

# Optimasi Naïve Bayes dengan PSO untuk Prediksi Kebutuhan ICU Pasien Covid-19

## *Naive Bayes Optimization with PSO for Predicting ICU Needs for Covid-19 Patients*

<sup>1</sup>Iqbal Harifal, <sup>2</sup>Lusiana Dwi Lestari\*,

<sup>3</sup>Taslim, <sup>4</sup>Yogi Yunefri, <sup>5</sup>Susi Handayani, <sup>6</sup>Moh. Fatkuroji, <sup>7</sup>Eka Sabna, <sup>8</sup>Kursiah Warti Ningsih  
<sup>1,2,3,4,5</sup>Teknik Informatika, Fakultas Ilmu Komputer,

Universitas Lancang Kuning, Jl. Yos Sudarso KM. 8, Umban Sari, Kec. Rumbai, Pekanbaru.

<sup>6</sup>Teknik Informatika, Fakultas Ilmu Komputer, Universitas Hang Tuah Pekanbaru,

Jalan Mustafa Sari No.5 Tangkerang Selatan Kec Bukit Raya, Pekanbaru

<sup>7,8</sup>Kesehatan Masyarakat, STIKes Payung Negeri Pekanbaru,

Jalan Tamtama No. 6 Labuh Baru Timur, Kecamatan Payung Sekaki, Kota Pekanbaru

\*e-mail: [lusianadwilestari1@gmail.com](mailto:lusianadwilestari1@gmail.com)

(received: 20 Juni 2022, revised: 5 Juli 2022, accepted: 21 Juli 2022)

### Abstrak

Covid-19 merupakan pandemi global yang membutuhkan respons dunia yang terkoordinasi di semua sistem kesehatan dan perawatan kesehatan nasional. Mengidentifikasi pasien yang berisiko tinggi terjangkit virus Covid-19 adalah hal penting untuk meningkatkan kewaspadaan sebelum pasien terinfeksi lebih lanjut oleh virus Covid-19 yang dapat menyebabkan penyakit pernapasan parah sehingga membutuhkan perawatan khusus di *intensive care units* (ICU). Penelitian ini bertujuan untuk memprediksi kebutuhan ICU pada pasien terjangkit virus Covid-19. Hasil nilai dari prediksi kebutuhan ICU dipergunakan sebagai bahan acuan rumah sakit untuk memenuhi kebutuhan ICU pada pasien terjangkit Covid-19 sehingga dapat meningkatkan persediaan ICU. Prediksi tersebut akan dilakukan dengan metode algoritma Naïve Bayes dengan dilakukan optimasi menggunakan algoritma PSO. Berdasarkan hasil penelitian, maka diperoleh population size 20 dengan nilai akurasi algoritma NBC sebesar 87,03%, population size 40 dengan nilai akurasi 87,28, population size 60 didapatkan hasil akurasi 87,13%, population size 80 dengan nilai akurasi 87,16% dan population size 100 didapatkan hasil 87,26% sehingga disetiap population mengalami peningkatan nilai akurasinya.

**Kata kunci:** Covid-19, Optimasi, Naïve Bayes, Prediksi, PSO

### Abstract

*Covid-19 is a global pandemic that requires a coordinated worldwide response across all national health and healthcare systems. Identifying patients who are at high risk of contracting the Covid-19 virus is important to increase awareness before patients are further infected by the Covid-19 virus which can cause severe respiratory illness that requires special treatment in intensive care units (ICU). This study aims to predict ICU needs in patients infected with the Covid-19 virus. The value results from the prediction of ICU needs are used as a reference for hospitals to meet ICU needs for patients infected with Covid-19 so that they can increase ICU supplies. The prediction will be carried out using the Naïve Bayes algorithm method with optimization using the PSO algorithm. Based on the results of the study, the population size 20 with an accuracy value of the NBC algorithm was 87.03%, population size 40 with an accuracy value of 87.28, population size 60 obtained an accuracy of 87.13%, population size 80 with an accuracy value of 87.16 % and population size 100, the results obtained are 87.26% so that each population has an increase in the accuracy value.*

**Keywords:** Covid-19, Optimization, Naïve Bayes, Prediction, PSO

## 1 Pendahuluan

Pada masa sekarang, Covid-19 merupakan pandemi global yang membutuhkan respons dunia yang terkoordinasi di semua sistem kesehatan dan perawatan kesehatan nasional. Corona virus disease 2019 atau sering disebut dengan Covid-19 ialah suatu penyakit infeksi yang menyerang saluran pernafasan akut [1]. Covid-19 merupakan penyakit yang menyebabkan pneumonia yang umumnya dapat dilihat dari pasien yang terjangkit Covid-19. Gejala umum yang diakibatkan oleh covid-19 diantaranya yaitu demam tinggi, batuk kering, nyeri otot, nafas sesak, hilangnya indra penciuman dan indra perasa, dan sakit kepala hingga kegagalan respirasi akut [2].

