

# Penerapan Metode Naïve Bayes Classifier untuk Klasifikasi Penyakit Infeksi Saluran Pernapasan Akut (ISPA)

## *Application of Naïve Bayes Classifier Method for Classification of Acute Respiratory Infection (ARI)*

Rahmat Thaib  
Program Studi Teknik Informatika  
Universitas Puhuwato  
Pohuwato, Indonesia  
rahmatthaib@email.com

Betrisandi  
Program Studi Teknik Informatika  
Universitas Puhuwato  
Pohuwato, Indonesia  
betris.sin@email.com

Diterima : Mei 2025  
Disetujui : Juli 2025  
Dipublikasi : Juli 2025

**Abstrak**— ISPA adalah salah satu masalah global yang mempengaruhi jutaan individu setiap tahunnya, termasuk di Kabupaten Pohuwato. Gejala yang ditimbulkan bervariasi seperti pilek, sakit tenggorokan, batuk, dan sesak napas yang umumnya disebabkan oleh infeksi bakteri atau virus. Jenis ISPA dapat diklasifikasikan menjadi ringan, sedang, dan berat. Permasalahan dalam penelitian ini adalah kurangnya pengetahuan masyarakat terkait penanganan ISPA dan tingginya angka kejadian penyakit ISPA. Penelitian ini bertujuan untuk mengklasifikasikan penyakit ISPA berdasarkan usia pasien dan lama rawat inap dengan harapan dapat membantu para tenaga medis di puskesmas dalam memberikan penanganan yang cepat serta pengobatan yang tepat kepada para penderita ISPA. Teknik pengumpulan data dilakukan dengan survei langsung ke Puskesmas Marisa, Kabupaten Pohuwato, Provinsi Gorontalo. Analisis data dilakukan dengan data *mining* yang menekankan pada klasifikasi penyakit ISPA dengan metode *Naïve Bayes Classifier*. Hasil dari penelitian ini menunjukkan akurasi yang tinggi yaitu 96.61% untuk prediksi ringan, 95,60% untuk prediksi sedang, dan 100% untuk prediksi berat. Jadi metode *Naïve Bayes Classifier* mampu menghasilkan prediksi yang tepat dalam mengidentifikasi kasus ISPA.

**Kata Kunci**—*Klasifikasi; ISPA; Naïve Bayes Classifier; Penyakit; Diagnosis*

*ARI is a global problem that affects millions of individuals every year, including in Pohuwato District. Symptoms include runny nose, sore throat, cough, and shortness of breath which are generally caused by bacterial or viral infections. Types of ARI can be classified into mild, moderate, and severe. The problem in this study is the lack of public knowledge related to ARI management and the high incidence of ARI disease. This study aims to classify ARI diseases based on patient age and length of hospitalisation in the hope that it can help medical personnel at the puskesmas in providing fast handling and appropriate treatment to patients with ARI. Data collection techniques were carried out by direct survey to the Marisa Health Centre, Pohuwato Regency, Gorontalo Province. Data analysis is done by data mining which emphasises the classification of ARI diseases with the Naïve Bayes Classifier method. The results of this study showed high accuracy of 96.61%*

*for mild prediction, 95.60% for moderate prediction, and 100% for severe prediction. So the Naïve Bayes Classifier method is able to produce correct predictions in identifying ARI cases.*

**Keywords**— *Classification; ARI; Naïve Bayes Classifier; Disease; Diagnosis*

### I. PENDAHULUAN

Infeksi Saluran Pernapasan Akut (ISPA), umumnya merupakan jenis penyakit yang banyak ditemukan di masyarakat dengan penderita yang tidak hanya orang dewasa dan remaja tetapi juga anak-anak dan balita. ISPA merupakan masalah kesehatan yang serius terutama pada anak berusia 1 hingga 5 tahun, dan menjadi penyebab utama kematian anak-anak di Negara yang berkembang [1]. Penyakit ISPA dapat disebabkan oleh berbagai jenis virus atau bakteri yang menyerang hidung atau trakea (saluran pernapasan) yang menimbulkan gejala bervariasi seperti batuk, pilek, sakit tenggorokan dan sesak napas [2,3]. Adapun jenis-jenis ISPA dapat dikelompokkan menjadi tiga kategori diantaranya ISPA ringan, ISPA sedang dan ISPA berat [4]. Pengklasifikasian penyakit ISPA didasarkan pada tingkat keparahan penyakitnya.

ISPA adalah penyakit yang menyerang saluran bernapas dan membuatnya tidak berfungsi dengan baik. Jika ISPA tidak diobati dengan benar, maka dapat menimbulkan masalah lain yang bisa berakibat sangat serius bagi yang mengalaminya merupakan infeksi [4]. ISPA tidak hanya terjadi pada orang dewasa, tetapi juga pada anak-anak [5]. ISPA merupakan penyakit balita berbahaya pertama di dunia dibandingkan penyakit lainnya [6] sehingga menjadi salah satu faktor yang berkontribusi pada tingginya tingkat kematian dan penyakit dikalangan anak di Indonesia [5].

