

**PREDIKSI TINDAKAN MEDIS PASIEN OTG, ODP, PDP DAN POSITIF  
COVID-19 MENGGUNAKAN METODE NAÏVE BAYES**



**LAPORAN PENELITIAN**

Sebagai Salah Satu Bentuk Pengamalan Tri Darma Perguruan Tinggi

**Oleh :**

<b>Nama</b>	<b>NIPY</b>
1. Arfan Haqiqi Sulasmoro, M.Kom	1. 02.009.054
2. Rais, S.Pd., M.Kom	2. 07.011.083
3. Istiqomah Dwi Andari, S.ST., M.Kes	3. 02.013.140

**PROGRAM STUDI DIII TEKNIK KOMPUTER  
POLITEKNIK HARAPAN BERSAMA  
TEGAL  
AGUSTUS 2021**

---

**No. SK Direktur : 098.05/PHB/V/2021 Tanggal, 31 Mei 2021  
No. Kontrak : 039.16/P3M.PHB/V/2021, Tanggal, 6 Mei 2020**

**HALAMAN PERSETUJUAN  
LAPORAN PENELITIAN**

**PREDIKSI TINDAKAN MEDIS PASIEN OTG, ODP, PDP DAN POSITIF  
COVID-19 MENGGUNAKAN METODE NAÏVE BAYES**

Sebagai Salah Satu Bentuk Pengamalan Tri Darma Perguruan Tinggi

Oleh :

- |                                       |               |
|---------------------------------------|---------------|
| 1. Arfan Haqiqi Sulasmoro, M.Kom      | 1. 02.009.054 |
| 2. Rais, S.Pd., M.Kom                 | 2. 07.011.083 |
| 3. Istiqomah Dwi Andari, S.ST., M.Kes | 3. 02.013.140 |

Tegal, Agustus 2021

Menyetujui,  
Ka.P3M  
Politeknik Harapan Bersama



Kusnadi, M.Pd  
NIPY. 04.015.217

Mengetahui,  
Ka Prodi DIII Teknik  
Komputer  
Politeknik Harapan Bersama



Rais, S.Pd, M.Kom  
NIPY. 07.011.083

**HALAMAN PENGESAHAN  
LAPORAN PENELITIAN**

1. **Judul Penelitian** : Prediksi Tindakan Medis Pasien Otg, Odp, Pdp Dan Positif Covid-19 Menggunakan Metode Naïve Bayes
2. **Ketua Peneliti:**
- a. Nama Lengkap : Arfan Haqiqi Sulasmoro, M.Kom
  - b. NIDN : 0623037704
  - c. NIPY : 02.009.054
  - d. Jabatan Fungsional : Asisten Ahli
  - e. Program Studi : DIII Teknik Komputer
  - f. Alamat e-mail :
3. **Jumlah Anggota** : 2
- Nama Anggota 1 : Rais, S.Pd., M.Kom
  - Nama Anggota 2 : Istiqomah Dwi Andari, S.ST., M.Kes
- Biaya Penelitian** : Rp. 3.157.000

Tegal, Agustus 2021

Reviewer 1



Arif Rakhman, SE, S.Pd, M.Kom  
NIPY. 05.016.291

Reviewer 2



M. Fikri Hidayattullah, M.Kom.  
NIPY. 09.016.307

Menyetujui,  
Ketua Prodi DIII Teknik Komputer  
Politeknik Harapan Bersama



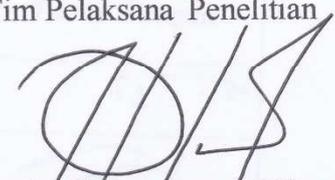
Rais, S.Pd., M.Kom  
NIPY. 07.011.083

Mengetahui,  
Wakil Direktur I  
Politeknik Harapan Bersama



Apt. Heru Nurcahyo, S.Farm, M.Sc  
NIPY. 10.007.038

Ketua Tim Pelaksana Penelitian



Arfan Haqiqi Sulasmoro, M.Kom  
NIPY. 02.009.054

Mengesahkan,  
Ketua P3M  
Politeknik Harapan Bersama



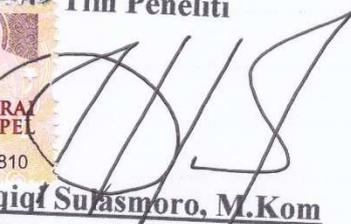
Kusrudi, M.Pd  
NIPY. 04.015.217

## PERNYATAAN

Dengan ini kami menyatakan bahwa:

1. Penelitian ini tidak pernah dibuat oleh peneliti lain dengan tema, judul, isi, metode, objek penelitian yang sama.
2. Penelitian ini bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi.
3. Dalam penelitian ini juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Tegal, Agustus 2021

Ketua Tim Peneliti  
  
Arfan Haqiqi Sulasmoro, M.Kom  
NIPY. 02.009.054

No	Nama	NIPY	Tanda Tangan
1	Rais, S.Pd., M.Kom	07.011.083	1 
2	Istiqomah Dwi Andari, S.ST., M.Kes	02.013.140	2 

## DAFTAR ISI

	Halaman
HALAMAN COVER .....	i
HALAMAN PERSETUJUAN .....	<b>Error! Bookmark not defined.</b>
HALAMAN PENGESAHAN .....	<b>Error! Bookmark not defined.</b>
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	vii
DAFTAR TABEL.....	viii
DAFTAR LAMPIRAN.....	ix
ABSTRAK.....	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	2
1.3 Batasan Masalah .....	3
1.4 Tujuan .....	3
1.5 Manfaat .....	3
2.1 Penelitian Terkait .....	5
2.2 Dasar Teori.....	6
2.2.1. Tindakan Medis .....	6
2.2.2. Pasien Dalam Pengawasan (PDP).....	6
2.2.3. Orang Dalam Pemantauan (ODP).....	7
2.2.4. Orang Tanpa Gejala (OTG) .....	7
2.2.5. <i>Naïve Bayes</i> .....	8
2.2.6. <i>K-Fold Cross Validation</i> .....	9
2.2.7. Cross Validation (CV) .....	9
2.2.8. Confusion Matrix .....	11
BAB III METODE PENELITIAN .....	13
3.1 Bahan Penelitian .....	13
3.2 Alat Penelitian.....	13
3.3 Prosedur Penelitian .....	13
3.4 Pengumpulan Data .....	14
3.5 Pengolahan Awal ( Preprocessing Data).....	14
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	16

4.1	Hasil Eksperimen dan Pengujian Metode .....	16
4.1.1	Proses Validasi.....	17
4.1.2	Metode <i>Naïve Bayes</i> .....	18
4.2	Evaluasi dan Validasi.....	19
4.3	Proses Pengujian dan Testing .....	19
4.4	Luaran Yang dicapai .....	20
BAB V .....		21
KESIMPULAN.....		21
5.1	Kesimpulan .....	21
5.2	Saran.....	21
DAFTAR PUSTAKA .....		22

## DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Ilustrasi metode Cross Validation menggunakan 3 fold .....	9
Gambar 2. 2 Ilustrasi 10 Fold Cross Validation .....	10
Gambar 4. 1 Proses Rapid Miner.....	16
Gambar 4. 2 Confusion Matrik .....	18

## DAFTAR TABEL

	Halaman
Tabel 2. 1 Model Confusion Matrix .....	12
Tabel 3. 1 Tabel variabel data gejala pasien .....	15
Tabel 4. 1 hasil proses validasi .....	17
Tabel 4. 2 <i>Class Precision</i> dan <i>Class Recall</i> .....	18
Tabel 4. 3 Nilai accuracy/parameter Naïve Bayes.....	19

## **DAFTAR LAMPIRAN**

Lampiran 1. Struktur organisasi peneliti

Lampiran 2. Realisasi Anggaran

Lampiran 3. Submit Jurnal

Lampiran 4. Draft Jurnal

Lampiran 5. Data Sheet

## ABSTRAK

Manajemen tindakan medis yang dilakukan dalam penanganan pasien yang ODP, OTG, PDP serta sudah positif Covid-19 dilakukan berdasarkan asumsi saja, seperti isolasi mandiri, rawat inap, atau dengan tindakan khusus di ICU. Kondisi tubuh pada setiap pasien berbeda-beda bisa jadi pasien dengan gejala yang sama namun penanganannya berbeda apa lagi pada pasien yang sudah lanjut usia. Banyak permasalahan yang terjadi dalam menentukan tindakan medis karena kondisi tubuh pasien berbeda. Sehingga perlu diangkat sebagai penelitian. Metode penelitian yang digunakan dalam penelitian ini adalah algoritma *Naive Bayes* dengan aplikasi pendukung Rapid Miner untuk melakukan proses pengujian data pasien sebanyak 500 data, 25 variabel atau gejala pasien dan 3 *output* sebagai bentuk tindakan medis. berdasarkan hasil analisis yang dilakukan pada penelitian ini, didapat hasil prediksi tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 dengan membandingkan data *training* dengan data *testing* menggunakan aplikasi Rapid Miner didapat tingkat akurasi sebesar 76.00%.

Kata kunci: *Pasien, Covid-19, Naive Bayes, Prediksi*

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

COVID-19 adalah penyakit yang menular, dan ditandai oleh gejala pada bagian pernapasan akut (*coronavirus 2 (severe acute respiratory syndrome coronavirus 2* atau SARS-CoV-2) [1]. Penyakit Coronavirus (COVID-19) telah meresahkan dunia khususnya Indonesia sebagai bencana non alamiah berupa [2]. Covid-19 telah ditetapkan sebagai pandemi global oleh WHO karena tingkat penyebarannya yang begitu cepat antar manusia. [3]. Penyebaran kasus COVID-19 yang terjadi secara cepat di 188 Negara [4]. Data terbaru World Health Organization mengenai COVID-19 secara global melaporkan bahwa sampai dengan tanggal 21 Oktober terdapat 350.424 kasus baru, 40.665.438 penderita dan 1.121.843 meninggal dunia [5][6]. Di Indonesia data BNPB berdasarkan data kementerian kesehatan sampai tanggal 21 Oktober 2020 penderita COVID-19 berjumlah 373.102 penderita, kasus aktif sebanyak 62.743 penderita dan meninggal sebanyak 12.857 penderita. Penderita COVID-19 terbanyak berada di DKI Jakarta (97.217 penderita), Jawa Barat (31.907 penderita), Jawa Timur (49.801 Penderita), Jawa Tengah (30.218 penderita), Sulawesi Selatan (17.690 penderita), Riau (12.319 penderita) dan Kalimantan timur (12.221 penderita) [7].

