

# Prediksi Penyebaran Virus COVID-19 Dari Hasil PCR Menggunakan Metode Naïve Bayes

Hetty Rohayani<sup>1\*</sup>, Sitti Nur Alam<sup>2</sup>, Muhammad Fauzi<sup>1</sup>, Rico<sup>3</sup>

<sup>1</sup>Fakultas Sains dan Teknologi, Informatika, Universitas Muhammadiyah Jambi, Jambi, Indonesia

<sup>2</sup>Program Studi Sistem Informasi, Universitas Yapis Papua, Jayapura Indonesia

<sup>3</sup>Fakultas Teknik dan Ilmu Komputer, Universitas Adiwangsa Jambi, Jambi, Indonesia

Email: <sup>1,\*</sup>hettyrohayani@gmail.com, <sup>2</sup>azkadzar@gmail.com, <sup>3</sup>fauzim876@gmail.com, <sup>4</sup>Reecho86@gmail.com

Email Penulis Korespondensi: hettyrohayani@gmail.com

Submitted: 23/11/2022; Accepted: 29/11/2022; Published: 30/11/2022

**Abstrak**—Pandemi Covid-19 di tahun 2020 yang terjadi di Indonesia menjadi masalah kesehatan yang kompleks dan membutuhkan penanganan cepat serta kolaborasi solusi dari berbagai disiplin ilmu. Pasien Covid-19 yang mendapatkan perawatan di rumah sakit memiliki kondisi dan tingkat keparahan yang berbeda-beda. Hal ini berpengaruh pada tindakan penanganan yang akan dilakukan oleh petugas medis. Banyaknya pasien serta kurangnya tenaga medis mengakibatkan perlunya dukungan teknologi untuk membantu mengklasifikasikan status pasien berdasarkan kondisinya agar penanganan dikonsentrasikan pada pasien yang sangat gawat dan membutuhkan penanganan cepat. Banyaknya karyawan pada Kejaksaan Negeri Jambi yang terpapar virus Covid-19 pada khususnya, dan perlu dilakukan prediksi hasil PCR. Penelitian ini menerapkan teknik prediksi dari bagian disiplin ilmu data mining untuk mengklasifikasikan status pegawai di Kejaksaan Negeri Jambi yang terkonfirmasi Covid-19. Pengklasifikasi menggunakan metode Naive Bayes diterapkan untuk membangun model berdasarkan dataset pasien yang terinfeksi Covid-19. Dataset hasil PCR Covid-19 diperoleh dari Kejaksaan Negeri Jambi dan diaplikasikan menggunakan RapidMiner. Model yang dibangun dapat memprediksi status terkonfirmasi Covid-19 berdasarkan usia, jenis kelamin, kondisi kesehatan, dan vaksinasi. Hasil penelitian ini menunjukkan bahwa klasifikasi metode Naive Bayes memiliki tingkat akurasi yang tinggi dalam mengklasifikasikan status terkonfirmasi Covid-19 yaitu 97%.

**Kata Kunci:** Prediksi; Data Mining; COVID-19; Algoritma Naïve Bayes; Rapidminer

**Abstract**—The Covid-19 pandemic in 2020 that occurred in Indonesia is a complex health problem and requires fast handling and collaborative solutions from various disciplines. Covid-19 patients who receive treatment in hospitals have different conditions and severity. This affects the treatment actions that will be carried out by medical officers. The large number of patients and the lack of medical personnel have resulted in the need for technological support to help classify patient status based on their condition so that treatment is concentrated on patients who are very critical and require rapid treatment. The number of employees at the Jambi District Attorney's Office were exposed to the Covid-19 virus in particular, and it is necessary to predict PCR results. This study applies a prediction technique from the data mining discipline section to classify the status of employees at the Jambi District Attorney's Office who are confirmed to be Covid-19. Classifiers using the Naive Bayes method were applied to build a model based on a dataset of patients infected with Covid-19. The Covid-19 PCR result dataset was obtained from the Jambi District Attorney and was applied using RapidMiner. The model built can predict the confirmed status of Covid-19 based on age, gender, health condition, and vaccinations. The results of this study indicate that the classification of the Naive Bayes method has a high level of accuracy in classifying the confirmed status of Covid-19, namely 97%.