Biasanya penyakit ISPA menyebar luas saat musim pancaroba karena meningkatnya pergerakan virus di udara serta menurunnya daya tahan tubuh yang mengakibatkan banyak yang rentan terhadap penyakit ini. Dalam penanganan penyakit ISPA, diagnosis yang cepat dan tepat serta

pengobatan yang intensif merupakan hal yang krusial untuk menghindari komplikasi yang lebih serius. Namun, biasanya proses diagnosis penyakit ISPA memerlukan waktu yang cukup lama, keahlian yang tinggi, serta biaya yang tidak murah. Dalam situasi ini, menemukan dan mengenali ISPA dengan cepat sangatlah krusial untuk memberikan perawatan yang sesuai dan mengurangi penyebaran ISPA. Untuk mendiagnosis ISPA, biasanya diperlukan pemeriksaan oleh tenaga medis yang memiliki pengalaman [7].

Pusat Kesehatan Masyarakat atau Puskesmas adalah lembaga yang penting untuk membantu perubahan kesehatan masyarakat agar bisa mencapai tingkat Kesehatan yang lebih baik [8]. Puskesmas sebagai salah satu tempat pelayanan kesehatan masyarakat yang menjadi acuan untuk pengobatan bagi warga masyarakat harus mampu mengidentifikasi jenis penyakit ISPA yang tepat [4].

Pengidentifikasian penyakit ISPA di puskesmas masih belum mendetail. Hal inilah yang menjadi penyebab penanganan terhadap penyakit ISPA belum maksimal. Oleh sebab itu, informasi terkait klasifikasi dari penyakit ISPA yang diderita oleh masyarakat sangat penting dan dibutuhkan untuk dapat membantu para tenaga medis di puskesmas dalam memberikan penanganan yang cepat serta pengobatan yang tepat kepada para penderita ISPA baik orang dewasa, anak-anak, dan balita berdasarkan jenis ISPA.

Data *mining* adalah bidang ilmu yang dapat mengidentifikasi pola dari sekumpulan data yang dapat menghasilkan informasi yang diinginkan dan sebuah proses penting dalam rangka menemukan [3]. Data *mining* melibatkan penelusuran terhadap tren atau pola yang relevan dalam basis data yang besar untuk mendukung para pengambil keputusan di masa mendatang. Pola-pola tersebut teridentifikasi oleh perangkat tertentu yang mampu menawarkan analisis data yang bermanfaat dan informatif. Analisis ini dapat dipelajari dengan lebih mendalam, yang mungkin juga menggunakan alat pendukung keputusan yang lainnya [9].

Dalam penambangan data, terdapat metode klasifikasi yang digunakan untuk meneliti suatu informasi sehingga dapat menghasilkan representasi kategori data yang relevan [10]. Klasifikasi merupakan sebuah metode untuk menemukan serangkaian model yang dapat menjelaskan dan membedakan kategori data, dengan harapan model tersebut mampu meramalkan kategori dari objek yang kelasnya masih tidak diketahui [11,12]. Berbagai metode klasifikasi yang sering digunakan secara umum mencakup pohon keputusan atau klasifikasi, pengklasifikasi *Bayesian* atau pengklasifikasi *Naïve Bayes*, Jaringan saraf, analisis *statistic*, algoritma genetik, *K-Nearest Neighbors* (KNN), metode berbasis aturan dan mesin *vector* pendukung atau *Support Vector Machine* (SVM) [12].

Salah satu metode klasifikasi yang telah terbukti efektif dalam berbagai aplikasi adalah *Naïve Bayes Classifier* (NBC). Metode ini berbasis pada Teorema *Bayes* dan mengasumsikan indenpendensi antara setiap pasangan fitur. Metode *Naïve Bayes Classifier* menggunakan analisis probabilitas dan statistika yang bisa digunakan untuk meramalkan kemungkinan di masa mendatang berdasarkan pengalaman yang sudah terjadi dimasa lampau [13]. Metode ini banyak diterapkan pada hal-hal yang berkenaan dengan diagnosa secara statistik yang berhubungan dengan probabilitistik serta kemungkinan dari penyakit dan gejala

gejala yang berkaitan [14]. Tidak heran, metode *Naïve Bayes Classifier* banyak digunakan dalam penelitian-penelitian sebelumnya.

Pada penelitian sebelumnya terkait dengan Impelementasi Metode *Naïve Bayes Classifier* pada Aplikasi Prediksi Penyebaran Wabah Penyakit ISPA (Studi Kasus: Wilayah Kota Pontianak), diperoleh persentase sebesar 82,97%[2]. Sedangkan pada penelitian ini analisis data dilakukan dengan data *mining* yang lebih menekankan pada pengklasifikasian penyakit ISPA agar dapat bermanfaat seperti yang sudah dijelaskan pada paragraf sebelumnya.