Covid-19 dapat menular ke semua orang yang memiliki kekebalan tubuh lemah. Namun yang paling beresiko terkomplikasi Covid-19 adalah orang yang lanjut usia, menderita penyakit komplikasi atau penyakit kronis. Resiko akan kematian yang disebabkan oleh Covid-19 sangat tinggi, terutama pada lanjut usia, kanker, memiliki diabetes, penyakit jantung, pembekuan darah atau yang telah menunjukkan tanda-tanda sepesis. Dengan tingkat kematian rata-rata 1%, tingkat kematian meningkat menjadi 6% pada orang dengan kanker, menderita hipertensi, atau penyakit pernapasan kronis, 7% untuk penderita diabetes, dan 10% pada penderita penyakit jantung. Sementara tingkat kematian di antara orang berusia 80 atau lebih beresiko 15% lebih tinggi [8].

Manajemen tindakan medis yang dilakukan dalam penanganan pasien yang ODP, OTG, PDP serta sudah positif Covid-19 dilakukan berdasarkan

asumsi saja, misalnya isolasi mandiri, rawat inap, atau dengan tindakan khusus di ICU. Kondisi tubuh pada setiap pasien berbeda-beda bias jadi pasien dengan gejala yang sama namun penanganannya berbeda apa lagi pada pasien yang sudah lanjut usia.

Berangkat dari permasalahan diatas, maka peneliti tertarik untuk melakukan klasifikasi prediksi guna mendapatkan informasi yang digunakan untuk memberikan rekomendasi kepada petugas kesehatan dalam memberikan tindakan yang tepat kepada pasien OTG, PDP, ODP atau yang dinyatakan positif Covid-19 berdasarkan usia. Sehingga nantinya tindakan yang diberikan akurat dan pasti.

Metode penelitian yang digunakan adalah klasifikasi dan prediksi. Banyak metode yang dapat digunakan diantaranya Regresi Logistik, *neural network*, *K-nearest Neighbor*, *Super Vector Machine*, *Naive Bayes*, *Decision Tree* dan *Random Forest* [9]. Namu pada penelitian kali ini akan menggunakan metode *Naive Bayes*.

Klasifikasi adalah teknik yang dilakukan untuk memprediksi *class* atau properti dari setiap *instance* data. Model prediksi memungkinkan untuk memprediksi nilai-nilai variabel yang tidak diketahui berdasarkan nilai variabel lainnya. Klasifikasi memetakan data ke dalam kelompok-kelompok kelas yang telah ditetapkan sebelumnya. Klasifikasi disebut juga dengan *supervised learning* karena kelas data telah ditentukan sebelumnya [10].

Penelitian *Naive Bayes* terkait penggunaan *Naive Bayes Classifier* telah banyak dilakukan. *Naive Bayes* memiliki beberapa kelebihan, yaitu cepat dalam perhitungan, algoritma yang sederhana dan berakurasi tinggi dan *Naive Bayes Classifier* lebih tepat diterapkan pada data yang besar dan dapat menangani data yang tidak lengkap (*missing value*) serta kuat terhadap atribut yang tidak relevan dan noise pada data [11].

## 1.2 Perumusan Masalah

Berdasarkan permasalahan diatas maka rumusan masalah penelitian ini adalah:

1. Menentukan parameter yang tepat dalam memprediksi tindakan medis untuk pasien OTG, OPD, PDP dan Positif Covid-19.
2. Mengoptimalkan pemilihan nilai parameternya sehingga akan berdampak pada tingkat akurasi yang dihasilkan.

### 1.3 Batasan Masalah

Sesuai dengan tujuan yang ingin dicapai, maka batasan permasalahan ini dibatasi dalam beberapa hal sebagai berikut :

1. Yang menjadi objek penelitian adalah pasien OTG, ODP, PDP dan Positif Covid-19 di Kota Tegal Tegal, Kabupaten Tegal dan Kabupaten Brebes
2. Metode yang digunakan adalah *Naive Bayes*
3. Yang menjadi *data traning* adalah data pasien OTG, ODP, PDP dan Positif Covid-19 di tiap rumah sakit di Kota Tegal, Kabupaten Tegal dan Kabupaten Brebes

### 1.4 Tujuan

Tujuan dari penelitian ini adalah diperolehnya model yang optimal dan akurat dalam memberikan tindakan medis yang tepat pada pasien OTG, ODP, PDP dan Positif Covid -19.

### 1.5 Manfaat

Manfaat dari penelitian ini dapat memberikan beberapa manfaat diantaranya adalah sebagai berikut :

1. Bagi Mahasiswa

Manfaat dari penelitian ini yaitu adanya pemahaman tentang algoritma *Naive Bayes* dan contoh penerapannya.

2. Bagi Pengembangan IPTEK

Hasil penelitian ini diharapkan dapat memberikan sumbangan berupa *improve* pengembangan model dengan menerapkan *Naive Bayes* sehingga hasil klasifikasi dan prediksi model ini menjadi lebih baik dan akurat.

### 3. Bagi Kebijakan Institusi

Hasil penelitian ini diharapkan dapat digunakan sebagai bahan pertimbangan bagi pemangku kebijakan untuk mengambil kebijakan dalam permasalahan yang terkait dengan tindakan medis.

## BAB II

### TINJAUAN PUSTAKA DAN DASAR TEORI

#### 2.1 Penelitian Terkait

Terkait jurnal penelitian yang dilakukan oleh Ela Nurelasari (2018) dengan judul penelitian “Penerapan Metode Metode *Neural Network* Berbasis Algoritma Genetika Untuk Prediksi Penyakit Kanker Payudara“. Pada penelitian ini menggunakan metode *Neural network* dapat mengatasi permasalahan tersebut, tetapi neural sebagai penentuan nilai parameter yang dioptimasi dengan Algoritma Genetika. Algoritma *Neural Network* berbasis algoritma genetika memiliki nilai akurasi yang lebih tinggi dibanding hanya menggunakan algoritma *Neural Network*. Hal ini terbukti dari peningkatan nilai akurasi untuk model algoritma *Neural Network* sebesar 95.42% dan nilai akurasi algoritma *Neural Network* berbasis Algoritma Genetika sebesar 96.85% dengan selisih akurasi 1.43%. Sehingga dapat disimpulkan bahwa penerapan teknik optimasi Algoritma Genetika dapat meningkatkan nilai akurasi pada algoritma *Neural Network* [12].

Penelitian lain yang dilakukan oleh Hilda Amalia, dkk (2019) dengan judul penelitiannya “Penerapan Feature Weighting Optimized Pada Naïve Bayes Untuk Prediksi Proses Persalinan” Dalam penelitian ini akan dilakukan klasifikasi proses persalinan ibu melahirkan dengan metode *data mining* yaitu *naive bayes* yang ditingkatkan kinerjanya dengan menggunakan metode *Optimize Weight* (PSO). *Naive Bayes* mampu menghasilkan nilai akurasi yang tinggi untuk pengolahan data persalinan ibu melahirkan yaitu sebesar 94%. Hasil akhir dari penelitian ini diperoleh nilai kinerja *naïve bayes* mampu ditingkatkan oleh metode *Optimize Weights* (PSO) menjadi lebih baik yakni 98% [13].

Pada penelitian lain yang dilakukan oleh Tutus Praningski & Indra Budi2 (2017) dengan judul penelitian “Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART, Naive Bayes, dan k-NN” Algoritma yang digunakan untuk melakukan klasifikasi penyakit kanker serviks adalah *Classification And Regression Trees* (CART), *Naive Bayes*, dan *k-Nearest Neighbor* (k-NN). Pengujian yang dilakukan terhadap algoritma *CART Decision Tree*, *Naive*

*Bayes*, dan k-NN, menggunakan formula *Confusion Matrix*, dengan menggunakan teknik pemecahan *dataset Holdout*. Hasil pengujian terhadap algoritma yang digunakan, menunjukkan algoritma *Naive Bayes* memiliki akurasi terbaik sebesar 94,44%, sedangkan tingkat akurasi yang dihasilkan algoritma CART dan k-NN adalah 88,89%, 85,04%. Performa yang didapatkan oleh masing-masing algoritma yang digunakan, memungkinkan penggunaan sistem prediksi penyakit kanker serviks untuk mendukung keputusan klinis pada pasien baru [11].

Dari beberapa penelitian terdahulu bahwa metode yang akurat untuk prediksi dan klasifikasi adalah metode *Naive Bayes* sehingga pada penelitian ini akan menggunakan metode *Naive Bayes* dalam memprediksi tindakan medis pada pasien OTG, ODP, PDP dan Positif & Covid

## 2.2 Dasar Teori

### 2.2.1. Tindakan Medis

Tindakan Medis Tindakan medis, yaitu suatu intervensi medis yang dilakukan pada seseorang berdasar atas indikasi medis tertentu yang dapat mengakibatkan integritas jaringan atau organ terganggu. Tindakan tersebut dapat berupa :

1. Tindakan terapeutik yang bertujuan untuk pengobatan
2. Tindakan diagnostik yang bertujuan untuk menegakan atau menetapkan diagnosis.
3. Tindakan medis hanya dapat dilakukan apabila telah dilakukan *informed consent*, yaitu persetujuan atau penolakan pasien yang bersangkutan terhadap tindakan medis yang akan diterimanya setelah memperoleh informasi lengkap tentang tindakan tersebut.