**Keywords:** Prediction; Data Mining; COVID-19; Naive Bayes Algorithm; Rapidminer

## 1. PENDAHULUAN

Covid-19 adalah virus baru yang muncul di akhir tahun 2019 di kota Wuhan, Tiongkok. Virus ini terus berkembang hingga menyebar ke seluruh dunia pada tahun 2020. *World Health Organization (WHO)* menginformasikan bahwa Covid-19 sebagai pandemi pada 11 Maret 2020. Virus ini menyebabkan infeksi pada paru-paru yang berdampak fatal berupa kematian [1]. Indonesia pertama kali mengonfirmasi kasus virus Covid-19 pada Senin, 2 Maret 2020 lalu [2]. Sehingga berbagai cara untuk melakukan pencegahan penyebaranpun dilakukan mulai dari Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM), menjaga jarak fisik, memakai masker, serta menerapkan pola hidup yang sehat. Bulan Juli 2022 kematian yang diakibatkan oleh virus Covid-19 di Indonesia [3] mengalami peningkatan, dengan total keseluruhan kasus Covid-19 mencapai 6,046 juta kasus, dengan 5,88 juta orang sembuh, hal ini terjadi karena masih minimnya masyarakat yang mematuhi berbagai himbauan yang dilakukan pemerintah untuk menghentikan penyebaran virus Covid-19 ini [1].

Perkembangan kondisi pandemi COVID-19 [4] yang mengkhawatirkan dengan ditandainya penambahan jumlah kasus terkonfirmasi positif dari hari ke hari khususnya di Indonesia dan umumnya di seluruh dunia [5] terus bertambah. Untuk menentukan apakah seseorang terjangkit COVID-19 [5], diperlukan pemeriksaan PCR. Hasil penelitian terbaru menunjukkan bahwa beberapa kasus dapat menunjukkan hasil positif yang terus menerus meskipun tidak ada gejala [6]. Oleh karena itu, meskipun sudah dinyatakan sembuh dari COVID-19, tetap harus menerapkan prosedur kesehatan. Pertambahan jumlah kasus terkonfirmasi sebagian besar disebabkan oleh penularan virus antar keluarga dan orang terdekat dalam satu komunitas, dimana seseorang dengan gejala ringan COVID-19 masih belum menyadari jika dirinya sudah mulai terinfeksi virus.

Banyaknya pasien yang di rumah sakit dan kapasitas tenaga medis menjadi permasalahan utama yang dihadapi di berbagai daerah. Pasien dengan tingkat kegawatan tinggi memerlukan prioritas penanganan dibanding

pasien dengan gejala sedang atau tanpa gejala. Tenaga medis memerlukan bantuan untuk mengklasifikasi status pasien berdasarkan data pasien secara otomatis untuk mengurangi kelelahan tenaga medis yang harus terus bertugas dan meminimalisir risiko penanganan yang terlambat terhadap pasien. Oleh karena itu dibutuhkan solusi teknologi berbasis data secara otomatis yang dapat membantu mengklasifikasikan status kegawatan berdasarkan data pasien[7].

Penelitian ini bertujuan memberikan solusi untuk mengkalasifikasikan sttus pasien Covid-19 pada Kejaksaan Negeri Jambi secara otomatis berdasarkan gejala yang dialami pasien. Klasifikasi dilakukan menggunakan metode Naïve-Bayes yang dikenal memiliki akurasi tinggi dengan pembelajaran terseperviisi berbasis distribusi dataset. Kontribusi utama pada penelotian ini adalah menghasilkan model prediktif status pasien Covid-19 pada Kejaksaan Negeri Jambi sehingga dapat disiapkan langkah-langkah yang cepat dan tepat untuk penanganan pasien oleh tenaga medis.