Dengan demikian, penggunaan metode *Naïve Bayes Classifier* pada penelitian ini diharapkan dapat membantu dalam mengklasifikasi penyakit ISPA. ISPA. Penelitian – penelitian terdahulu telah menunjukkan potensi *Naïve Bayes Classifier* dalam mendukung diagnosis berbagai penyakit termasuk ISPA. Metode akurasi untuk metode *Naïve Bayes Classifier* menghasilkan nilai yang bervariasi tergantung pada persentase pembagian data uji dan data latih [15]. Penelitian yang telah dilakukan sebelumnya berhasil menerapkan metode *Naïve Bayes Classifier* dalam klasifikasi ISPA berdasarkan penyakit yang diderita, umur dan lama rawat inap. Hasil penelitian menunjukkan bahwa *Naïve Bayes Classifier* dapat mengklasifikasikan ke dalam 3 jenis ISPA yaitu ringan, sedang dan berat dengan hasil akurasi sebesar 93,33% [4].

Dalam penelitian ini akan melakukan klasifikasi atau pengelompokkan penyakit ISPA berdasarkan faktor – faktor penentu penyakit ISPA yang diderita pasien yaitu umur pasien dan lama idap pasien sebagai fitur unik dalam penelitian ISPA. Hasil penelitian ini diharapkan mampu memberikan informasi mengenai ketepatan penggunaan *Naïve Bayes Classifier* dalam mengelompokkan penyakit ISPA berdasarkan jenis-jenisnya sehingga membantu pihak puskesmas dalam penanganan penyakit ISPA yang cepat dan tepat.

## II. METODE

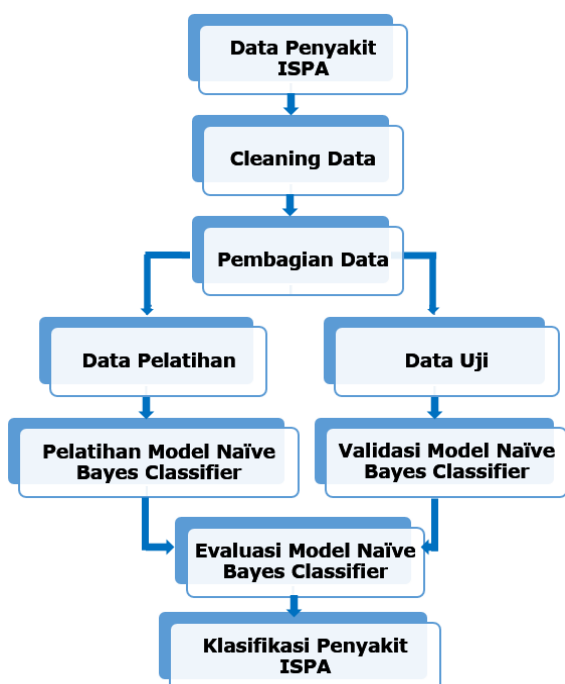
Metode pada penelitian ini menggunakan metode eksperimen dengan harapan bisa memberikan hasil yang efektif. Beberapa tahapan yang akan dilakukan diantaranya yaitu tahapan pertama penelitian dimulai dari identifikasi masalah adanya penyakit ISPA yang membutuhkan suatu sistem dalam klasifikasi berdasarkan gejala. Tahap kedua pengumpulan gejala dan diagnosa ISPA. Tahap ketiga pemrosesan data meliputi pembersihan data, *encoding* fitur dan normalisasi. Pada tahap ini data yang diperoleh dari puskesmas akan dilakukan pemrosesan data meliputi pembersihan data duplikat dan *missing values*, selanjutnya mengubah format data agar sesuai dengan kebutuhan algoritma seperti menyamakan skala data numerik serta menentukan *variable* yang paling relevan untuk klasifikasi yaitu umur, jenis penyakit dan lama inap. Tahap keempat pembagian data menggunakan metode *cross validation* yaitu data latih (*training*) sebanyak 70 % dan data uji (*testing*) sebanyak 30%. Tahap kelima yaitu pelatihan model dengan menerapkan metode *Naïve Bayes Classifier*. Metode ini menggunakan analisis probabilitas dan diagnosa secara statistika termasuk mendukung diagnosis berbagai penyakit sehingga memudahkan dalam mengkasifikasikan penyakit ISPA dengan akurasi yang baik dan efektif. Tahap terakhir yaitu evaluasi model yaitu akurasi, presisi, *recall* dan

intrepretasi hasil serta kesimpulan. Gambaran tahapan tersebut dapat dilihat pada gambar 1 dibawah ini.