### 2.2.2. Pasien Dalam Pengawasan (PDP)

1. Orang dengan Infeksi Saluran Pernapasan Akut (ISPA) yaitu demam ( $\geq 38^{\circ}\text{C}$ ) atau riwayat demam; disertai salah satu gejala/ tanda penyakit pernapasan seperti: batuk/ sesak nafas/ sakit tenggorokan/ pilek/ pneumonia ringan hingga berat# dan tidak ada penyebab lain berdasarkan gambaran klinis yang meyakinkan dan pada 14 hari

terakhir sebelum timbul gejala memiliki riwayat perjalanan atau tinggal di negara/wilayah yang melaporkan transmisi lokal.

2. Orang dengan demam ( $\geq 38^{\circ}\text{C}$ ) atau riwayat demam atau ISPA dan pada 14 hari terakhir sebelum timbul gejala memiliki riwayat kontak dengan kasus konfirmasi COVID-19.
3. Orang dengan ISPA berat/pneumonia berat yang membutuhkan perawatan di rumah sakit dan tidak ada penyebab lain berdasarkan gambaran klinis yang meyakinkan.

### **2.2.3. Orang Dalam Pemantauan (ODP)**

1. Orang yang mengalami demam ( $\geq 38^{\circ}\text{C}$ ) atau riwayat demam; atau gejala gangguan sistem pernapasan seperti pilek/sakit tenggorokan/batuk dan tidak ada penyebab lain berdasarkan gambaran klinis yang meyakinkan dan pada 14 hari terakhir sebelum timbul gejala memiliki riwayat perjalanan atau tinggal di negara/wilayah yang melaporkan transmisi lokal.
2. Orang yang mengalami gejala gangguan sistem pernapasan seperti pilek/sakit tenggorokan/batuk dan pada 14 hari terakhir sebelum timbul gejala memiliki riwayat kontak dengan kasus konfirmasi COVID-19.

### **2.2.4. Orang Tanpa Gejala (OTG)**

Seseorang yang tidak bergejala dan memiliki risiko tertular dari orang konfirmasi COVID-19. Orang tanpa gejala (OTG) merupakan kontak erat dengan kasus konfirmasi COVID-19.

Kontak Erat adalah seseorang yang melakukan kontak fisik atau berada dalam ruangan atau berkunjung (dalam radius 1 meter dengan kasus pasien dalam pengawasan atau konfirmasi) dalam 2 hari sebelum kasus timbul gejala dan hingga 14 hari setelah kasus timbul gejala. Termasuk kontak erat adalah:

1. Petugas kesehatan yang memeriksa, merawat, mengantar dan membersihkan ruangan di tempat perawatan kasus tanpa menggunakan APD sesuai standar.

2. Orang yang berada dalam suatu ruangan yang sama dengan kasus (termasuk tempat kerja, kelas, rumah, acara besar) dalam 2 hari sebelum kasus timbul gejala dan hingga 14 hari setelah kasus timbul gejala.
3. Orang yang bepergian bersama (radius 1 meter) dengan segala jenis alat angkut/kendaraan dalam 2 hari sebelum kasus timbul gejala dan hingga 14 hari setelah kasus timbul gejala.

### 2.2.5. *Naïve Bayes*

Naïv Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistic yang ditemukan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes dengan asumsi independensi (ketidaktergantungan) yang kuat (naif). Dengan kata lain, dalam *Naïve Bayes*, model yang digunakan adalah model dengan fitur independen.

Algoritma Naïve Bayes sendiri memiliki dua tahap dalam proses klarifikasi, yaitu tahap pembelajaran dan tahap pengujian klasifikasi. Pada tahap pembelajaran dilakukan proses analisis terhadap sampel dokumen berupa pemeliharaan *vocabulary*, yaitu kata yang mungkin muncul dalam koleksi sampel dokumen yang sedapat mungkin dapat menjadi presentasi dokumen. Selanjutnya adalah penentuan probabilitas prior bagi kategori berdasarkan sampel dokumen. Pada tahap klasifikasi ditentukan nilai kategori dokumen berdasarkan term yang muncul dalam dokumen yang diklasifikasi.

Untuk melakukan perhitungan nilai kelas yang akan dibandingkan (*yes or no*) menggunakan algoritma *Naïve Bayes* maka dilakukan perhitungan probabilitas  $P(V_j)$  dengan rumus sebagai berikut :

$$P(P_j) = \frac{|doc_j|}{|Contoh|} \dots\dots\dots (1)$$

Docj adalah kebanyakan dikumen yang memiliki kategori j dalam pembelajaran, sedangkan contoh adalah banyaknya dokumen dalam contoh yang digunakan untuk pembelajaran. Untuk nilai  $P(W_k|V_j)$ , yaitu probabilitas kata  $W_k$  dalam kategori j, ditentukan dengan rumus :

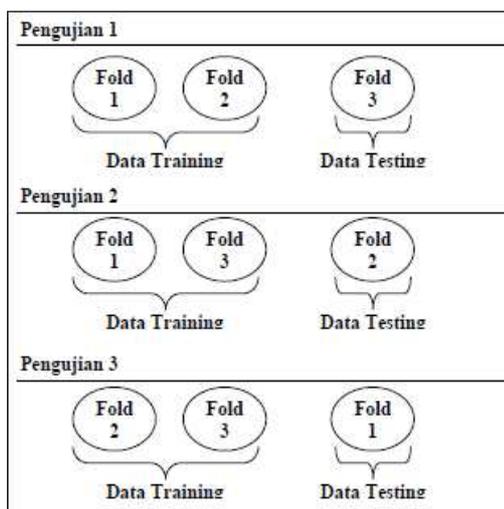
$$P(W_k|V_j) = \frac{N_k+1}{N+|vocabulary|} \dots \dots \dots (2)$$

$N_k$  adalah frekuensi munculnya kata  $W_k$  dalam dokumen yang berkategori  $V_j$  ditambah 1, hal ini berfungsi untuk menghindari angka nol dalam data atau bias disebut *Laplace Smoothing*, sedangkan nilai  $N$  adalah banyaknya seluruh data dalam dokumen berkategori  $V_j$ , dan *vocabulary* adalah banyaknya kata dalam pembelajaran.

### 2.2.6. K-Fold Cross Validation

Cross-validasi adalah metode statistik yang mengevaluasi dan membandingkan algoritma pembelajaran dengan membagi data menjadi dua yaitu data *training* dan data *testing*. Bentuk dari *cross validation* adalah *k-fold cross validation*.

Metode *k-fold cross validation* yang sering dipakai adalah *3-fold cross validation* dan *5-fold cross validation*. Dalam *cross-validation*, tentukan nilai *folds* atau partisi untuk data. Prinsip dari *k-fold cross validation* adalah membagi tiap kelompok data menjadi  $k$  bagian kelompok data yang selanjutnya data tersebut secara bergantian akan digunakan untuk *training* dan *testing* sejumlah  $k$  pengujian. Berikut gambar dari metode *Cross Validation*.



Gambar 2. 1 Ilustrasi metode Cross Validation menggunakan 3 fold

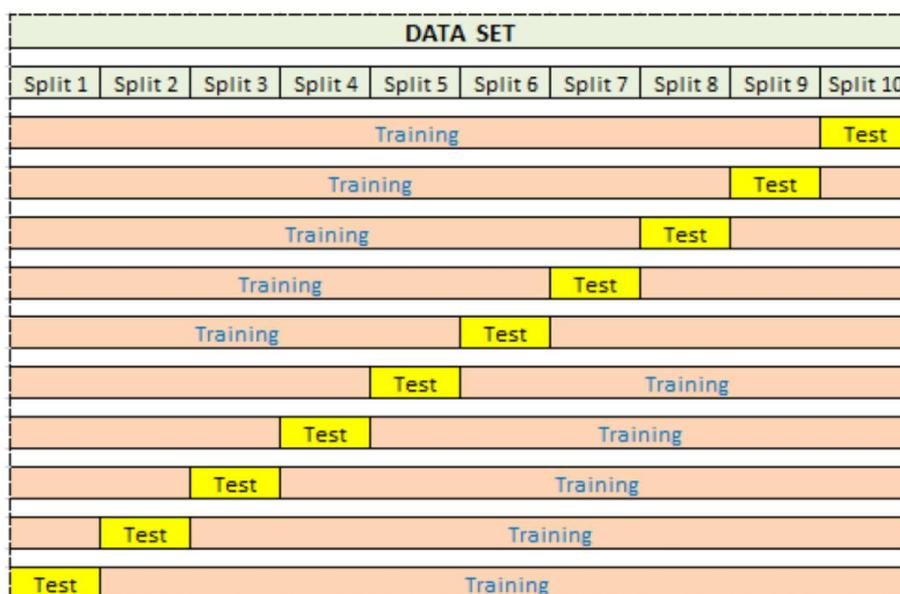
### 2.2.7. Cross Validation (CV)

*Cross validation* merupakan pengujian standar yang dilakukan untuk memprediksi *error rate*. *Cross validation* akan membagi *data*

*training* secara random ke dalam beberapa bagian dengan perbandingan yang sama kemudian *error rate* akan dihitung bagian demi bagian, kemudian menghitung rata-rata seluruh *error rate* untuk mendapatkan *error rate* keseluruhan.

Cross-Validasi (CV) tes ada di sejumlah varian tetapi ide umum adalah untuk membagi data pelatihan menjadi beberapa partisi atau lipatan. Classifier dievaluasi *oleh* akurasi klasifikasi lebih satu partisi setelah belajar dari yang lain. Prosedur ini kemudian diulang sampai semua partisi telah digunakan untuk evaluasi. Beberapa jenis yang paling umum adalah *10-fold*, *n-fold* dan CV bootstrap.