Pada penelitian yang dilakukan oleh Iqbal dkk, memprediksi kebutuhan ICU pada pasien terjangkit virus Covid-19, sehingga disetiap population mengalami peningkatan nilai akurasinya[7]. Saleh [8] melakukan penelitian dengan melakukan Penerapan metode naïve bayes, maka diperoleh hasil persentase 78,3333% untuk keakuratan prediksi. Penerapan metode Naive Bayes diharapkan dapat memprediksi pembelian cat dari supplier, data pembelian cat berhasil diklasifikasikan dengan benar [9].

Data mining adalah teknik yang digunakan untuk membangun model pembelajaran mesin [10][11][12]. Pembelajaran mesin (*machine learning*) [13] adalah teknik kecerdasan buatan modern yang belajar membangun model dengan menggunakan data empiris. Data Mining [14] digunakan untuk menemukan pola dalam kumpulan besar data mentah. Data Mining [15] menerapkan teknik Machine Learning untuk menarik pengetahuan pada data. Dalam penelitian ini penulis menerapkan teknik data mining untuk mengklasifikasikan dataset Covid-19 menggunakan Algoritma *Naive Bayes Classifier (NBC)* karena NBC telah berhasil diterapkan dalam banyak tugas klasifikasi berbasis probabilitas bersyarat pada populasi data [7].

Dalam penelitian [1] Model yang dibangun dapat memprediksi status penanganan pasien berdasarkan usia dan jenis kelamin yang memiliki kemungkinan tertinggi untuk ditangani dengan cara isoman atau harus dirujuk kerumah sakit. Data yang digunakan diaplikasikan menggunakan Rapidminer dengan validasi yang digunakan adalah spilt validation serta ratio data training 70% dan data uji 30%, dengan hasil penelitian menunjukkan bahwa klasifikasi menggunakan algoritma Naive Bayes Classifier memiliki tingkat akurasi yang tinggi dalam melakukan klasifikasi data status pasien yaitu 83,33 %.

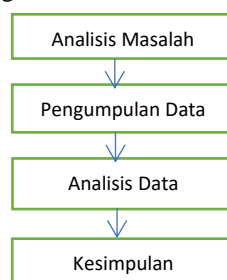
Penelitian ini bertujuan memberikan solusi untuk mengklasifikasikan status pasien Covid-19 secara otomatis berdasarkan kondisi pasien pada instansi Kejaksaan Negeri Jambi. Klasifikasi dilakukan menggunakan algoritma Naïve-Bayes [16] yang dikenal memiliki akurasi tinggi dengan pembelajaran tersuperviisi berbasis distribusi dataset. Pemilihan algoritma yang tepat diharapkan dapat menghasilkan model klasifikasi yang baik dan berguna untuk diimplementasikan secara riil pada penanganan dan skrining status kegawatan pasien Covid-19. Sehingga diharapkan dapat membantu memprediksi status pasien untuk mendapatkan penanganan yang tepat secara langsung oleh tim medis.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

#### 2.1.1 Analisis Masalah

Analisis masalah merupakan tahap awal dari penelitian ini. Analisis masalah dilakukan untuk mengidentifikasi permasalahan yang muncul dari hasil PCR yang dilakukan. Seperti hasil PCR yang terkonfirmasi positif. Menanggapi permasalahan tersebut, diperlukan suatu informasi untuk membangun sistem klasifikasi penentuan status pasien COVID – 19 di Kejaksaan Negeri Jambi:



**Gambar 1.** Tahapan Penelitian

#### 2.1.2 Pengumpulan Data

Pengumpulan data merupakan tahap inti dari suatu penelitian dimana data ini akan menentukan probabilitas hasil swap PCR. Dalam penelitian ini, dataset Penentuan hasil PCR di Kejaksaan Negeri Jambi memiliki jumlah total