Mengidentifikasi pasien yang berisiko tinggi terjangkit virus Covid-19 adalah hal penting untuk meningkatkan kewaspadaan sebelum pasien terinfeksi lebih lanjut, karena Covid-19 dapat menyebabkan penyakit pernapasan parah yang mungkin memerlukan perawatan di *intensive care units* (ICU) [3]. Mengingat tingginya kasus virus Covid-19 di dunia, prediksi persediaan ruangan ICU menjadi prioritas utama bagi rumah sakit, sehingga pasien yang sangat membutuhkan ruang ICU, tidak terlantar karena kekurangan ruangan ICU yang disediakan rumah sakit.

Riset yang dilakukan oleh SARS- RAS dari *Italian Society of Hypertension* menyatakan bahwa, tingkat penerimaan *intensive care units* (ICU) pada masa pandemik Covid-19 di tiap negara yang ada dunia berbeda, sehingga menjadi salah satu tantangan utama yang harus diselesaikan oleh sistem perawatan kesehatan di dunia [4]. Karena belum ditemukannya kriteria yang menjadi latar belakang untuk menentukan ketersediaan ruang ICU, maka perlu dilakukan riset untuk mencari kriteria-kriteria yang menjadi pendukung pada pasien Covid-19 yang sangat membutuhkan perawatan intensif di ruang ICU. Pada penelitian ini, akan dilakukan prediksi terhadap pasien yang membutuhkan perawatan khusus di ruang ICU. Salah satu cara untuk menyelesaikan masalah tersebut, yaitu dengan menggunakan teknik data mining.

Data mining merupakan metode pengolahan berskala besar yang digunakan untuk menemukan pola baru dengan cara memilah beberapa data menggunakan teknologi penalaran pola serta teknik-teknik analisa yang terdapat pada data mining untuk menghasilkan informasi data yang akurat [5]. Untuk melakukan prediksi ketersediaan ICU pada pasien terjangkit virus Covid-19, data yang digunakan yaitu data COVID-19 patient pre-condition dataset yang bersumber dari website kaggle ([www.kaggle.com](http://www.kaggle.com)) yang akan dilakukan perhitungan menggunakan algoritma Naïve Bayes dengan optimasi PSO. Naive Bayes merupakan algoritma yang sering digunakan untuk melakukan prediksi karena perhitungan Naïve Bayes sederhana [6]. Dengan adanya asumsi bahwa setiap attribute bersifat independent satu sama lain, dimana setiap attribute mendapat perlakuan yang sama. PSO adalah teknik optimasi di mana solusi potensial terus dihitung menggunakan standar kualitas tertentu. Algoritma ini mengoptimalkan masalah dengan memindahkan partikel atau solusi yang diharapkan dalam ruang masalah menggunakan fungsi tertentu dari posisi dan kecepatan partikel [7].

Penelitian ini bertujuan untuk memprediksi kebutuhan ICU pada pasien terjangkit virus Covid-19 menggunakan metode NBC dengan optimasi PSO [8]. Hasil dari prediksi kebutuhan ICU dipergunakan sebagai bahan acuan untuk memenuhi ketersediaan ruang ICU untuk pasien terjangkit Covid-19 yang sangat membutuhkan, sehingga dapat mengurangi angka kematian pasien yang terlambat ditangani karena kurangnya persediaan ruang ICU yang ada di rumah sakit [9].

## 2 Tinjauan Literatur

Berdasarkan penelitian sebelumnya yang dilakukan oleh Kuncahyo, Istiadi, dkk dengan judul Optimasi naive Bayes classifier untuk klasifikasi teks pada e-government menggunakan particle swarm optimization, pada penelitian ini menggunakan metode K-NN dan NBC untuk mengetahui kinerja optimasi PSO dengan menggunakan 10 fold, sehingga didapatkan nilai NBC yang tidak menggunakan PSO sebesar 64,38% dan K-NN 75%, setelah dilakukan perhitungan menggunakan optimasi PSO maka didapatkan hasil NBC 87,44% [10].

Penelitian sebelumnya yang berhubungan dengan penelitian yaitu mendeteksi penyakit jantung menggunakan algoritma NBC berbasis PSO (*Particle Swarm Optimization*) menghasilkan nilai akurasi untuk nilai NBC sebesar 82,14% dengan nilai AUC (*Area Under Cover*) 0.686 dengan status "*Poor Clasification*" sedangkan hasil NBC yang menggunakan PSO mendapatkan hasil akurasi 92.86% dan nilai AUC 0.839 dengan status "*Good Clasification*" [11].