Perbedaan antara ketiga jenis CV terletak pada cara data yang dipartisi. Tinggalkan-satu-out mirip dengan-n kali lipat CV, di mana n adalah singkatan dari jumlah *kasus* pada set data. *Leave-one-out* atau *n-fold* CV dilakukan dengan meninggalkan satu contoh untuk pengujian dan pelatihan tentang kasus lain. Prosedur ini *kemudian* dilakukan sampai semua kasus telah ditinggalkan sekali. Telah dikemukakan bahwa desain *10-fold cross-validation* memperkenalkan sumber korelasi karena salah menggunakan contoh untuk pelatihan dalam satu percobaan dan untuk pengujian di tempat lain.



Gambar 2. 2 Ilustrasi 10 Fold Cross Validation

Pada gambar 2.2 terlihat bahwa tiap percobaan akan menggunakan satu data *testing* dan  $k-1$  bagian akan menjadi data *training*, kemudian data *testing* itu akan ditukar dengan satu buah data *training* sehingga untuk tiap percobaan akan didapatkan data *testing* yang berbeda-beda.

### 2.2.8. Confusion Matrix

Evaluasi dengan *confusion matrix* menghasilkan nilai *accuracy*, *precision*, dan *recall*. *Accuracy* dalam klasifikasi adalah persentase ketepatan *record data* yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Sedangkan *precision* atau *confidence* adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. *Recall* atau *sensitivity* adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar, Powers[37].

*True Positive* adalah jumlah *record* positif yang diklasifikasikan sebagai *positive*, *false positive* adalah jumlah *record negative* yang diklasifikasikan sebagai *positive*, *false negative* adalah jumlah *record positive* yang diklasifikasikan sebagai *negative*, *true negative* adalah jumlah *record negative* yang diklasifikasikan sebagai *negative*, kemudian masukkan data uji. Setelah data uji dimasukkan ke dalam *confusion matrix*, hitung nilai-nilai yang telah dimasukkan tersebut untuk dihitung jumlah *sensitivity (recall)*, *Specifity*, *precision*, dan *accuracy*. *Sensitivity* digunakan untuk membandingkan jumlah  $t_{pos}$  terhadap jumlah *record* yang positif sedangkan *Specifity*, *precision* adalah perbandingan jumlah  $t_{neg}$  terhadap jumlah *record* yang *negative*. Untuk menghitung digunakan persamaan dibawah ini, Han dan Kamber.

Keterangan :

- $t_{pos}$  : Jumlah *true positives*
- $t_{neg}$  : Jumlah *true negative*
- pos : Jumlah *record positives*
- neg : Jumlah *tupel negatives*
- $f_{pos}$  : Jumlah *false positives*
- $f_{neg}$  : Jumlah *false negatives*

Tabel 2. 1 Model Confusion Matrix

<i>Correct Classification</i>	<i>Classified as</i>	
	+	-
+	<i>True Positive (t<sub>pos</sub>)</i>	<i>False Negative (f<sub>neg</sub>)</i>
-	<i>False Positive (f<sub>pos</sub>)</i>	<i>True Negative (t<sub>neg</sub>)</i>

Berikut adalah persamaan model *confusion matrix*:

1. Nilai akurasi (acc) adalah proporsi jumlah prediksi yang benar. Dapat dihitung dengan menggunakan persamaan:

$$\text{Accuracy} = \text{Sensitivity} \frac{pos}{(pos+neg)} + \text{Specity} \frac{neg}{(pos+neg)} \dots\dots (1)$$

2. Sensitivity digunakan untuk membandingkan proporsi  $t_{pos}$  terhadap tupel yang positif, yang dihitung dengan menggunakan persamaan:

$$\text{Sensitifity} = \frac{t_{pos}}{pos} \dots\dots\dots (2)$$

3. Specifity digunakan untuk membandingan proporsi  $t_{neg}$  terhadap tupel yang negatif, yang dihitung dengan menggunakan persamaan:

$$\text{Specifity} = \frac{t_{neg}}{neg} \dots\dots\dots (3)$$

4. PPV (*Positive Predictive Value*) adalah proporsi kasus dengan hasil diagnosa positif, yang dihitung dengan menggunakan persamaan:

$$PPV = \frac{t_{pos}}{t_{pos} + f_{pos}} \dots\dots\dots (4)$$

5. NPV (*Negative Predictive Value*) adalah proporsi kasus dengan hasil diagnosa negatif, yang dihitung dengan menggunakan persamaan:

$$NPV = \frac{t_{neg}}{t_{neg} + f_{neg}} \dots\dots\dots (5)$$

Confusion matrix memberikan keputusan yang diperoleh dalam *training* dan *testing*, confusion matrix memberikan penilaian *performance* klasifikasi berdasarkan objek dengan benar atau salah.

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Bahan Penelitian**

Bahan penelitian yang digunakan adalah data pasien covid yang didapat dari rumah sakit, diantaranya: RSUD Dr. Soesilo, RSUD Suradadi, Rumah Sakit DKT Pagongan, RSI PKU Muhammadiyah, RS Harapan Sehat, RSUD Kardinah, RS Mitra Siaga serta RSUI Harapan Anda untuk data pasien Covid-19.

#### **3.2 Alat Penelitian**

Adapun alat bantu yang digunakan untuk dalam pengumpulan data penelitian ini adalah menggunakan observasi.

#### **3.3 Prosedur Penelitian**

Penelitian ini menggunakan jenis penelitian eksperimen, dengan tahapan penelitian sebagai berikut:

1. Pengumpulan Data

Pengumpulan data merupakan langkah awal pada suatu penelitian. Data yang digunakan pada penelitian ini adalah data rekamedik pasien OTG, PDP, ODP dan Positif Covid-19 sepanjang tahun 2020 dan 2021.

2. Pembersih Data

Pada tahap ini akan dilakukan pembersihan data yang tidak lengkap, kosong (*null*), *noise*, dan data yang tidak konsisten. Dalam langkah ini, data yang bernilai kosong (*null*), akan dibersihkan dengan cara dihapus secara manual, dan akan dilakukan penghapusan atribut atau mengganti data tersebut

3. Seleksi Data

Pada tahap ini, akan dilakukan penyelesaian data untuk mengurangi data yang tidak relevan, dan berlebihan (*redundant*). Menurut Tan, Steinbach, dan Kumar (2006), atribut yang tidak relevan adalah atribut yang berisi informasi yang tidak berguna untuk melakukan penambangan data, sedangkan atribut yang berlebihan (*redundant*) adalah atribut yang

menduplikasikan banyak atau semua informasi yang terdapat di dalam satu atau lebih pada atribut lainnya

#### 4. Pengolahan Awal

Pengolahan awal (*Preprocessing*) merupakan tahap untuk mempersiapkan data yang telah diperoleh dari tahap pengumpulan data, yang akan digunakan pada tahap selanjutnya.

#### 5. Metode yang Diusulkan

Tahapan ini akan membahas metode yang akan digunakan pada penelitian. Metode yang digunakan adalah *Naïve Bayes*

#### 6. Eksperimen dan Pengujian

Tahapan ini akan membahas tahapan penelitian dan teknik pengujian yang akan digunakan.

#### 7. Evaluasi dan Validasi Penelitian

Tahapan ini akan membahas hasil evaluasi dari eksperimen yang telah digunakan, menggunakan *K-Fold Cross Validation*.

### 3.4 Pengumpulan Data

Dataset yang digunakan pada penelitian ini adalah berupa data nilai pasien PDP, ODP, OTG dan Positif Covid-19. Data pasien diambil per tanggal 16 Juli 2021 dari rumah sakit RSUD Dr. Soesilo, RSUD Suradadi, Rumah Sakit DKT Pagongan, RSI PKU Muhammadiyah, RS Harapan Sehat, RSUD Kardinah, RS Mitra Siaga serta RSUI Harapan Anda terkait data pasien Covid-19.

### 3.5 Pengolahan Awal ( Preprocessing Data)

Pada tahapan *preprocessing* data, untuk mempermudah penggunaan data karena variabel yang akan digunakan dalam penelitian ini sebanyak 25 variabel gejala dan 3 label yang dijadikan *output*. Yang diambil dari nilai data pasien dari rumah sakit, jurnal serta hasil wawancara dengan variabel seperti dibawah ini:

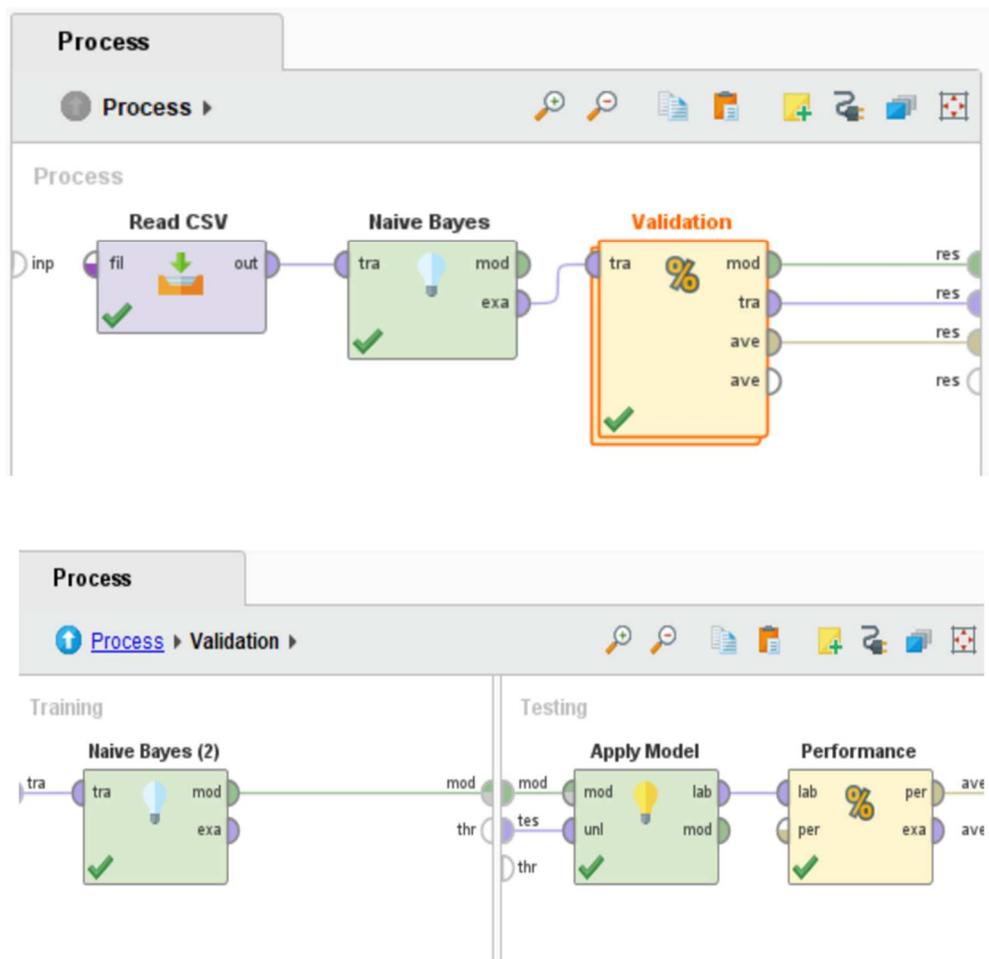
Tabel 3. 1 Tabel variabel data gejala pasien

