2582	144	0.589094	0.137012
NO	1801	907	0.137012	0.862988

12. Menghitung Probabilitas Atribut Attended Large Gathering

$\frac{\text{Jumlah Yes Attended Large Gathering Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2442}{4383} = 0.557153$

$\frac{\text{Jumlah Yes Attended Large Gathering No Covid}}{\text{Jumlah Class NO Covid}} = \frac{68}{1051} = 0.0647$

$\frac{\text{Jumlah No Attended Large Gathering Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{1941}{4383} = 0.442847$

$\frac{\text{Jumlah No Attended Large Gathering No Covid}}{\text{Jumlah Class No Covid}} = \frac{983}{1051} = 0.9353$

Attended Large Gathering	YES	NO	P(YES)	P(NO)
YES	2442	68	0.557153	0.0647
NO	1941	983	0.442847	0.9353

13. Menghitung Probabilitas Atribut Visited Public Exposed Places

$\frac{\text{Jumlah Yes Visited Public Exposed Places Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{2403}{4383} = 0.548255$

$\frac{\text{Jumlah Yes Visited Public Exposed Places No Covid}}{\text{Jumlah Class NO Covid}} = \frac{417}{1051} = 0.396765$

$\frac{\text{Jumlah No Visited Public Exposed Places Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{1980}{4383} = 0.451745$

$\frac{\text{Jumlah No Visited Public Exposed Places No Covid}}{\text{Jumlah Class No Covid}} = \frac{634}{1051} = 0.603235$

Visited Public Exposed Places	YES	NO	P(YES)	P(NO)
YES	2403	417	0.548255	0.396765
NO	1980	634	0.451745	0.603235

14. Menghitung Probabilitas Atribut Family working in Public Exposed Places

$$\frac{\text{Jumlah Yes Family working in Public Exposed Places Yes Covid } 1994}{\text{Jumlah Class Yes Covid } 4383} = 0.45494$$

$$\frac{\text{Jumlah Yes Family working in Public Exposed Places No Covid } 268}{\text{Jumlah Class NO Covid } 1051} = 0.254995$$

$$\frac{\text{Jumlah No Family working in Public Exposed Places Yes Covid } 2389}{\text{Jumlah Class Yes Covid } 4383} = 0.54506$$

$$\frac{\text{Jumlah No Family working in Public Exposed Places No Covid } 783}{\text{Jumlah Class No Covid } 1051} = 0.745005$$

Family working in Public Exposed Places	YES	NO	P(YES)	P(NO)
YES	1994	268	0.45494	0.254995
NO	2389	783	0.54506	0.745005

15. Menghitung Probabilitas Atribut Wearing Masks

$$\frac{\text{Jumlah Yes Wearing Masks Yes Covid } 0}{\text{Jumlah Class Yes Covid } 4383} = 0$$

$$\frac{\text{Jumlah Yes Wearing Masks No Covid } 0}{\text{Jumlah Class NO Covid } 1051} = 0$$

$$\frac{\text{Jumlah No Wearing Masks Yes Covid } 4383}{\text{Jumlah Class Yes Covid } 4383} = 1$$

$$\frac{\text{Jumlah No Wearing Masks No Covid } 1051}{\text{Jumlah Class No Covid } 1051} = 1$$

Wearing Masks	YES	NO	P(YES)	P(NO)
YES	0	0	0	0
NO	4383	1051	1	1

16. Menghitung Probabilitas Atribut Sanitization from Market

$$\frac{\text{Jumlah Yes Sanitization from Market Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{0}{4383} = 0$$

$$\frac{\text{Jumlah Yes Sanitization from Market No Covid}}{\text{Jumlah Class NO Covid}} = \frac{0}{1051} = 0$$

$$\frac{\text{Jumlah No Sanitization from Market Yes Covid}}{\text{Jumlah Class Yes Covid}} = \frac{4383}{4383} = 1$$

$$\frac{\text{Jumlah No Sanitization from Market No Covid}}{\text{Jumlah Class No Covid}} = \frac{1051}{1051} = 1$$

Sanitization from Market	YES	NO	P(YES)	P(NO)
YES	0	0	0	0
NO	4383	1051	1	1

Perhitungan naïve bayes dengan kasus pada tabel 3.2

No	Atribut	Kasus	P Yes	P No
1	Breathing Problem	Yes	0,768652	0,238820
2	Fever	Yes	0,857175	0,490961
3	Dry Cough	Yes	0,884782	0,408183
4	Sore throat	No	0,162902	0,729781
5	Running Nose	No	0,458134	0,450999
6	Asthma	Yes	0,484600	0,371075
7	Chronic Lung Disease	Yes	0,458134	0,529971
9	Heart Disease	No	0,529090	0,563273
13	Gastrointestinal	Yes	0,468629	0,472883
14	Abroad travel	Yes	0,559206	0,000000
15	Contact with COVID Patient	No	0,410906	0,862988
16	Attended Large Gathering	Yes	0,557153	0,064700
17	Visited Public Exposed Places	Yes	0,548255	0,396765
18	Family working in Public Exposed Places	Yes	0,454940	0,254995
19	Wearing Masks	No	1,000000	1,000000
20	Sanitization from Market	No	1,000000	1,000000
21	COVID-19		0,000367	0,000000

Dari tabel 3.2 diatas dapat diketahui nilai kemungkinan Yes (Covid) adalah 0,000367 dan kemungkinan No (Covid) adalah 0, sehingga kemungkinan hasil prediksi dari kasus adalah Yes (Covid) karena memiliki nilai yang lebih tinggi dari No (Covid)

3. Tahap evaluasi adalah mengevaluasi satu atau lebih model yang disampaikan dalam tahap pemodelan untuk kualitas dan efektivitas sebelum menerapkannya untuk digunakan di lapangan.. Pada tahap

ini yang akan dilakukan adalah menentukan model sebenarnya mencapai tujuan yang ditetapkan untuk itu pada tahapan pertama, menetapkan apakah beberapa aspek penting dari masalah bisnis atau penelitian belum diperhitungkan secara memadai., mengambil keputusan tentang penggunaan hasil data mining.

3.2. Alat dan Bahan

Penelitian ini menggunakan perangkat keras Laptop Asus Intel Celeron N4000 up to 2,6 Ghz, sedangkan perangkat lunak yang digunakan Microsoft Excel dan Rapidminer Studio aplikasi yang digunakan untuk pengolahan data