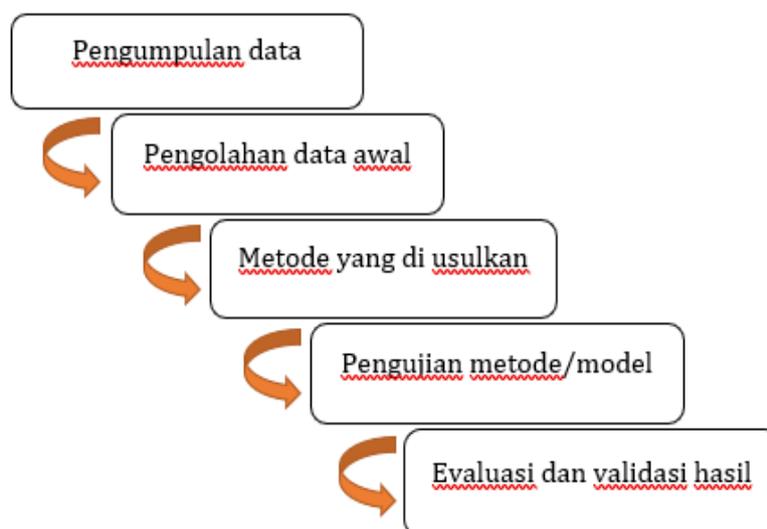
Penelitian Analisa sentiment yang bersumber dari twitter menggunakan *software rapid miner* untuk melakukan implementasinya untuk mendapatkan hasil yang baik. Hasil nilai akurasi yang paling tinggi didapatkan sebesar 76,38% dengan nilai AUC sebesar 0,787 dengan algoritma SVM menggunakan optimasi PSO. Pada algoritma NBC menggunakan optimasi PSO mengalami peningkatan nilai sebesar 6,38% dibandingkan tanpa PSO, sedangkan algoritma SVM mengalami kenaikan sebesar 3,83% dibanding tanpa PSO. Pada penelitian ini juga membuktikan bahwa hasil yang didapatkan menggunakan optimasi PSO lebih akurat dibanding tanpa optimasi PSO [12].

Penelitian yang dilakukan menggunakan dengan Classifier Random Forest yang digunakan untuk prediksi covid-19 berdasar hasil rontgen thorax dilakukan dengan beberapa tahapan dan didapatkan dua tahapan yang terbaik yaitu *feature extraction Histogram* dan *Hu-Moment* dengan nilai akurasi sebesar 84% dengan nilai standar deviasi sebesar 0.015847 dan didapatkan juga nilai kappa sebesar 0.713 [13].

Selanjutnya penelitian berbasis web untuk memprediksi jumlah pasien covid-19 menggunakan metode Trend Least Square akan dilakukan dengan menghitung nilai MAPE (*Mean Absolute Percentage Error*) dan didapatkan hasil prediksi pasien sembuh 84,2% dengan status buruk, pasien positif 52,4% dengan status buruk dan hasil prediksi pasien meninggal 40,9% dengan status layak/memadai [14].

### 3 Metode Penelitian

Pada metode penelitian ini didapatkan sumber data set rekam medis pasien Covid-19 dengan melakukan optimasi nilai Naïve Bayes menggunakan algoritma PSO untuk mendapatkan hasil nilai yang akurat [15]. Dalam penelitian ini akan dilakukan beberapa langkah-langkah atau tahapan penelitian seperti yang digambarkan pada Gambar 1.



Gambar 1. Metode Penelitian

#### 1. Pengumpulan data

Penelitian ini menggunakan Informasi Dataset Pra-Kondisi pasien Covid-19 yang teridentifikasi memerlukan perawatan ICU yang diambil dari situs Kaggle ([www.kaggle.com](http://www.kaggle.com)) yang diunggah pada tahun 2020. Data ini mempunyai 566.603 baris dan 23 kolom, data yang akan digunakan hanyalah sebagian dari beberapa kriteria data yang digunakan pada proses penelitian, yang bertujuan untuk menghindari error dalam pengolahan data. Selanjutnya, data akan dibagi menjadi 2 bagian, yaitu 60% untuk *data training* dan 40% *data testing*. Kriteria data yang digunakan adalah jenis kelamin, tipe pasien, pneumonia, diabetes, COPD, asma, inmsupr, hipertensi, penyakit lain, kardiovaskular, kegemukan, ginjal kronis, tembakau, kontak covid lain, covid res, dan ICU.

## 2. Pengolahan Data Awal

Pada tahapan pengolahan data awal akan dilakukan pembersihan data untuk memperoleh dan mengolah data mentah menjadi data yang bersih bebas dari noise atau outlire. Adaun beberapa tahapannya yatu:

### a. Data validation

Data validation digunakan untuk mengidentifikasi adanya data yang masih memiliki noise atau outlier, sehingga dapat diketahui data mana saja yang tidak lengkap.

### b. Data transformation

Data transformation terdiri dari beberapa data kategori, yang selanjutnya data yang didapatkan akan di transformasikan menjadi data numerik.