<b>Nama</b>	<b>Keterangan</b>
X1	Demam di atas 30 derajat celcius
X2	Riwayat demam
X3	Batuk
X4	Pilek
X5	Sakit Tenggorokan
X6	Sesak Nafas
X7	Kedinginan yang menggigil
X8	Sakit kepala
X9	Mudah lemah
X10	Nyeri otot
X11	Mual-mual
X12	Muntah-muntah
X13	Nyeri Abdomen
X14	Diare
X15	Gejala lain
X16	Memiliki riwayat perjalanan penyakit lain
X17	Memiliki kontak erat dengan kasus suspek Covid-19
X18	Memiliki kontak erat dengan kasus konfirmasi covid 19
X19	Hilang indra pencuman
X20	ISPA berat
X21	Hilang konsentrasi
X22	Demam terus menerus
X23	Umur
X24	Jenis kelamin
X25	Kota/Kabupaten asal
Y	Isoman, Isolasi Rumah Sakit, ICU

## BAB IV HASIL PENELITIAN DAN PEMBAHASAN

### 4.1 Hasil Eksperimen dan Pengujian Metode

Untuk mendapatkan model yang sesuai dengan yang diharapkan, pada tahapan eksperimen dilakukan dengan melakukan percobaan dengan Rapid Miner seperti contoh dibawah ini :



Gambar 4. 1 Proses Rapid Miner

Setelah itu dilakukan pengujian menggunakan tools Rapid Miner, kemudian :

1. Menambahkan dua operator read excel untuk memasukan data *training* dan data *testing*
2. Menambahkan operator *naive bayes* yang digunakan sebagai metode dalam penelitian.

3. Menambahkan operator *split validation* digunakan untuk proses pengolahan data. Yang didalamnya terdapat operator *Naïve Bayes*, *Apply Model* dan *Performance* untuk menentukan hasil prediksinya.

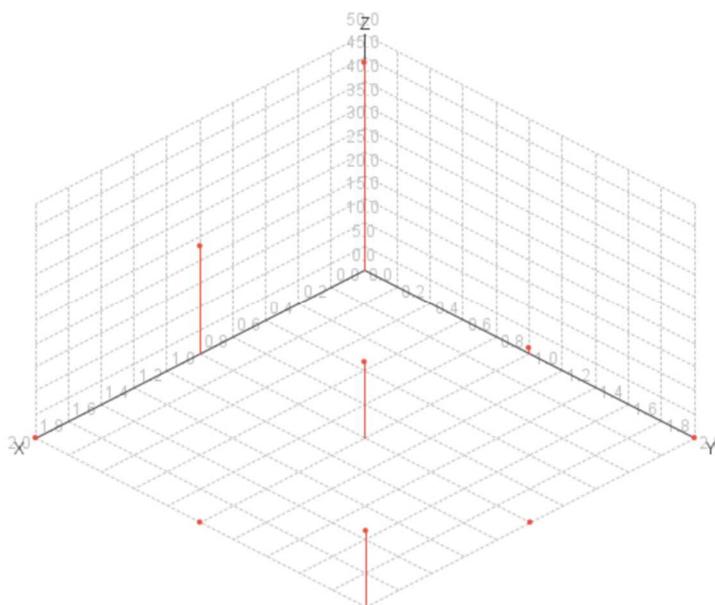
#### 4.1.1 Proses Validasi

Pada tahapan validasi ini menggunakan data responden sebanyak 500 pasien, 25 Variabel *input* dan 3 *output* dengan *split relative*, *split ratio* dari 0.1-0.9, *sampling type automatic* didapat hasil seperti tabel dibawah ini :

Tabel 4. 1 hasil proses validasi

No	Split	Sampling Type	Split Ratio	Accuracy
1	<i>relative</i>	<i>automatic</i>	0.1	70.44%
2	<i>relative</i>	<i>automatic</i>	0.2	73.50%
3	<i>relative</i>	<i>automatic</i>	0.3	73.14%
4	<i>relative</i>	<i>automatic</i>	0.4	74.25%
5	<i>relative</i>	<i>automatic</i>	0.5	63.60%
6	<i>relative</i>	<i>automatic</i>	0.6	73.13%
7	<i>relative</i>	<i>automatic</i>	0.7	72.67%
<b>8</b>	<b><i>relative</i></b>	<b><i>automatic</i></b>	<b>0.8</b>	<b>76.00%</b>
9	<i>relative</i>	<i>automatic</i>	0.9	74.00%

Dari beberapa percobaan yang dilakukan mendapatkan akurasi yang lebih baik adalah 76.00% dengan split ratio 0.8 dengan bentuk *Confusion Matrik* seperti dibawah ini :



Gambar 4. 2 Confusion Matrik

#### 4.1.2 Metode *Naïve Bayes*

Proses klasifikasi dengan Rapid Miner menggunakan metode *Naïve Bayes* yang digunakan untuk memprediksi/klasifikasi tindakan pasien sehingga didapat akurasi hasil prediksi 76.00% dari hasil data *testing*. Pengujian berdasarkan *Confusion Matrix* menghasilkan nilai *accuracy*, *precision* dan *recall* yang tinggi dengan nilai *accuracy* sebesar 76.00%. dengan *Class precision* untuk Pred. Isoman 65.67%, Pred. Isolasi rumah sakit 94.12%, Pred. ICU 100% sedangkan *Class Recall* untuk *True* Isoman 97.78%, *True*. Isolasi rumah sakit 41.03%, *True* ICU 100%. Dengan tabel dibawah ini:

Tabel 4. 2 *Class Precision* dan *Class Recall*

	true Isoman	true Isolasi rumah sakit	true ICU	class precision
pred. Isoman	44	23	0	65.67%
pred. Isolasi rumah sakit	1	16	0	94.12%
pred. ICU	0	0	16	100.00%
class recall	97.78%	41.03%	100.00%	

## 4.2 Evaluasi dan Validasi

Setelah bebrapa kombinasi nilai parameter yang telah didapatkan, maka setelah divalidasi pada model yang telah digunakan maka didapatkan hasil nilai accuracy atau parameter terbaik

Tabel 4. 3 Nilai accuracy/parameter Naïve Bayes

Split Ratio	= 0.8
accuracy	= 76.00%

## 4.3 Proses Pengujian dan Testing

Pengujian dan testing dilakukan untuk melihat sejauh mana model dari yang dibuat dapat berjalan pada dataset yang dijadikan untuk testing, hasil analisa antara data *testing* dengan data *training* pada Rapid Miner dapat dilihat pada Tabel 4.2. Untuk menghitung akurasinya sebagai berikut :

Jumlah data yang di uji	: 500
Jumlah data yang diprediksi benar “Isoman”	: 44
Jumlah data yang diprediksi benar “Isolasi rumah sakit”	: 16
Jumlah data yang diprediksi benar “ICU”	: 16
Jumlah data yang diprediksi salah “Isoman”	: 0
Jumlah data yang diprediksi salah “Isolasi rumah sakit”	: 23
Jumlah data yang diprediksi salah “ICU”	: 0

$$\begin{aligned}
 \text{Akurasi} &= \frac{44+16+16}{44+16+16+1+0+23} \times 100\% \\
 &= \frac{76}{100} \times 100\% \\
 &= 76,00\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Eror} &= \frac{23+0+1}{44+16+16+1+0+23} \times 100\% \\
 &= \frac{24}{100} \times 100\% \\
 &= 24,00\%
 \end{aligned}$$

Dari perhitungan tersebut disimpulkan bahwa prediksi dengan menggunakan metode *Naïve Bayes* untuk menentukan tindakan medis pada

pasien ODP, PDP, OTG dan Positif Covid-19 menghasilkan tingkat akurasi sebesar 76.00% dan tingkat eror 24.00%.

#### **4.4 Luaran Yang dicapai**

Dari penelitian ini akan dipublikasikan di jurnal ELKOM STEKOM Semarang pada bulan Desember 2021.

## **BAB V**

### **KESIMPULAN**

#### **5.1 Kesimpulan**

Dari penelitian yang dilakukan mengenai penerapan metode *Naive Bayes* untuk prediksi tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 dapat disimpulkan bahwa berdasarkan hasil analisis yang dilakukan pada penelitian ini, didapat hasil prediksi tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 dengan membandingkan data *training* dengan dengan data *testing* menggunakan aplikasi Rapid Miner didapat tingkat akurasi sebesar 76.00%.