Gambar 1. Tahapan Penelitian

Berdasarkan tahapan penelitian yaitu identifikasi permasalahan maka pendekatan dan strategi pemecahan masalah dalam penelitian ini yaitu klasifikasi yang merupakan salah satu metode dalam data *mining* yang bertujuan untuk memprediksi label atau kategori dari suatu data berdasarkan data sebelumnya. Proses ini termasuk dalam pembelajaran terawasi dimana model dilatih dengan data yang sudah memiliki label/kategori. Dataset nantinya akan diproses menggunakan *software* rapidminer dengan *tools* AI Studio versi 2024.1.0 dengan tautan berikut, <https://rapidminer.com/products/studio/>. Penelitian ini nantinya akan melakukan klasifikasi penyakit ISPA berdasarkan gejala dan faktor – faktor penentu. Pengklasifikasian penyakit ISPA dilakukan dengan cara menganalisis dan mencari kemiripan data berdasarkan pada faktor-faktor penentu seperti usia pasien dan masa idap pasien. Selanjutnya hasil klasifikasi tersebut diberikan kode tertentu untuk mempermudah dalam memprediksi data. Gambaran aliran eksperimen seperti pada gambar 2 dibawah.



Gambar 2. Aliran Eksperimen

Data penyakit ISPA yang diperoleh dari Puskesmas Marisa, Kabupaten Pohuwato, Gorontalo, melalui wawancara langsung dengan tenaga medis diklasifikasikan menjadi 5 kelompok berdasarkan tingkat keparahan penyakit, mulai dari penyakit ISPA yang paling ringan sampai penyakit ISPA yang paling berat. Setiap penyakit memiliki kode masing-masing yaitu A01 (*Influenza Like Illness*), A02 (*Rhinosinusitis/sinusitis*), A03 (*Tonsilitis, Faringitis, dan Laringitis*), A04 (*Faringitis akuta*), dan A05 (*Pneumonia*). Selanjutnya hasil klasifikasi dibagi menjadi dua data yakni data latih dan data uji kemudian diuji coba pada *software* rapidminer menggunakan *Naïve Bayes Classifier*.

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Data penyakit ISPA berjumlah 787 data pasien. Beberapa contoh data pasien dapat dilihat pada Tabel 1.

TABEL 1. DATA PENDERITA ISPA

Pasien	Umur	Jenis Penyakit	Lama Idap	Klasifikasi ISPA
Pasien 1	0-11 bulan	A01	B1	Ringan
Pasien 2	0-11 bulan	A01	B1	Ringan
Pasien 3	0-11 bulan	A01	B1	Ringan
Pasien 4	0-11 bulan	A01	B1	Ringan
Pasien 5	1-4 tahun	A05	B1	Berat
Pasien 6	1-4 tahun	A01	B2	Ringan
Pasien 7	1-4 tahun	A01	B1	Ringan
Pasien 8	9-10 tahun	A02	B2	Sedang
Pasien 9	9-10 tahun	A02	B2	Sedang
Pasien 10	9-10 tahun	A02	B2	Sedang
Pasien 11	11-59 tahun	A02	B2	Sedang
Pasien 12	11-59 tahun	A02	B2	Sedang
Pasien 13	11-59 tahun	A02	B2	Sedang
Pasien 14	60 $\geq$ tahun	A02	B2	Ringan
Pasien 15	60 $\geq$ tahun	A02	B2	Sedang

Dari Tabel 1 bisa dilihat bahwa pasien yang mengalami ISPA memiliki kriteria usia, jenis penyakit, lama idap dan Klasifikasi ISPA. Usia pasien mulai 0 bulan, untuk memudahkan dalam pengolahan data untuk kriteria jenis penyakit kategori numerik ada 5 seperti yang sudah dijelaskan sebelumnya. Untuk masa idap terdiri dari 2 yaitu baru idap dengan kode B1 dan lama idap dengan kode B2 sedangkan kriteria kategori ISPA terdiri 3 yaitu ringan, sedang dan berat.

#### B. Pemrosesan Data

Data akan diproses menggunakan aplikasi Rapidminer dengan terlebih dahulu melakukan *cleaning* data dengan tujuan untuk memastikan keakuratan data yang digunakan dalam analisis serta dapat diandalkan. Proses yang dilakukan adalah mendeteksi kesalahan data dan merevisi atau menghapus data sesuai dengan kebutuhan. Hasil dari *cleaning* data tidak ditemukan adanya kesalahan data selanjutnya data ISPA ini siap untuk diolah.

#### C. Pembagian Data

Pada tahapan ini dilakukan pembagian data dengan *random split* yang tujuannya membagi data ke dalam dua kelompok, pertama data latih (*training*) dan kedua data uji (*testing*). Data latih digunakan untuk melatih model dan

menemukan pola dengan persentase 70% berjumlah 551 sedangkan data uji untuk mengetahui data yang masih belum teridentifikasi kategori kelasnya lalu dikelompokkan dari data latih yang telah dibangun dengan persentase 30% berjumlah 236 untuk data uji.

**D. Pelatihan Model**

Pelatihan mode dilakukan dengan menerapkan model atau algoritma yang dipilih yaitu *Naive Bayes Classifier* dengan cara mencari pola klasifikasi berdasarkan kriteria atau faktor penentu yaitu usia, jenis penyakit dan lamanya sakit. Klasifikasi *Naive Bayes* merupakan algoritma yang memiliki akurasi tinggi meskipun merupakan metode yang cukup sederhana.