### c. Normalisasi data

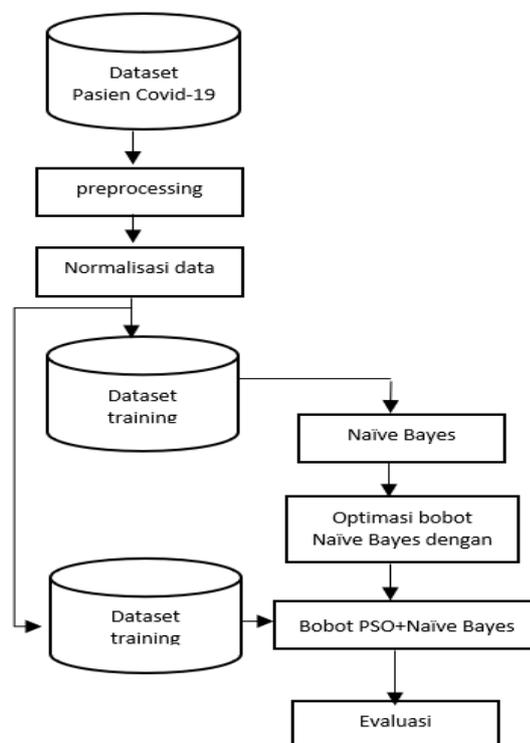
Pada penelitian ini digunakan normalisai min-max yang bertujuan untuk menghasilkan nilai atribut yang seimbang dalam rentang tertentu. Rentang nilai yang tidak seimbang pada tiap atribut akan mempengaruhi hasil dari data mining.

## 3. Metode yang diusulkan

Data yang diuji pada penelitian ini menggunakan algoritma Naïve Bayes dengan melakukan perhitungan terhadap variable yang digunakan dengan optimasi PSO, alat bantu dalam mencari perhitungan data menggunakan *software rapid miner* [16].

## 4. Pengujian metode

Pada tahap pengujian metode, akan dilakukan perhitungan menggunakan atribut yang tersedia menggunakan algoritma naïve bayes dan akan didapatkan hasil prediksi yang memiliki hasil akurat. Untuk tahapan pengujian ditunjukkan bawah Gambar 2 berikut.



**Gambar 2. Tahapan Pengujian**

## 5. Evaluasi dan validasi hasil

Evaluasi digunakan untuk menguji data menggunakan data training yang didapatkan dari data akhir yang akurat. Validasi merupakan bagian terpenting dari penelitian untuk mengetahui seberapa jauh kemampuan permodelan yang akan digunakan untuk mengambil suatu keputusan. Untuk mengetahui hasil dari evaluasi dan validasi dapat dilihat dari hasil confusion matrik [17].

## 4 Hasil dan Pembahasan

### a. Data Awal

Data awal yang digunakan adalah data set rekam medis pasien Covid-19 yang teridentifikasi memerlukan perawatan ICU yang diunggah Kaggle pada tahun 2020. Variable yang digunakan pada penelitian ini dapat dilihat pada Tabel 1.

**Tabel 1. Data Awal**

No	Atribut	Keterangan atribut	Keterangan Variabel
1	Id	Nomor id pasien	Contoh "16169f"
2	Sex	Jenis kelamin	1-Female (wanita), 2-Male (pria)
3	patient_type	Tipe rawat pasien	1-Outpatient (rawat jalan), 2-Inpatient (rawat inap)
4	entry_date	Tanggal masuk rumah sakit	Contoh "04-05-2020"
5	date_symptoms	Tanggal gejala pertama	Contoh "02-05-2020"
6	date_died	Tanggal kematian	Contoh "22-04-2020"
7	Intubed	Pasien pernah menggunakan bantuan alat pernafasan	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
8	Pneumonia	Memiliki riwayat radang paru-paru	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
9	Age	Umur	Contoh "27"
10	Pregnancy	Hamil	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
11	Diabetes	Kencing manis	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
12	Copd	Paru obstruktif kronik	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
13	Asthma	Asma	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
14	Inmsupr	Kekurangan imun dalam tubuh	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
15	Hypertension	Hipertensi	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
16	other_disease	Memiliki riwayat penyakit lain	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
17	Cardiovascular	Penyakit jantung	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
18	Obesity	Obesitas	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
19	renal_chronic	Ginjal kronis	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
20	Tobacco	Perokok	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
21	contact_other_covid	Kontak dengan pasien covid	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)
22	covid_res	Status covid	1-Positif, 2-Negatif, 3-Menunggu hasil
23	Icu	Masuk unit perawatan intensif	1-Yes (ya), 2-No (tidak) dan 97,98,99-NA (tidak diketahui)

Adapun dataset awal dari data pasien Covid-19 dapat dilihat pada Tabel 2 berikut.