#### **5.2 Saran**

Penelitian yang dilakukan tentunya tak lepas dari sebuah kekurangan. Oleh karena itu, ada beberapa saran yang penulis berikan yaitu Jenis pengelompokan data dalam penelitian ini bersifat kategorik, untuk hasil yang lebih optimal mungkin untuk peneliti selanjutnya dapat menambahkan dengan data numerik.

## DAFTAR PUSTAKA

- [1] T. Dey and A. Sinha, "Ethnicity and COVID-19 - A commentary on 'World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)' (Int J Surg 2020;76:71-6)," *Int. J. Surg.*, vol. 83, no. August, pp. 75–76, 2020, doi: 10.1016/j.ijisu.2020.08.046.
- [2] W. Wiguna and D. Riana, "Diagnosis of Coronavirus Disease 2019 (Covid-19) Surveillance Using C4.5 Algorithm," *J. Pilar Nusa Mandiri*, vol. 16, no. 1, pp. 71–80, 2020, doi: 10.33480/pilar.v16i1.1293.
- [3] B. Etikasari, T. D. Puspitasari, A. A. Kurniasari, and L. Perdanasari, "Sistem informasi deteksi dini Covid-19," *J. Tek. Elektro dan Komput.*, vol. 9, no. 2, pp. 101–108, 2020.
- [4] L. Wynants *et al.*, "Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal," *BMJ*, vol. 369, 2020, doi: 10.1136/bmj.m1328.
- [5] M. Pradana, S. Syahputra, A. Wardhana, B. R. Kartawinata, and C. Wijayangka, "The Effects of Incriminating COVID-19 News on the Returning Indonesians' Anxiety," *J. Loss Trauma*, vol. 25, no. 8, pp. 1–6, 2020, doi: 10.1080/15325024.2020.1771825.
- [6] World Health Organization (WHO) Indonesia, "Coronavirus Disease 2019 (COVID-19) WHO Indonesia Situation Report," *Indones. Situat. Reports*, vol. 2019, no. April, pp. 1–9, 2020, [Online]. Available: <https://www.who.int/indonesia/news/novel-coronavirus/situation-reports>.
- [7] National Disaster Management Agency, "Head of BNPB Decision Letter No. 9A/2020." p. 3, 2020.
- [8] N. Anung Ahadi Pradana, Casman, "Pengaruh kebijakan," *Pengaruh Kebijak. Soc. Distancing Pada Wabah Covid-19 Terhadap Kelompok Rentan Di Indones.*, vol. 09, no. 02, pp. 61–67, 2020.
- [9] F. Sodik, B. Dwi, and I. Kharisudin, "Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python," vol. 3, pp. 689–694, 2020.
- [10] M. Informatika and K. B. Aceh, "ANALISIS TREND TOPIK PENELITIAN PADA WEB OF SCIENCE DAN SINTA UNTUK PENENTUAN TEMA," pp. 13–25, 2018.
- [11] T. Praningki and I. Budi, "Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART, Naive Bayes, dan k-NN," *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 83, 2018, doi: 10.24076/citec.2017v4i2.100.
- [12] E. Nurelasari, "Penerapan Metode Metode Neural Network Berbasis Algoritma Genetika Untuk Prediksi Penyakit Kanker Payudara," *Speed – Sentra Penelit. Eng. dan Edukasi*, vol. 10, no. 3, pp. 1–11, 2018.
- [13] H. Amalia, A. B. Pohan, and S. Masripah, "Penerapan Feature Weighting

Optimized Pada Naïve Bayes Untuk Prediksi Proses Persalinan,” *J. Pilar Nusa Mandiri*, vol. 15, no. 1, pp. 15–20, 2019, doi: 10.33480/pilar.v15i1.3.

## STRUKTUR ORGANISASI PENGUSUL

### 1. Ketua

Nama : Arfan Haqiqi Sulasmoro, M.Kom  
NIPY : 02.009.054  
NIDN : 0623037704  
Pangkat/ Golongan : IIIB  
Jabatan Fungsional : Asisten Ahli  
Jabatan Struktural : Dosen Tetap  
Bidang Ilmu : Teknik Komputer  
Pengalaman Penelitian Administrasi : a. Pengembangan Aplikasi Sistem Program Studi DIII Teknik Komputer Politeknik Harapan Bersama Sebagai Optimalisasi Pelayanan Prodi  
b. Jemuran Ikan Asin Otomatis Berbasis Internet Of Things Untuk Daerah Pesisir Pantai Pantura  
c. Rancang Bangun E-Commerce Sebagai Sarana Peningkatan Perekonomian Bagi Industri Kecil Dan Menengah Di Kabupaten Brebes

### 2. Anggota1

Nama : Rais, S.Pd, M.Kom  
NIPY : 07.011.083  
NIDN : 0614108501  
Pangkat Golongan: : 3B  
Jabatan Fungsional : Asisten Ahli  
Jabatan Struktural : KA. Prodi  
Bidang Ilmu : Teknik Komputer  
Pengalaman Penelitian : a. Rancang Bangun *Website* Pengajuan Kerja Praktek Dan Tugas Akhir Di Prodi DIII Teknik Komputer  
b. Rancang Bangun E-Commerce Sebagai Sarana Peningkatan Perekonomian Bagi Industri Kecil Dan Menengah Di Kabupaten Brebes

### 3. Anggota2

Nama : Istiqomah Dwi Andari, S.ST., M.Kes  
NIPY : 02.013.140  
Pangkat Golongan: : 3A  
Jabatan Fungsional : -  
Jabatan Struktural : Kasubag PKM  
Bidang Ilmu : Kebidanan  
Pengalaman Penelitian : Pengaruh status gizi dengan prestasi belajar anak sekolah dasar di MI Bina Amanah Kabupaten

## Realisas Anggaran

No	Kebutuhan	Jumlah	Satuan	Harga	Total
A. Kosumsi					
1	Makan siang	20	Kali	30.000	600.000
Sub Total A					<b>600.000</b>
B. Belanja Operasional Lainnya					
1	Fotocopy & jilid	6	Eksemplar	50.000	300.000
2	Box file bindex folio	1	Buah	45.000	45.000
3	Keretas A4	3	Buah	55.000	165.000
4	Joyko binder clipp	8	Buah	10.000	80.000
5	Materai 10.000	5	Buah	12.000	60.000
Sub Total B					<b>650.000</b>
C. Perjalanan					
1	Observasi Lapangan	3	Hari	100.000	300.000
2	Pengambilan Data	4	Hari	100.000	400.000
Sub Total C					700.000
D. Honorarium					
1	Petugas Pengumpulan data	20	Jam	20.000	400.000
1	Petugas Analisis Data	20	Jam	20.000	400.000
Sub Total D					<b>800.000</b>
D. Publikasi					
1	Jurnal Lokal	1		500.000	500.000
Sub Total D					<b>500.000</b>
<b>Total Pengeluaran (A+B+C+D)</b>					<b>3.250.000</b>

# Submit Jurnal

Elkom : Jurnal Elektronika dan Komputer Tasks 0 English View Site rais

**Submit an Article**

1. Start 2. Upload Submission 3. Enter Metadata 4. Confirmation 5. Next Steps

**Submission Files** [Search](#) [Upload File](#)

 951-1	rais, Draff_Jurnal_Penelitian_Rais_Politeknik_Tegal.doc	August 12, 2021	Article Text
---	---	-----------------	--------------

[Save and continue](#) [Cancel](#)

# Prediction Of Medical Actions For Covid Patients Using Naïve Bayes Method

Arfan Haqiqi<sup>1</sup>, Rais<sup>2</sup>, Istiqomah Dwi Andari<sup>3</sup>, Siti Fatimah<sup>4</sup>

<sup>1</sup>Program Studi DIII Teknik Komputer, Politeknik Harapan bersama  
Jl. Mataram No. 09, Pesurungan Lor, Kec. Margadana, Kota Tegal, Indonesia 52142  
e-mail : [Arfan.hqq@gmail.com](mailto:Arfan.hqq@gmail.com)

<sup>2</sup>Program Studi DIII Teknik Komputer, Politeknik Harapan bersama  
Jl. Mataram No. 09, Pesurungan Lor, Kec. Margadana, Kota Tegal, Indonesia 52142  
e-mail : [rais@poltektegal.ac.id](mailto:rais@poltektegal.ac.id)

<sup>3</sup>Program Studi DIII Kebidanan, Politeknik Harapan bersama  
Jl. Mataram No. 09, Pesurungan Lor, Kec. Margadana, Kota Tegal, Indonesia 52142  
e-mail : [istyandari44@gmail.com](mailto:istyandari44@gmail.com)

<sup>4</sup>Program Studi DIII Kebidanan, STIKES BREBES  
Jl. Jatibarang KM 8 Janegara Kec. Jatibarang Kab. Brebes, Indonesia 52261  
e-mail : [helga.abhinaya@gmail.com](mailto:helga.abhinaya@gmail.com)

## ARTICLE INFO

Article history:

Received 30 Mei 2020

Received in revised form 2 Juni 2020

Accepted 10 Juni 2020

Available online 12 Juni 2020

## ABSTRACT

Management of medical actions carried out in handling patients who are ODP (people under monitoring), OTG (asymptomatic people), PDP (patient under monitoring) and positive Covid-19 patients is carried out based on assumptions, such as self-isolation, hospitalization, or special treatments in the ICU (Intensive Care Unit) room. The condition of the body in each patient is different, a patient may have same symptoms but the treatment is different, especially in elderly patients. Many problems occur in determining medical action because the patient's body condition is different. Therefore, it needs to be appointed as a research. The research method used in this study was Naive Bayes algorithm with supporting application Rapid Miner. It was applied to carry out the process of testing on patient data as much as 500 data, 25 variables or patient symptoms and 3 outputs as a form of medical action. Based on the results of the analysis carried out in this study, prediction of medical actions for ODP, PDP, OTG and positive Covid-19 patients were obtained by comparing training data with testing data using Rapid Miner application. It resulted that an accuracy rate of 76.00% was obtained.