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (1)$$

Dengan:

- X: data dengan *class* yang belum diketahui
- C<sub>i</sub>: hipotesis data X merupakan suatu *class* spesifik
- P(C<sub>i</sub>|X): probabilitas hipotesis C<sub>i</sub> berdasarkan kondisi X (posteriori probability)
- P(C<sub>i</sub>): probabilitas hipotesis C<sub>i</sub> (prior probability)
- P(X|C<sub>i</sub>): probabilitas X berdasarkan kondisi pada hipotesis C<sub>i</sub>
- P(X): probabilitas dari X

*Naive Bayesian Classifier* mengasumsikan bahwa keberadaan sebuah atribut (variabel) tidak ada kaitannya dengan beradaan atribut (variabel) yang lain karena asumsi atribut tidak saling terkait (*conditionally independent*), ditulis dengan rumus:

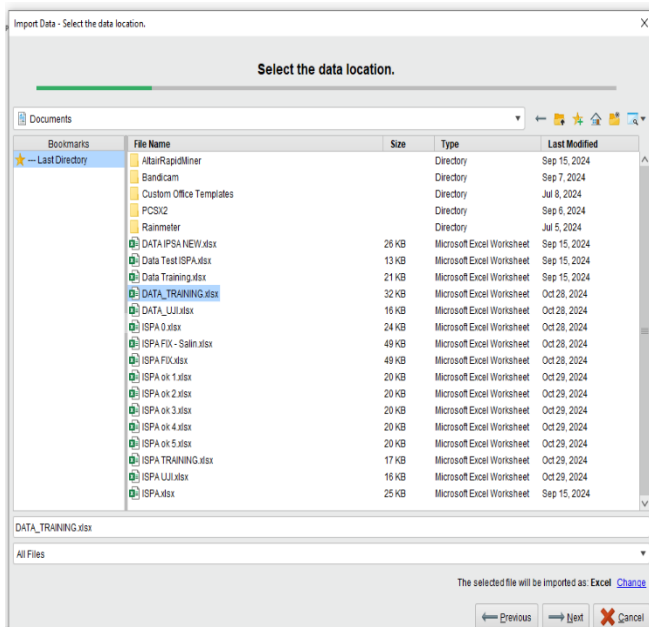
$$P(X|C_i) = \prod_{k=1}^n P(X_k|C_i) \quad (2)$$

Setelah diperoleh hasil dari seluruh data pada setiap class, maka hasil akhirnya dapat menggunakan rumus:  $P(X|C_i) = \arg \max P(X_i|C_i) * P(C_i)$ .

**E. Validasi Model**

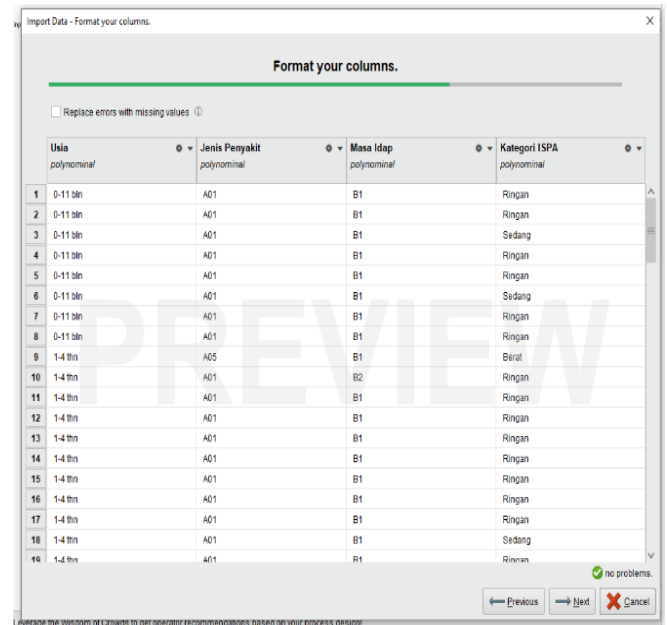
Berikut tampilan desain untuk validasi model pada rapidminer untuk klasifikasi penyakit ISPA seperti yang ditunjukkan pada gambar 1 sampai gambar 11 berikut.

Langkah pertama yang dilakukan yaitu proses *import data excel*.