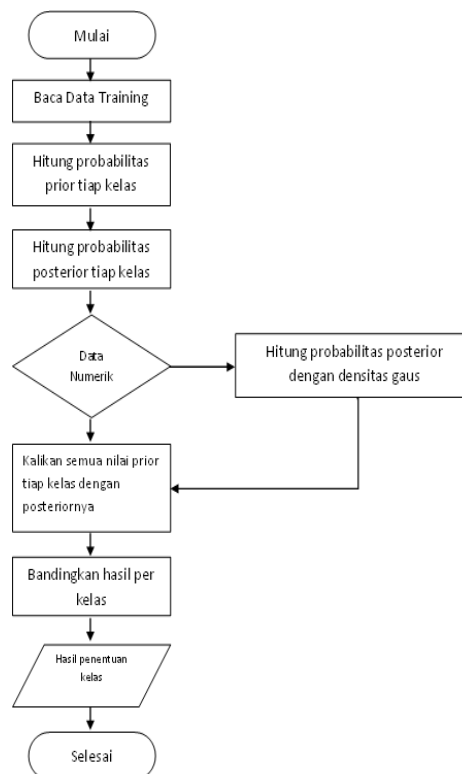
data sebanyak 50 data pegawai Kejaksaan Negeri Jambi yang diantaranya memiliki variable seperti; jenis kelamin, umur, kondisi kesehatan, dan vaksinasi.

### 2.1.3 Analisis Data

Data yang didapatkan kemudian dianalisa untuk menentukan variabel yang digunakan dalam pembuatan aplikasi. Sebagian variable data yang ada dalam bentuk angka, untuk penggunaan model algoritma Naïve bayes data angka harus ditransformasikan ke dalam klasifikasi atau kelompok berdasar interval [17]

## 2.2. Metode Penelitian

*Naïve Bayes Classifier (NBC)* [18] adalah klasifikasi dengan probabilitas sederhana yang didasarkan pada teorema bayes. Teorema bayes memiliki tambahan kata yaitu “Naïve” dengan makna setiap atribut atau variable bersifat independent (bebas). Berikut adalah alur kerja dari metode Naïve bayes :



**Gambar 2.** Alur metode Algoritma Naïve Bayes

Bentuk umum algoritma NBC dapat dilihat pada persamaan 1.

$$(H|X) = P(X|H)P(H) P(X) \quad (1)$$

Keterangan :

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

$P(H|X)$  = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

$P(H)$  = Probabilitas hipotesis H (prior prob.)

$P(X|H)$  = Probabilitas X berdasarkan kondisi tersebut

$P(X)$  = Probabilitas dari X

Metode pengembangan data mining yang digunakan untuk menganalisis data dalam penerapan data mining ini menggunakan proses tahapan yang terdiri dari Data, Data Cleaning, Data transformation, Data mining, evaluation.

Berikut merupakan hal –hal yang perlu di lakukan dalam penelitian berdasarkan tahapan knowledge discovery in databases :

#### 1. Data

Data merupakan sekumpulan data operasional yang diperlu sebelum dilakukan sebelum tahap penggalian informasi dimulai.

#### 2. Data Cleaning

Proses data cleaning merupakan proses Pembersihan data yang bertujuan untuk menghilangkan data yang tidak memiliki nilai (null), data yang salah input, data yang tidak relevan, duplikat data dan data yang tidak konsisten.

## 3. Data transformation

Data transformation dilakukan dengan memberikan inisialisasi terhadap data yang memiliki nilai nominal menjadi bernilai numerik.

## 4. Data Mining

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

## 5. Evaluation

Pada tahap evaluasi, akan diketahui apakah hasil daripada tahap data mining dapat menjawab tujuan yang telah ditetapkan. Untuk itu akan dilakukan profilisasi pada setiap cluster yang telah terbentuk, untuk diketahui karakteristik pada kelompok tersebut. Disamping itu untuk diketahui kesesuaian dengan jalur perminatan akan dilakukan analisis lebih lanjut untuk dihubungkan dengan atribut [19].