**Tabel 2. Data Pasien Covid-19**

No	Sex	Patient_type	Pneumonia	Diabetes	copd	Asthma	Inmsupr	...	Icu
1	2	1	2	2	2	2	2	...	97
2	2	1	2	2	2	2	2	...	97
3	1	2	2	2	2	2	2	...	2
4	2	2	1	2		2	2	...	2
5	1	2	2	1	2	2	2	...	2
...	...	...	...	...	...	...	...	...	...
566603	1	1	2	2	2	1	2	...	2

b. Normalisasi Dataset

Dataset yang akan diolah selanjutnya melewati tahap pembersihan data dengan menghilangkan data yang missing dan tidak teridentifikasi. Dari total 566603 data kasus setelah dilakukan pembersihan data menjadi 32663 data kasus. Agar didapat data dengan rentang nilai yang sama maka tahap berikutnya dilakukan normalisasi terhadap dataset Covid-19 dengan menggunakan algoritma min-max. Setelah proses normalisasi, data dibagi menjadi 60% data training dan 40% data testing.

c. Data Training

Untuk data training digunakan data 60% dari total data yang sudah di preprocessing. Data training yang digunakan dapat dilihat pada Tabel 3.

**Tabel 3. Data Training**

No	Sex	Patient_type	Pneumonia	Diabetes	Copd	Asthma	Inmsupr	...	Icu
1	1	0	0	0	0	0	0	...	0
2	1	0	0	0	0	0	0	...	0
3	0	0	1	0	0	0	0	...	0
4	1	0	1	0	0	0	0	...	0
5	1	0	1	1	0	0	0	...	0
...	...	...	...	...	...	...	...	...	...
19598	1	0	1	1	0	0	0	...	0

d. Data Testing

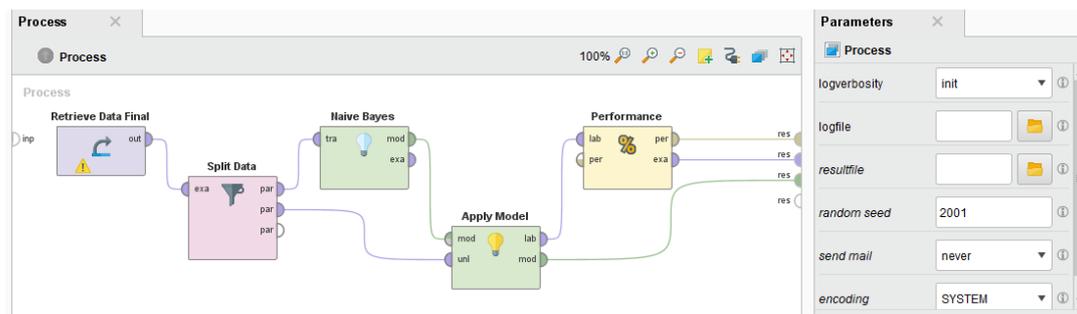
Data testing yang digunakan sebesar 40% dari total data yang sudah di pre processing. Hasil dari data testing dapat dilihat pada Tabel 4 berikut:

**Tabel 4. Data Testing**

No	Sex	Patient_type	Pneumonia	Diabetes	copd	Asthma	Inmsupr	...	Icu
19599	0	0	1	0	0	0	0	...	0
19600	0	0	0	0	0	0	0	...	0
19601	1	0	0	0	0	0	1	...	0
19602	1	0	1	1	0	0	0	...	0
19603	0	0	1	1	0	0	0	...	0
...	...	...	...	...	...	...	...	...	...
32663	0	0	0	0	0	0	0	...	0

e. Memasukkan data dari Repository serta operator yang dibutuhkan ke dalam kolom process

Pada proses ini, akan dimasukkan operator split data yang berfungsi untuk membagi data menjadi 40% data testing dan 60% data training yang akan di masukkan operator Naïve Bayes, Apply Model, Performance ke dalam kolom proses yang ada pada *software rapid miner* seperti terlihat pada Gambar 3.



**Gambar 3. Proses Pengolahan Data Naïve Bayes**

f. Hasil Pengolahan data dengan algoritma NBC.

Hasil pengolahan dataset pasien Covid-19 dengan algoritma NBC menghasilkan tingkat akurasi sebesar 86,53% dengan confusion matrix seperti terlihat pada Tabel 5 berikut.

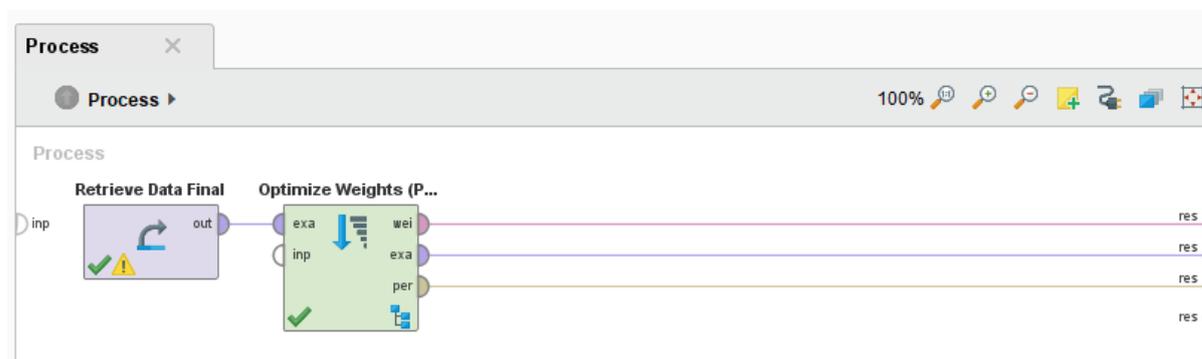
**Tabel 5. Confusion Matrix Dari Algoritma NBC**