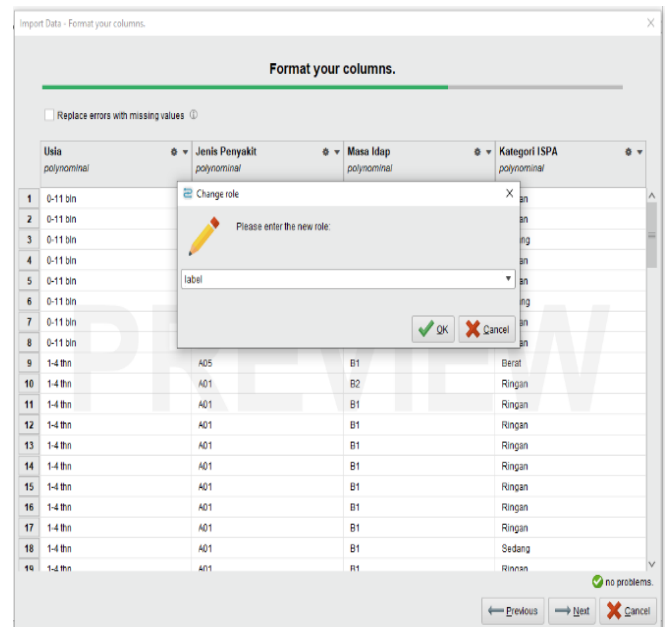
Gambar 3. Proses *Import Data Excel Step 1*

Langkah kedua yaitu memilih atribut pada data yang akan diproses selanjutnya. Pemilihan atribut adalah proses untuk memilih fitur atau atribut yang paling relevan dari seluruh data yang tersedia untuk digunakan dalam proses pemodelan yaitu usia, jenis penyakit, rawat inap dan kategori ISPA seperti pada gambar 4 berikut.



Gambar 4. Proses *Import Data Excel Step 2*

Proses selanjutnya adalah memberi label pada kategori ISPA. Tujuan dari pemberian label adalah untuk menentukan kelas atau kategori dari data yang dianalisis. Dengan adanya label yang benar saat pelatihan model dapat berjalan lebih akurat. Tanpa label, model tidak tahu apakah prediksi yang dihasilkan benar atau tidak sehingga proses pembelajaran menjadi tidak efektif. Adapun label dalam data penelitian ini diklasifikasikan menjadi tiga yaitu ISPA kategori ringan, sedang, dan berat.



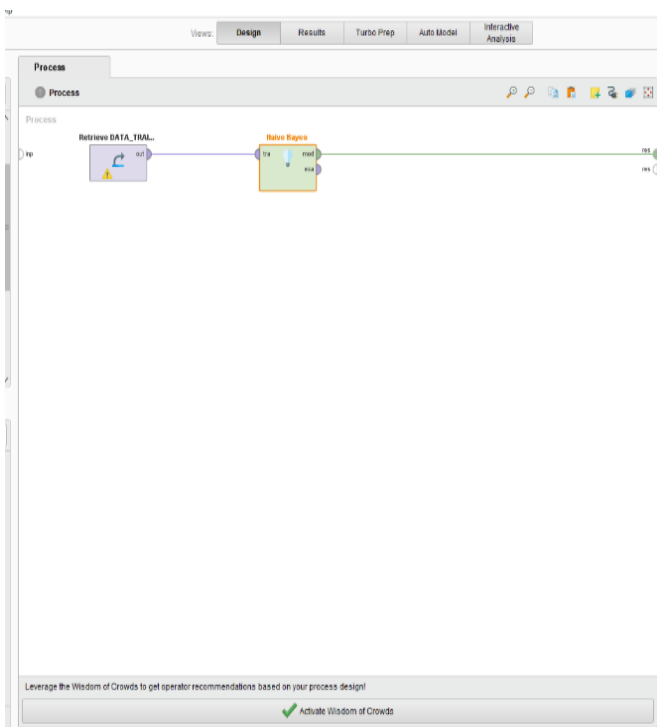
Gambar 5. Proses *Import Data Excel* Step 3

Tampilan gambar 6 berikut adalah atribut yang telah diberi label.

	Usia	Jenis Penyakit	Masa Idap	Kategori ISPA
1	0-11 bln	A01	B1	Ringan
2	0-11 bln	A01	B1	Ringan
3	0-11 bln	A01	B1	Sedang
4	0-11 bln	A01	B1	Ringan
5	0-11 bln	A01	B1	Ringan
6	0-11 bln	A01	B1	Sedang
7	0-11 bln	A01	B1	Ringan
8	0-11 bln	A01	B1	Ringan
9	1-4 thn	A05	B1	Berat
10	1-4 thn	A01	B2	Ringan
11	1-4 thn	A01	B1	Ringan
12	1-4 thn	A01	B1	Ringan
13	1-4 thn	A01	B1	Ringan
14	1-4 thn	A01	B1	Ringan
15	1-4 thn	A01	B1	Ringan
16	1-4 thn	A01	B1	Ringan
17	1-4 thn	A01	B1	Ringan
18	1-4 thn	A01	B1	Sedang
19	1-4 thn	A01	B1	Ringan