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pengumpulan Dataset

Dataset hasil PCR Covid-19 di Indonesia diperoleh dari Kejaksaan Negeri Jambi. Dataset inilah yang nantinya akan dilatih untuk menghasilkan model prediksi terhadap status pasien. Dataset disiapkan dan dibersihkan di mana hanya atribut relevan yang diekstraksi dan digunakan dalam proses klasifikasi. Dataset pasien terdiri dari 50 baris data dengan empat atribut. Dataset pasien digambarkan pada tabel 1.

**Tabel 1.** Dataset hasil swab PCR sebagai data training

Nama	Gender	Umur	Kondisi	Vaksin	Hasil Swab
Hakim Albana	Laki-Laki	young	sehat	sudah	negatif
Pison Darma	Laki-Laki	young	sehat	belum	negatif
Rusyidi Sastrawan	Laki-Laki	older	sehat	belum	negatif
Firman	Laki-Laki	older	sakit	belum	positif
Boby Setiawan	Laki-Laki	young	sakit	belum	positif
Dewangga Adi	Laki-Laki	young	sakit	sudah	positif
Susi Triati	perempuan	young	sehat	sudah	positif
Muhammad Gempa	Laki-Laki	young	sehat	sudah	positif
Lorenza	perempuan	young	sehat	sudah	positif
Teti Kurnia	perempuan	young	sehat	sudah	negatif
Ramario	Laki-Laki	older	sehat	sudah	negatif
Juniarty	perempuan	young	sehat	belum	negatif
Faidol Mubarak	Laki-Laki	young	sehat	belum	negatif
Muntasriyah	perempuan	older	sehat	belum	negatif
Dewi Ariningrum	perempuan	young	sehat	sudah	positif
Winastri	perempuan	older	sehat	sudah	positif
Merry Agustin	perempuan	older	sakit	belum	positif
Nur Zalipah	perempuan	older	sakit	sudah	negatif
Susi Indriani	perempuan	older	sehat	belum	negatif
Mariana	perempuan	older	sehat	sudah	negatif
Royandri	Laki-Laki	young	sehat	sudah	negatif
Aprizal	Laki-Laki	young	sehat	sudah	negatif
M. Robi	Laki-Laki	older	sehat	belum	negatif
Sri Maryani	perempuan	young	sehat	sudah	negatif
Rizka Amalia	perempuan	young	sehat	sudah	negatif
M. Amir	Laki-Laki	older	sehat	sudah	negatif
Putra Ramadhan	Laki-Laki	older	sehat	sudah	negatif
Dede Supriadi	Laki-Laki	young	sakit	sudah	negatif
Fazar Rudi	Laki-Laki	older	sakit	sudah	negatif
Fikri Iqbal	Laki-Laki	older	sakit	belum	positif
Andi Wira	Laki-Laki	older	sakit	belum	negatif

Pelatihan dan pengujian pertama

#### 1. Hasil perhitungan probabilitas tiap kelas

**Tabel 2.** Probabilitas Prior kelas positif dan negatif

Probabilitas Kelas	
Hasil swab	Nilai

Positif	0.34
Negatif	0.66
	100 %

$$P(Y = \text{positif}) = \frac{17}{50} = 0.34$$

$$P(Y = \text{negatif}) = \frac{33}{50} = 0.66$$

## 2. Hasil Perhitungan Probabilitas Posterior tiap Atribut per Kelas

**Tabel 3.** Probabilitas Posterior Atribut Gender

P (Gender)	positif	negatif
Laki-laki	0.588235294	0.515151515
Perempuan	0.411764706	0.484848485

$$P(\text{Gender} = \text{laki-laki} | Y = \text{positif}) = \frac{10}{17} = 0.588235$$

$$P(\text{Gender} = \text{Perempuan} | Y = \text{positif}) = \frac{7}{17} = 0.411764$$

$$P(\text{Gender} = \text{laki-laki} | Y = \text{negatif}) = \frac{17}{33} = 0.515151$$

$$P(\text{Gender} = \text{Perempuan} | Y = \text{negatif}) = \frac{16}{33} = 0.484848$$