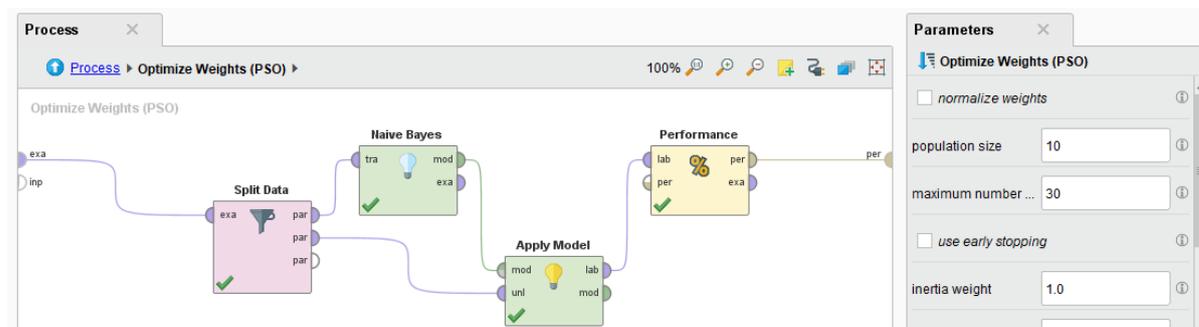
	True 0	True 1	Class precision
Pred 0	11305	1760	86,53%
Pred 1	0	0	0%
Class recall	100%	0%	

Algoritma NBC menghasilkan nilai accuracy sebesar 86,53% dengan data testing sebesar 40% (13065 Data). Pada 13065 data terdapat 11305 data yang tepat prediksinya dengan nilai aktual icu = 0, 1760 data dengan nilai aktual icu = 1 hasil nilai prediksi icu = 0.

g. Optimasi hasil klasifikasi Algoritma NBC dengan *Partical Swarm Optimization*

Berdasarkan optimasi dari metode PSO maka akan dipastikan kembali posisi klasifikasi pada NBC yang bisa berubah karena bobot masing-masing atributnya. Untuk mendapatkan hasil dari optimasi, maka tahap yang harus dilakukan yaitu, memasukkan data ke dalam operator PSO (Gambar 4). Data akan dibagi menjadi 0,4 (40%) data testing dan 0,6 (60%) data training dengan memasukkan population size di tiap optimasinya.





**Gambar 4. Proses Pengolahan Data Menggunakan Optimasi PSO**

1) Hasil Optimasi dengan *Population Size* = 20 (*Execution Time* : 43 s)

Dengan melakukan optimasi menggunakan *Particle Swarm Optimization (PSO)* pada Algoritma NBC menghasilkan nilai accuracy sebesar 87,03% dengan *population size* = 20 yang semula sebesar 87,00% pada *population size* = 10 menggunakan data testing sebesar 40% (13065 Data). Pada 13065 data, terdapat 11370 data yang tepat prediksinya pada actual icu = 0, sementara 1695 data yang aktualnya icu = 1 menghasilkan prediksi icu = 0. Hasil confusion matrix sebagaimana terlihat pada Tabel 6.

**Tabel 6. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO 10 Population Size**

	True 0	True 1	Class precision
Pred 0	11370	1695	87,03%
Pred 1	0	0	0%
Class recall	100%	0%	

2) Hasil Optimasi dengan *Population Size* = 40 (*Execution Time* : 02 m 37 s)

Dengan melakukan optimasi menggunakan *Particle Swarm Optimization (PSO)* pada Algoritma NBC pada *population size* = 40, menghasilkan nilai accuracy sebesar 87,28% yang semula sebesar 87,15% pada *population size* = 30 menggunakan data testing sebesar 40% (13065 Data). Pada 13065 data terdapat 11403 data yang tepat prediksinya pada aktual icu = 0, sementara 1662 data yang aktualnya icu = 1 menghasilkan prediksi icu = 0. Hasil Confusion matrix dapat dilihat pada Tabel 7.