**Keywords:** *Pasien, Covid-19, Naïve Bayes, Prdiksi*

## 1. Introduction

COVID-19 adalah penyakit yang menular, dan ditandai oleh gejala pada bagian pernapasan akut (coronavirus 2 (*severe acute respiratory syndrome coronavirus 2* atau SARS-CoV-2) [1]. Penyakit Coronavirus (COVID-19) telah meresahkan dunia khususnya Indonesia sebagai bencana non alamiah berupa [2]. Covid-19 telah ditetapkan sebagai pandemi global oleh WHO karena tingkat penyebarannya yang begitu cepat antar manusia. [3]. Penyebaran kasus COVID-19 yang terjadi secara cepat di 188 Negara [4]. Data terbaru World Health Organization mengenai COVID-

*Received Mei 23, 2021; Revised Juni 29, 2021; Accepted Juli 12, 2021*

19 secara global melaporkan bahwa sampai dengan tanggal 21 Oktober terdapat 350.424 kasus baru, 40.665.438 penderita dan 1.121.843 meninggal dunia [5][6]. Di Indonesia data BNPB berdasarkan data kementerian kesehatan sampai tanggal 21 Oktober 2020 penderita COVID-19 berjumlah 373.102 penderita, kasus aktif sebanyak 62.743 penderita dan meninggal sebanyak 12.857 penderita. Penderita COVID-19 terbanyak berada di DKI Jakarta (97.217 penderita), Jawa Barat (31.907 penderita), Jawa Timur (49.801 Penderita), Jawa Tengah (30.218 penderita), Sulawesi Selatan (17.690 penderita), Riau (12.319 penderita) dan Kalimantan timur (12.221 penderita) [7].

Covid-19 dapat menular ke semua orang yang memiliki kekebalan tubuh lemah. Namun yang paling beresiko terkomplikasi Covid-19 adalah orang yang lanjut usia, menderita penyakit komplikasi atau penyakit kronis. Resiko akan kematian yang disebabkan oleh Covid-19 sangat tinggi, terutama pada lanjut usia, kanker, memiliki diabetes, penyakit jantung, pembekuan darah atau yang telah menunjukkan tanda-tanda sepesis. Dengan tingkat kematian rata-rata 1%, tingkat kematian meningkat menjadi 6% pada orang dengan kanker, menderita hipertensi, atau penyakit pernapasan kronis, 7% untuk penderita diabetes, dan 10% pada penderita penyakit jantung. Sementara tingkat kematian di antara orang berusia 80 atau lebih beresiko 15% lebih tinggi [8].

Manajemen tindakan medis yang dilakukan dalam penanganan pasien yang ODP, OTG, PDP serta sudah positif Covid-19 dilakukan berdasarkan asumsi saja, misalnya isolasi mandiri, rawat inap, atau dengan tindakan khusus di ICU. Kondisi tubuh pada setiap pasien berbeda-beda bias jadi pasien dengan gejala yang sama namun penanganannya berbeda apa lagi pada pasien yang sudah lanjut usia.

Berangkat dari permasalahan diatas, maka peneliti tertarik untuk melakukan klasifikasi prediksi guna mendapatkan informasi yang digunakan untuk memberikan rekomendasi kepada petugas kesehatan dalam memberikan tindakan yang tepat kepada pasien OTG, PDP, ODP atau yang dinyatakan positif Covid-19 berdasarkan usia. Sehingga nantinya tindakan yang diberikan akurat dan pasti.

Metode penelitian yang digunakan adalah klasifikasi dan prediksi. Banyak metode yang dapat digunakan diantaranya Regresi Logistik, *neural network*, *K-nearest Neighbor*, *Super Vector Machine*, *Naive Bayes*, *Decision Tree* dan *Random Forest* [9]. Namun pada penelitian kali ini akan menggunakan metode *Naive Bayes*.

Klasifikasi adalah teknik yang dilakukan untuk memprediksi *class* atau properti dari setiap *instance* data. Model prediksi memungkinkan untuk memprediksi nilai-nilai variabel yang tidak diketahui berdasarkan nilai variabel lainnya. Klasifikasi memetakan data ke dalam kelompok-kelompok kelas yang telah ditetapkan sebelumnya. Klasifikasi disebut juga dengan *supervised learning* karena kelas data telah ditentukan sebelumnya [10].

Penelitian *Naive Bayes* terkait penggunaan *Naive Bayes Classifier* telah banyak dilakukan. *Naive Bayes* memiliki beberapa kelebihan, yaitu cepat dalam perhitungan, algoritma yang sederhana dan berakurasi tinggi dan *Naive Bayes Classifier* lebih tepat diterapkan pada data yang besar dan dapat menangani data yang tidak lengkap (*missing value*) serta kuat terhadap atribut yang tidak relevan dan noise pada data [11].

## 2. Research Method

### 3.6 Bahan Penelitian

Bahan penelitian yang digunakan adalah data pasien covid yang didapat dari rumah sakit, diantaranya: RSUD Dr. Soesilo, RSUD Suradadi, Rumah Sakit DKT Pagongan, RSI PKU

Muhammadiyah, RS Harapan Sehat, RSUD Kardinah, RS Mitra Siaga serta RSUI Harapan Anda untuk data pasien Covid-19.

### 3.7 Alat Penelitian

Adapun alat bantu yang digunakan untuk dalam pengumpulan data penelitian ini adalah menggunakan observasi.

### 3.8 Prosedur Penelitian

Penelitian ini menggunakan jenis penelitian eksperimen, dengan tahapan penelitian sebagai berikut:

#### 8. Pengumpulan Data

Pengumpulan data merupakan langkah awal pada suatu penelitian. Data yang digunakan pada penelitian ini adalah data rekamedik pasien OTG, PDP, ODP dan Positif Covid-19 sepanjang tahun 2020 dan 2021.

#### 9. Pembersih Data

Pada tahap ini akan dilakukan pembersihan data yang tidak lengkap, kosong (*null*), *noise*, dan data yang tidak konsisten. Dalam langkah ini, data yang bernilai kosong (*null*), akan dibersihkan dengan cara dihapus secara manual, dan akan dilakukan penghapusan atribut atau mengganti data tersebut

#### 10. Seleksi Data

Pada tahap ini, akan dilakukan penyelesaian data untuk mengurangi data yang tidak relevan, dan berlebihan (*redundant*). Menurut Tan, Steinbach, dan Kumar (2006), atribut yang tidak relevan adalah atribut yang berisi informasi yang tidak berguna untuk melakukan penambangan data, sedangkan atribut yang berlebihan (*redundant*) adalah atribut yang menduplikasikan banyak atau semua informasi yang terdapat di dalam satu atau lebih pada atribut lainnya

#### 11. Pengolahan Awal

Pengolahan awal (*Preprocessing*) merupakan tahap untuk mempersiapkan data yang telah diperoleh dari tahap pengumpulan data, yang akan digunakan pada tahap selanjutnya.

#### 12. Metode yang Diusulkan

Tahapan ini akan membahas metode yang akan digunakan pada penelitian. Metode yang digunakan adalah *Naïve Bayes*

#### 13. Eksperimen dan Pengujian

Tahapan ini akan membahas tahapan penelitian dan teknik pengujian yang akan digunakan.

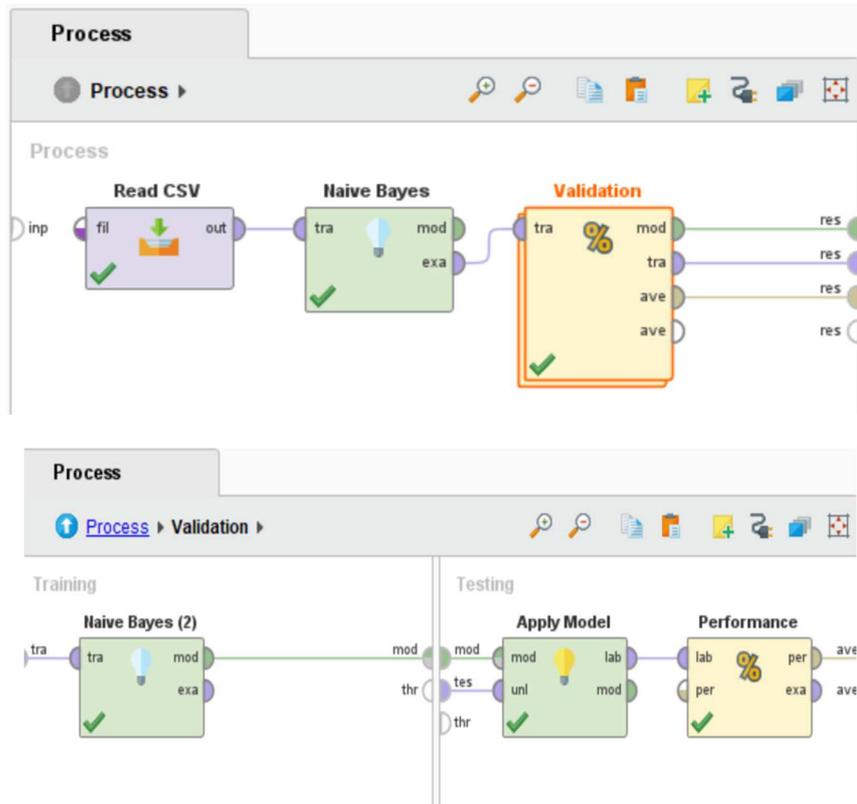
#### 14. Evaluasi dan Validasi Penelitian

Tahapan ini akan membahas hasil evaluasi dari eksperimen yang telah digunakan, menggunakan *K-Fold Cross Validation*.