**Tabel 4.** Probabilitas Posterior Atribut Umur

P (Umur)	positif	negatif
Young	0.470588235	0.545454545
older	0.529411765	0.454545455

$$P(\text{Umur} = \text{young} | Y = \text{positif}) = \frac{8}{17} = 0.470588$$

$$P(\text{Umur} = \text{older} | Y = \text{positif}) = \frac{9}{17} = 0.529411$$

$$P(\text{Umur} = \text{young} | Y = \text{negatif}) = \frac{18}{33} = 0.545454$$

$$P(\text{Umur} = \text{older} | Y = \text{negatif}) = \frac{15}{33} = 0.454545$$

**Tabel 5.** Probabilitas Posterior Atribut Kondisi

P (Kondisi)	positif	negatif
Sehat	0.588235294	0.696969697
sakit	0.411764706	0.303030303

$$P(\text{Kondisi} = \text{sehat} | Y = \text{positif}) = \frac{10}{17} = 0.588235$$

$$P(\text{Kondisi} = \text{sakit} | Y = \text{positif}) = \frac{7}{17} = 0.411764$$

$$P(\text{Kondisi} = \text{sehat} | Y = \text{negatif}) = \frac{23}{33} = 0.696969$$

$$P(\text{Kondisi} = \text{sakit} | Y = \text{negatif}) = \frac{10}{33} = 0.303030$$

**Tabel 6.** Probabilitas Posterior Atribut Vaksinasi

P (Vaksin)	positif	negatif
sudah	0.529411765	0.545454545
belum	0.470588235	0.454545455

$$P(\text{Vaksin} = \text{sudah} | Y = \text{positif}) = \frac{9}{17} = 0.529411$$

$$P(\text{Vaksin} = \text{belum} | Y = \text{positif}) = \frac{8}{17} = 0.470588$$

$$P(\text{Vaksin} = \text{sudah} | Y = \text{negatif}) = \frac{18}{33} = 0.545454$$

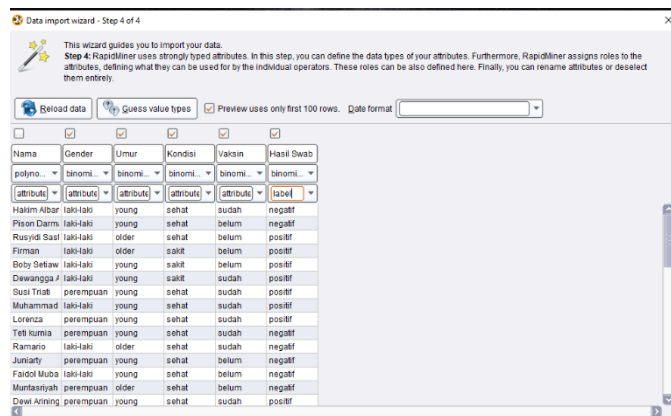
$$P(\text{Vaksin} = \text{belum} | Y = \text{negatif}) = \frac{15}{33} = 0.454545$$

## 3.2. Implementasi Data Mining Pada Rapidminer

### 1. Pre-processing Data

Tahap preprocessing data meliputi tahap sebagai berikut :

Import data pada aplikasi rapid miner : mengimpor data excel yang telah disiapkan sebagai data training dan menghilangkan kelas yang tidak relevan. Pada kelas tersebut disesuaikan atribut yang sesuai seperti kelas hasil swap dengan atribut label dengan tipe binominal (karena memiliki nilai hanya 2).

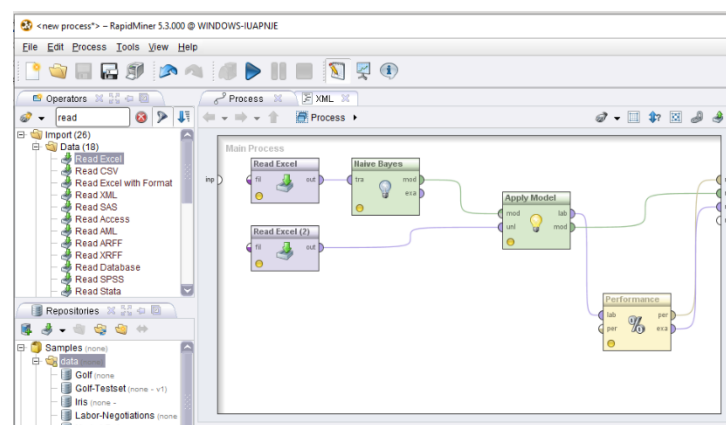


Gambar 3. Tahap pre-processing data training

Mengimpor data excel yang telah disiapkan sebagai data testing dan menghilangkan kelas yang tidak relevan.