**Tabel 7. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO 30 Population Size**

	True 0	True 1	Class precision
Pred 0	11403	1662	87,28%
Pred 1	0	0	0%90
Class recall	100%	0%	

3) Hasil Optimasi dengan *Population Size* = 60 (*Execution Time* : 05 m 05 s)

Dengan melakukan optimasi menggunakan *Particle Swarm Optimization (PSO)* pada Algoritma NBC pada *population size* = 60, menghasilkan nilai accuracy sebesar 87,13% yang semula sebesar 87,16% pada *population size* = 50 menggunakan data testing sebesar 40% (13065 Data). Pada 13065 data terdapat 11383 data yang tepat prediksinya pada actual icu = 0, sementara 1682 data yang aktualnya icu = 1 menghasilkan prediksi icu = 0. Hasil confusion matrix dapat dilihat pada Tabel 8.

**Tabel 8. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO 50 Population Size**

	True 0	True 1	Class precision
Pred 0	11383	1682	87,13%
Pred 1	0	0	0%
Class recall	100%	0%	

4) Hasil Optimasi dengan *Population Size* = 80 (*Execution Time* : 08 m 53 s)

Dengan melakukan optimasi menggunakan *Particle Swarm Optimization (PSO)* pada Algoritma NBC menghasilkan nilai accuracy sebesar 87,16% dengan *population size* = 80 yang semula sebesar 87,27% pada *population size* = 70 menggunakan data testing sebesar 40% (13065 Data). Pada 13065 data terdapat 11388 data yang tepat prediksinya pada actual icu = 0, sementara 1677 data yang aktualnya icu = 1 menghasilkan prediksi icu = 0. Hasil confusion matrix dapat dilihat pada Tabel 9.

**Tabel 9. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO 70 Population Size**

	True 0	True 1	Class precision
Pred 0	11388	1677	87,16%
Pred 1	0	0	0%
Class recall	100%	0%	

5) Hasil Optimasi dengan *Population Size* = 100 (*Execution Time* : 14 m 20 s)

Dengan melakukan optimasi menggunakan *Particle Swarm Optimization (PSO)* pada Algoritma NBC menghasilkan nilai accuracy sebesar 87,26% dengan *population size* = 100 yang semula sebesar 87,22% pada *population size* = 90 menggunakan data testing sebesar 40% (13065 Data). Pada 13065 data terdapat 11400 data yang tepat prediksinya pada actual icu = 0, sementara 1665 data yang aktualnya icu = 1 menghasilkan prediksi icu = 0. Hasil confusion matrix dapat dilihat pada Tabel 10.

**Tabel 10. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO 90 Population Size**

	True 0	True 1	Class precision
Pred 0	11400	1665	87,26%
Pred 1	0	0	0%
Class recall	100%	0%	

Hasil keseluruhan dari prediksi Naïve Bayes sebelum dan Sesudah Optimasi dengan PSO dapat dilihat pada Tabel 11.

**Tabel 11. Confusion Matrix Dari Algoritma NBC Dengan Optimasi PSO**

<i>Population Size</i>	<i>Accuracy</i>	<i>Waktu Eksekusi (Menit)</i>
-	86,53%	00:00
20	87,03%	00:43
40	87,28%	02:37
60	87,13%	05:05
80	87,16%	08:53
100	87,26%	14:20

## 5 Kesimpulan

Penelitian ini bertujuan untuk memprediksi kebutuhan ICU pada pasien terjangkit virus Covid-19 menggunakan metode Naïve Bayes pada optimasi PSO dengan *software rapid miner*. Hasil nilai dari prediksi kebutuhan ICU dipergunakan sebagai bahan acuan untuk memenuhi kebutuhan ICU untuk pasien terjangkit Covid-19 sehingga dapat mengurangi kurangnya persediaan ICU di rumah sakit. Hasil akurasi yang didapat sebesar 86,53% dengan total data sebanyak 32663. Data dibagi menjadi 60% data training dan 40% data testing. Berikutnya dilakukan optimasi bobot pada Naïve Bayes dengan menggunakan algoritma *Particle Swarm Optimization*. Percobaan dilakukan dengan ragam jumlah populasi yaitu 20, 40, 60, 80 dan 100. Peningkatan jumlah populasi pada PSO memperlihatkan nilai akurasi yang semakin meningkat akan tetapi waktu eksekusi juga semakin meningkat. Nilai akurasi terbaik didapatkan dengan population size 100 yaitu sebesar 87,26% dengan waktu eksekusi 14:20.

## Ucapan Terima Kasih

Terima kasih peneliti ucapkan kepada semua pihak yang ikut serta dalam terlaksananya penelitian yang sudah peneliti lakukan, terutama pimpinan dan staff Fakultas Ilmu Komputer Universitas Lancang Kuning, tim peneliti dan teman-teman seperjuangan yang selalu memberi dukungan untuk menyelesaikan penelitian ini. Terimakasih juga peneliti ucapkan kepada Tim Jurnal Sistemasi.