## 3. Results and Analysis

### 4.5 Hasil Eksperimen dan Pengujian Metode

Untuk mendapatkan model yang sesuai dengan yang diharapkan, pada tahapan eksperimen dilakukan dengan melakukan percobaan dengan Rapid Miner seperti contoh dibawah ini :



Gambar 1. Proses Rapid Miner

Setelah itu dilakukan pengujian menggunakan tools Rapid Miner, kemudian :

4. Menambahkan dua operator read excel untuk memasukan data *training* dan data *testing*
5. Menambahkan operator *naive bayes* yang digunakan sebagai metode dalam penelitian.
6. Menambahkan operator *split validation* digunakan untuk proses pengolahan data. Yang didalamnya terdapat operator *Naive Bayes*, *Apply Model* dan *Performance* untuk menentukan hasil prediksinya.

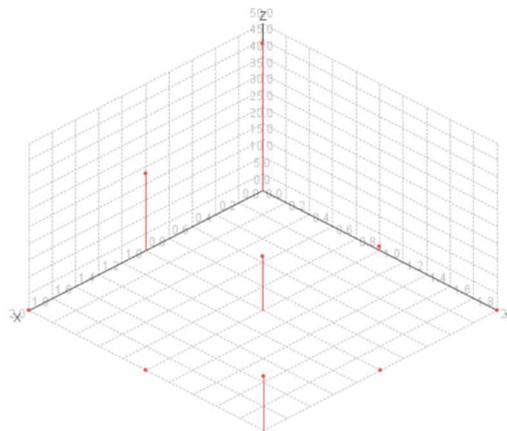
**4.1.3 Proses Validasi**

Pada tahapan validasi ini menggunakan data responden sebanyak 500 pasien, 25 Variabel *input* dan 3 *output* dengan *split relative*, *split ratio* dari 0.1-0.9, *sampling type automatic* didapat hasil seperti tabel dibawah ini :

Tabel 1. hasil proses validasi

No	Split	Sampling Type	Splite Ratio	Accuracy
1	<i>relative</i>	<i>automatic</i>	0.1	70.44%
2	<i>relative</i>	<i>automatic</i>	0.2	73.50%
3	<i>relative</i>	<i>automatic</i>	0.3	73.14%
4	<i>relative</i>	<i>automatic</i>	0.4	74.25%
5	<i>relative</i>	<i>automatic</i>	0.5	63.60%
6	<i>relative</i>	<i>automatic</i>	0.6	73.13%
7	<i>relative</i>	<i>automatic</i>	0.7	72.67%
<b>8</b>	<b><i>relative</i></b>	<b><i>automatic</i></b>	<b>0.8</b>	<b>76.00%</b>
9	<i>relative</i>	<i>automatic</i>	0.9	74.00%

Dari bebrapa percobaan yang dilakukan mendapatkan akurasi yang lebih baik adalah 76.00% dengan split ratio 0.8 dengan bentuk *Confusion Matrik* seperti dibawah ini :



Gambar 2. Confusion Matrik

#### 4.1.4 Metode Naïve Bayes

Proses klasifikasi dengan Rapid Miner menggunakan metode *Naïve Bayes* yang digunakan untuk memprediksi/klasifikasi tindakan pasien sehingga didapat akurasi hasil prediksi 76.00% dari hasil data *testing*. Pengujian berdasarkan *Confusion Matrix* menghasilkan nilai *accuracy*, *precision* dan *recall* yang tinggi dengan nilai *accuracy* sebesar 76.00%. dengan *Class precision* untuk Pred. Isoman 65.67%, Pred. Isolasi rumah sakit 94.12%, Pred. ICU 100% sedangkan *Class Recall* untuk *True* Isoman 97.78%, *True*. Isolasi rumah sakit 41.03%, *True* ICU 100%. Dengan tabel dibawah ini:

Tabel 2. *Class Precision* dan *Class Recall*

	true Isoman	true Isolasi rumah sakit	true ICU	class precision
pred. Isoman	44	23	0	65.67%
pred. Isolasi rumah sakit	1	16	0	94.12%
pred. ICU	0	0	16	100.00%
class recall	97.78%	41.03%	100.00%	

#### 4.6 Evaluasi dan Validasi

Setelah bebrapa kombinasi nilai parameter yang telah didapatkan, maka setelah divalidasi pada model yang telah digunakan maka didapatkan hasil nilai *accuracy* atau parameter terbaik

Tabel 3. Nilai *accuracy*/parameter Naïve Bayes

Split Ratio	= 0.8
<i>accuracy</i>	= 76.00%

#### 4.7 Proses Pengujian dan Testing

Pengujian dan testing dilakukan untuk melihat sejauh mana model dari yang dibuat dapat berjalan pada dataset yang dijadikan untuk testing, hasil analisa antara data *testing* dengan data *training* pada Rapid Miner dapat dilihat pada Tabel 4.2. Untuk menghitung akurasi sebagai berikut :

Jumlah data yang di uji	: 500
Jumlah data yang diprediksi benar "Isoman"	: 44
Jumlah data yang diprediksi benar "Isolasi rumah sakit"	: 16
Jumlah data yang diprediksi benar "ICU"	: 16
Jumlah data yang diprediksi salah "Isoman"	: 0
Jumlah data yang diprediksi salah "Isolasi rumah sakit"	: 23
Jumlah data yang diprediksi salah "ICU"	: 0

$$\begin{aligned} \text{Akurasi} &= \frac{44+16+16}{44+16+16+1+0+23} \times 100\% \\ &= \frac{76}{100} \times 100\% \\ &= 76,00\% \end{aligned}$$

$$\begin{aligned} \text{Eror} &= \frac{23+0+1}{44+16+16+1+0+23} \times 100\% \\ &= \frac{24}{100} \times 100\% \\ &= 24,00\% \end{aligned}$$

Dari perhitungan tersebut disimpulkan bahwa prediksi dengan menggunakan metode *Naive Bayes* untuk menentukan tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 menghasilkan tingkat akurasi sebesar 76.00% dan tingkat eror 24.00%.

#### 4. Conclusion

Dari penelitian yang dilakukan mengenai penerapan metode *Naive Bayes* untuk prediksi tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 dapat disimpulkan bahwa berdasarkan hasil analisis yang dilakukan pada penelitian ini, didapat hasil prediksi tindakan medis pada pasien ODP, PDP, OTG dan Positif Covid-19 dengan membandingkan data *training* dengan dengan data *testing* menggunakan aplikasi Rapid Miner didapat tingkat akurasi sebesar 76.00%.

#### References

- [1] T. Dey and A. Sinha, "Ethnicity and COVID-19 - A commentary on 'World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)' (Int J Surg 2020;76:71-6)," *Int. J. Surg.*, vol. 83, no. August, pp. 75–76, 2020, doi: 10.1016/j.ijsu.2020.08.046.
- [2] W. Wiguna and D. Riana, "Diagnosis of Coronavirus Disease 2019 (Covid-19) Surveillance Using C4.5 Algorithm," *J. Pilar Nusa Mandiri*, vol. 16, no. 1, pp. 71–80, 2020, doi: 10.33480/pilar.v16i1.1293.
- [3] B. Etikasari, T. D. Puspitasari, A. A. Kurniasari, and L. Perdanasari, "Sistem informasi deteksi dini Covid-19," *J. Tek. Elektro dan Komput.*, vol. 9, no. 2, pp. 101–108, 2020.
- [4] L. Wynants *et al.*, "Prediction models for diagnosis and prognosis of covid-19: Systematic

- review and critical appraisal,” *BMJ*, vol. 369, 2020, doi: 10.1136/bmj.m1328.
- [5] M. Pradana, S. Syahputra, A. Wardhana, B. R. Kartawinata, and C. Wijayangka, “The Effects of Incriminating COVID-19 News on the Returning Indonesians’ Anxiety,” *J. Loss Trauma*, vol. 25, no. 8, pp. 1–6, 2020, doi: 10.1080/15325024.2020.1771825.
- [6] World Health Organization (WHO) Indonesia, “Coronavirus Disease 2019 (COVID-19) WHO Indonesia Situation Report,” *Indones. Situat. Reports*, vol. 2019, no. April, pp. 1–9, 2020, [Online]. Available: <https://www.who.int/indonesia/news/novel-coronavirus/situation-reports>.
- [7] National Disaster Management Agency, “Head of BNPB Decision Letter No. 9A/2020.” p. 3, 2020.
- [8] N. Anung Ahadi Pradana, Casman, “Pengaruh kebijakan,” *Pengaruh Kebijak. Soc. Distancing Pada Wabah Covid-19 Terhadap Kelompok Rentan Di Indones.*, vol. 09, no. 02, pp. 61–67, 2020.
- [9] F. Sodik, B. Dwi, and I. Kharisudin, “Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python,” vol. 3, pp. 689–694, 2020.
- [10] M. Informatika and K. B. Aceh, “ANALISIS TREND TOPIK PENELITIAN PADA WEB OF SCIENCE DAN SINTA UNTUK PENENTUAN TEMA,” pp. 13–25, 2018.
- [11] T. Praningki and I. Budi, “Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART, Naive Bayes, dan k-NN,” *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 83, 2018, doi: 10.24076/citec.2017v4i2.100.



Data Sheet Penelitian

Respon	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	Y
1	0	1	0	0	1	0	0	1	0	0	1	0	0	1	1	0	0	1	1	1	0	1	1	1	0	Isoman
2	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	Isoman
3	0	1	0	1	1	0	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	Isoman
4	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	Isolasi rumah sakit
5	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	Isoman
6	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	Isoman
7	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	ICU
8	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Isoman
9	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	ICU
10	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	ICU
11	0	1	0	0	1	0	0	0	0	0	1	0	0	1	1	0	0	1	1	1	0	0	1	1	0	Isoman
12	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	Isoman
13	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0	Isolasi rumah sakit
14	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	Isoman
15	1	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Isoman
16	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	Isoman
17	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	ICU
18	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	Isoman
19	1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	Isoman
20	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Isoman
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	Isolasi rumah sakit
22	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	1	0	1	1	1	0	Isolasi rumah sakit