## 2. Pengolahan Data

Tahap pengolahan data meliputi tahap sebagai berikut :



Gambar 4. Tahap pengolahan data pada rapid miner

Dalam pengolahan data dilakukan proses data menggunakan algoritma Naïve Bayes yang meliputi proses training dan testing (*apply model dan performance*). Setelah itu dilakukan proses pemodelan dengan menggunakan Algoritma Naïve Bayes untuk mendapatkan nilai *accuracy*, *confidence*, *class recall* dan *prediction* berdasarkan *performance vector*. Nilai akurasi yang didapatkan dengan algoritma naïve bayes ini adalah 97%. Dengan nilai akurasi recall untuk hasil positif adalah 3 %, sedangkan nilai akurasi recall untuk hasil negatif adalah 97%. Nilai akurasi precision untuk pred. positif adalah 3 %, sedangkan nilai akurasi precision untuk pred. negatif adalah 93.4%. Dapat disimpulkan bahwa algoritma naïve bayes ini memiliki nilai akurasi yang tinggi sehingga hasil dari prediksinya akurat untuk dijadikan metode dalam penelitian Data Mining. Berikut prediksi yang dapat ditemukan pada algoritma Naïve bayes

Row No.	Hasil S...	confidence(negatif)	confidence(positif)	prediction(Hasil Swab)	Gender	Umur	Kondisi	Vaksin
1	negatif	0.641	0.359	negatif	laki-laki	older	sehat	sudah
2	negatif	0.693	0.307	negatif	laki-laki	young	sehat	belum
3	negatif	0.626	0.374	negatif	laki-laki	older	sehat	belum
4	negatif	0.584	0.416	negatif	laki-laki	young	sakit	belum
5	negatif	0.584	0.416	negatif	laki-laki	young	sakit	belum
6	negatif	0.526	0.474	negatif	laki-laki	older	sakit	sudah
7	negatif	0.764	0.236	negatif	perempuan	young	sehat	sudah

Gambar 5. Nilai prediksi dari data testing

Nilai confidence yang terdapat pada tiap-tiap atribut dapat ditemukannya probabilitas dari masing-masing data testing sehingga bisa diprediksi hasil PCR pegawai Kejaksaan Negeri yang terkonfirmasi COVID-19.

#### 4. KESIMPULAN

Berdasarkan analisis perhitungan data mining menggunakan algoritma naïve bayes, dapat ditarik kesimpulan bahwa pengelompokan kelas hasil swab “positif” lebih sedikit dari pengelompokan kelas hasil swab “negatif”. Hasil observasi terhadap dataset hasil swab di Kejaksaan Negeri Jambi yang telah dilakukan melalui proses perhitungan menggunakan metode klasifikasi naïve Bayes dengan atribut yang telah dijelaskan di pembahasan sebelumnya, didapatkan hasil bahwa nilai akurasi sebesar 97 % sehingga dapat memprediksi cukup akurat.