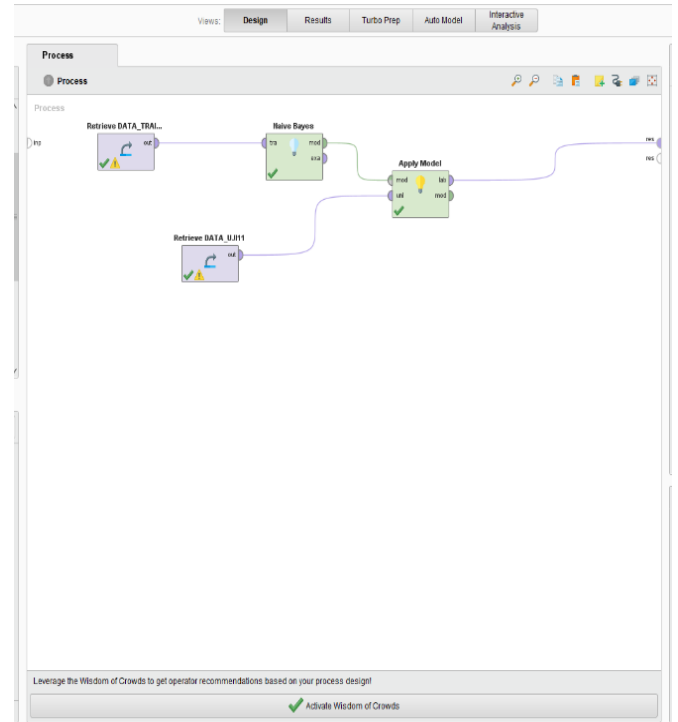
Gambar 6. Proses *Import Data Excel* Step 4

Setelah *import* data selesai, maka proses selanjutnya yaitu pemodelan. Data pelatihan dihubungkan ke operator *Naive Bayes* untuk dipelajari pola datanya sehingga akan menghasilkan model seperti yang terlihat pada gambar 7.



Gambar 7. Proses Pemodelan Step 1

Pada gambar berikut, terdapat model dari *Naive Bayes* dan data uji dihubungkan ke *Apply Model* lalu di *result* untuk melihat hasil prediksi dari *Naive Bayes* apakah sesuai dengan data uji.



Gambar 8. Proses Pemodelan Step 2

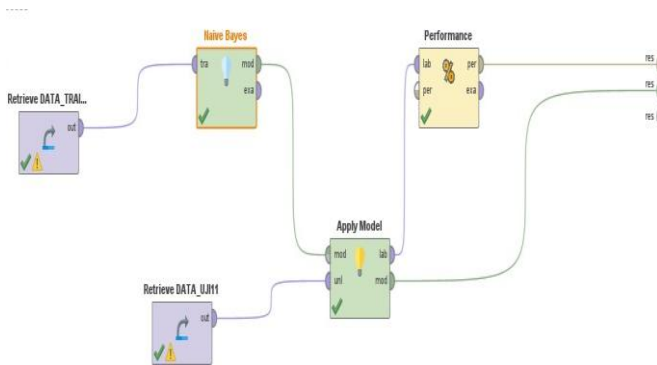
Gambar 9 berikut ini menunjukkan hasil dari pengujian data yang terhubung dengan model dan di hasilnya. Lebih banyak data yang sesuai akan menghasilkan klasifikasi yang lebih akurat.

Row No.	Kategori ISPA	predictionK...	confidence...	confidence...	confidence...	Usia	Jenis Pempa.	Masa idap
1	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
2	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
3	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
4	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
5	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
6	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
7	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
8	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
9	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
10	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
11	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
12	Ringan	Ringan	1.000	0.000	0.000	0-11 bln	A01	B1
13	Ringan	Ringan	0.000	0.020	0.000	1-4 thn	A01	B2
14	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
15	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
16	Ringan	Ringan	0.000	0.020	0.000	1-4 thn	A01	B2
17	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
18	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
19	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
20	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
21	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
22	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
23	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
24	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
25	Ringan	Ringan	1.000	0.000	0.000	1-4 thn	A01	B1
26	Sedang	Sedang	0.004	0.996	0.000	5-6 thn	A02	B2

Gambar 9. Proses Pemodelan Step 3

Data ISPA hasil klasifikasi dari data uji yang dihubungkan ke *apply* model selanjutnya dihubungkan ke

performance dan result untuk melihat tingkat akurasi yang dihasilkan. Untuk penjelasan yang lebih rinci, dapat dilihat pada gambar 10 berikut.



Gambar 10. Desain Model Rapidminer

Akurasi adalah ukuran seberapa sering model membuat prediksi yang benar dibandingkan dengan seluruh jumlah prediksi yang dibuat. Nilai akurasi sering digunakan untuk mengevaluasi performa model klasifikasi. Desain pemodelan dari rapidminer menghasilkan akurasi sebesar 96,61%. Semakin tinggi akurasi semakin sering model membuat keputusan yang benar. Dalam praktik medis akurasi tinggi bisa memberikan keyakinan awal kepada tenaga medis bahwa sistem mampu membantu proses deteksi dini penyakit ISPA.