## Referensi

- [1] D. Y. Liliana, H. Maulana, and A. Setiawan, "Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes," *Multinetics*, vol. 7, no. 1, pp. 48–53, 2021, doi: 10.32722/multinetics.v7i1.3786.
- [2] L. Herlina, W. R. Agustin, and I. K. Mustikarani, "Pengalaman Perawat dalam Memberikan Perawatan End of Life pada Pasien Covid-19 di Ruang ICU Rumah Sakit Umum Daerah Dr. Moewardi," *Univ. Kusuma Husada Surakarta*, 2021.
- [3] A. K. Amirullah and K. Kartinah, "Penanganan Kecemasan Pasien Survivor Covid-19 Intensive Care Unit: Literature Review," in *Prosiding Seminar Nasional Keperawatan Universitas Muhammadiyah Surakarta 2020 (Profesi Ners XXII)*, 2020, pp. 8–15.
- [4] E. Nurelasari, "Komparasi Algoritma Naïve Bayes dengan Support Vector Machine Berbasis Particle Swarm Optimization untuk Prediksi Kesuburan," *Bina Insa. ICT J.*, vol. 5, no. 1, pp. 61–70, 2018.
- [5] Ratino, N. Hafidz, S. Anggraeni, and W. Gata, "Sentimen Analisis Informasi Covid-19 menggunakan Support Vector Machine dan Naïve Bayes," *J. JUPITER*, vol. 12, no. 2, pp. 1–11, 2020.
- [6] T. Mardiana, "Optimasi Naïve Bayes dengan Particle Swarm Optimization dan Stratified untuk Prediksi Kredit Macet pada Koperasi," *J. Ris. Inform.*, vol. 1, no. 1, pp. 43–50, 2018, doi: 10.34288/jri.v1i1.13.
- [7] T. S. N. Koeswara, M. S. Mardiyanto, and M. A. Ghani, "Penerapan Particle Swarm Optimization (PSO) dalam Pemilihan Atribut untuk Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis dengan Metode Naïve Bayes," *J. Speed – Sentra Penelit. Eng. dan Edukasi*, vol. 12, no. 1, pp. 1–10, 2020.
- [8] F. Novaldy and A. Herliana, "Penerapan PSO pada Naïve Bayes untuk Prediksi Harapan Hidup Pasien Gagal Jantung," *J. Responsif Ris. Sains dan Inform.*, vol. 3, no. 1, pp. 37–43, 2021, doi: 10.51977/jti.v3i1.396.
- [9] F. Ratnawati, "Implementasi Algoritma Naïve Bayes terhadap Analisis Sentimen Opini Film pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [10] K. S. Nugroho, F. Marisa, I. Istiadi, and F. Marisa, "Optimasi Naïve Bayes classifier untuk klasifikasi Teks pada e-Government menggunakan Particle Swarm Optimization/ Naïve Bayes Classifier Optimization for Text Classification on e-Government using Particle Swarm Optimization," *J. Teknol. dan Sist. Komput.*, vol. 8, no. January, pp. 21–26, 2020, doi:

- 10.14710/jtsiskom.8.1.2020.21-26.
- [11] D. Evanko, "Optical Imaging of the Native Brain," *Nat. Methods*, vol. 7, no. 1, p. 34, 2010, doi: 10.1038/nmeth.f.284.
- [12] F. Astuti and R. Taufan, "Sentimen Analisis Vaksinasi Covid - 19 pada Twitter menggunakan Algoritma Klasifikasi Berbasis PSO/ Sentiment Analysis of Covid-19 Vaccination on Twitter using Classification Algorithms based on PSO," *Sist. J. Sist. InformasiEMASI*, vol. 11, pp. 364–376, 2022.
- [13] A. Mustopa *et al.*, "Ekstraksi Fitur dengan Classifier Random Forest untuk Memprediksi Covid 19 Berdasarkan Hasil Rontgen Thorax Feature Extraction with Random Forest Classifier to Predict Covid 19 Based on Results Thorax X - ray," *Sist. J. Sist. Inf.*, vol. 11, pp. 515–525, 2022.
- [14] J. S. Widjaya, D. A. R, and S. R. Puspitasari, "Sistem Prediksi Jumlah Pasien Covid-19 menggunakan Metode Trend Least Square Berbasis Web," *Sist. J. Sist. Inf.*, vol. 10, pp. 39–51, 2021.
- [15] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, "Optimasi Naive Bayes Classifier dengan Menggunakan Particle Swarm Optimization pada Data Iris," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 180, 2017, doi: 10.25126/jtiik.201743251.
- [16] T. Arifin and D. Ariesta, "Prediksi Penyakit Ginjal Kronis menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization," *J. Tekno Insentif*, vol. 13, no. 1, pp. 26–30, 2019, doi: 10.36787/jti.v13i1.97.
- [17] S. Sumarni and S. Rustam, "Klasifikasi Topik Tugas Akhir Mahasiswa menggunakan Algoritma Particle Swarm Optimization dan k-Nearest Neighbor," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 168–175, 2020, doi: 10.33096/ilkom.v12i2.604.168-175.