#### REFERENCES

- [1] D. Pramana and M. Mustakim, “Prediksi Status Penanganan Pasien Covid-19 dengan Algoritma Naïve Bayes Classifier di Provinsi Riau,” *J. Sist. Komput. dan Inform.*, vol. 3, no. 2, pp. 202–208, 2021.
- [2] D. Susanna, “When will the COVID-19 pandemic in indonesia end?,” *Natl. Public Heal. J. (Jurnal Kesehat. Masy. Nas.)*, vol. 15, no. 4, pp. 160–162, 2020.
- [3] Y. Indrianingrum and A. Wiranta, “Correlation of Factors Causing the Death of COVID-19 Patients and Enforcement of Regulations in Handling COVID-19 in the City of Bogor,” *J. Bina Praja*, vol. 13, no. 3, pp. 471–484, 2021.
- [4] Y. C. Wu, C. S. Chen, and Y. J. Chan, “The outbreak of COVID-19: An overview,” *J. Chinese Med. Assoc.*, vol. 83, no. 3, pp. 217–220, 2020.
- [5] T. Singhal, “A Review of Coronavirus Disease-2019 (COVID-19),” *Indian J. Pediatr.*, vol. 87, no. 4, pp. 281–286, 2020.
- [6] F. A. Nugroho, A. F. Solikin, M. D. Anggraini, and K. Kusriani, “Sistem Pakar Diagnosa Virus Corona Dengan Metode Naïve Bayes,” *J. Teknol. Inf. dan Komun.*, vol. 9, no. 1, p. 81, 2021.
- [7] I. Harifal and L. D. Lestari, “Optimasi Naïve Bayes dengan PSO untuk Prediksi Kebutuhan ICU Pasien Covid-19 Naïve Bayes Optimization with PSO for Predicting ICU Needs for,” *J. Sist. Informasi, Sist.*, vol. 11, no. September, pp. 724–734, 2022.
- [8] A. Saleh, “Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga,” *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 73–81, 2019.
- [9] F. Harahap, N. E. Saragih, E. T. Siregar, and H. Sariangisah, “Penerapan Data Mining Dengan Algoritma Naïve Bayes Classifier Dalam Memprediksi Pembelian Cat,” *J. Ilm. Inform.*, vol. 09, no. 01, pp. 19–23, 2021.
- [10] H. Rohayani and M. C. Umam, “Prediksi Penentuan Program Studi Berdasarkan Nilai Siswa dengan Algoritma Backpropagation,” *J. Inf. Syst. Res.*, vol. 3, no. 4, pp. 651–657, 2022.
- [11] Á. H. de Mendonça, A. E. Iribarne, and A. Yurkina, *Data mining for the masses*, vol. 8, no. 16, 2017.
- [12] S. Nagabhushana, *Data Warehousing, OLAP and Data Mining*. New Age International Publisher, 2013.
- [13] H. Rohayani, H. L. H. S. Warnars, T. Mauritsius, and E. Abdurrachman, “Wind speed forecasting in big data and machine learning: from presents, opportunities and future trends,” *Commonications Math. Biol. Neurosci. J.*, vol. 3, no. 1, pp. 1–16, 2021.
- [14] N. H. Son, “Introduction to Rough sets and Data mining.”
- [15] M. G. Feshki and O. S. Shijani, “Improving the heart disease diagnosis by evolutionary algorithm of PSO and Feed Forward Neural Network,” *2016 Artif. Intell. Robot. IRANOPEN 2016*, pp. 48–53, 2016.
- [16] Y. Huang and L. Li, “Naïve Bayes classification algorithm based on small sample set,” *CCIS2011 - Proc. 2011 IEEE Int. Conf. Cloud Comput. Intell. Syst.*, no. 2, pp. 34–39, 2011.
- [17] S. Rahayu and A. S. RMS, “Penerapan Metode Naïve Bayes Dalam Pemilihan Kualitas Jenis Rumput Taman CV. Rumput Kita Landscape,” *J. Teknol. Inf. dan Komun.*, vol. 9, no. 2, pp. 162–171, 2018.
- [18] K. Vembandasamy, R. Sasipriya, and E. Deepa, “Heart Diseases Detection Using Naïve Bayes Algorithm,” *Int. J. Innov. Sci. Eng. Technol.*, vol. 2, no. 9, pp. 441–444, 2015.
- [19] Heliyanti Susana, “Penerapan Model Klasifikasi Metode Naïve Bayes Terhadap Penggunaan Akses Internet,” *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022.