Berikut adalah tampilan hasil desain model dari rapidminer.

accuracy: 96.61%

	true Ringan	true Sedang	true Berat	class precision
pred. Ringan	121	3	0	97.50%
pred. Sedang	5	106	0	95.50%
pred. Berat	0	0	1	100.00%
class recall	96.03%	97.25%	100.00%	

Gambar 11. Hasil Desain Model dari Rapidminer

Berdasarkan hasil, true berat 100 % dikarenakan data untuk ISPA berat itu sedikit sehingga model mempelajari data tersebut dengan sempurna sehingga menghasilkan nilai True (TP atau TN) 100 %.

#### F. Evaluasi Kinerja

Untuk mengukur evaluasi kinerja dari pengklasifikasian Bayes menggunakan alat ukur evaluasi berupa confusion matrix dengan tujuan untuk mempermudah dalam menganalisis Algoritma. Data yang diperoleh berjumlah 787 pasien. Berdasarkan hasil evaluasi 121 pasien diprediksi ringan dengan kategori ringan, 3 diprediksi ringan dengan kategori sedang. 5 pasien diprediksi sedang dengan kategori ringan, sedangkan 105 pasien diprediksi sedang dengan kategori sedang. 1 pasien diprediksi berat dengan kategori berat. Berdasarkan hasil evaluasi confusion matrix diperoleh hasil akurasi sebesar 96.61% dengan rincian sebagai berikut.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$(4)$$

$$Recall = \frac{TP}{TP+FN}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Keterangan:

- TP = True Positif
- FP = False Positif
- TN = True Negatif
- FN = False Negatif

*Precision*: Proposisi data sampel bernilai benar yang berhasil diprediksi secara tepat.

$$\frac{TP}{TP+FP} = \frac{121}{121+3} \times 100\% = 97,58\%$$

*Recall*: Proposisi data sampel bernilai benar yang diprediksi secara benar.

$$\frac{TP}{TP+FN} = \frac{121}{121+5} \times 100\% = 96,03\%$$

Hasil dari *precision* dan *recall* dapat dihitung nilai akurasinya:

$$\frac{\text{Banyaknya prediksi yang benar}}{\text{Total banyaknya prediksi}} \times 100\%$$

$$\frac{121+106+1}{121+3+5} \times 100\% = \frac{228}{236} = 96,61$$

Berikut merupakan hasil *Confusion Matrix* yang dituliskan dalam tabel 2.

TABEL 2. HASIL CONFUSION MATRIX

	True Ringan	True Sedang	True Berat	Class Precision
Pred. Ringan	121	3	0	97.58%
Pred. Sedang	5	106	0	95.50%
Pred. Berat	0	0	1	100.00%
Class Recall	96.03%	97.25%	100.00%	

#### IV. KESIMPULAN

Dari hasil confusion matrix terlihat bahwa akurasi dari metode *Naive Bayes Classifier* adalah 96,61 % dengan presisi 97,58 % untuk prediksi ringan, 95,50% untuk prediksi sedang dan 100% untuk prediksi berat. *Recall* 96,03% untuk true ringan, 97,25 % untuk true sedang dan 100 % untuk true berat. Hasil ini menunjukkan bahwa metode ini dapat digunakan sebagai alat bantu dalam sistem pendukung keputusan medis. Hal ini memberikan manfaat bagi tenaga medis dalam mempercepat proses diagnosis awal terhadap pasien ke dalam tiga jenis penyakit ISPA yaitu ringan, sedang, dan berat. Hasil dari penelitian ini menunjukkan bahwa *Naive Bayes Classifier* mampu menghasilkan prediksi yang tepat dalam mengidentifikasi kasus ISPA.

Untuk peneliti selanjutnya disarankan untuk menambahkan fitur medis yang lengkap seperti riwayat imunisasi dan hasil laboratorium. Hal ini akan meningkatkan akurasi dan kemampuan prediktif model dengan menggabungkan beberapa algoritma klasifikasi seperti *random forest* untuk memperoleh hasil yang lebih stabil dan

akurat. Diharapkan dengan adanya model klasifikasi ISPA dapat digunakan sebagai alat bantu dalam diagnosis yang cepat dan akurat sehingga mendukung tenaga medis dalam pengambilan keputusan klinis. Penelitian ini dapat menjadi dasar dalam pengembangan sistem pendukung keputusan berbasis kecerdasan buatan seperti sistem pakar.

### UCAPAN TERIMA KASIH

Terima kasih kepada Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat (DRTPM) sebagai pemberi dana hibah pada Program Penelitian Dosen Pemula (PDP) Afirmasi tahun 2024.

### REFERENSI

- [1] I. Silviana, "Hubungan Pengetahuan Ibu tentang Penyakit ISPA dengan Perilaku Pencegahan ISPA pada Balita di PHPT Muara Angke Jakarta Utara Tahun 2014," *J. Ilm.*, pp. 402–411, 2014.
- [2] M. Bari, S. H. Sitorus, and U. Ristian, "IMPLEMENTASI METODE NAÏVE BAYES PADA APLIKASI PREDIKSI PENYEBARAN WABAH PENYAKIT ISPA (Studi Kasus: Wilayah Kota Pontianak)," *J. Coding, Sist. Komput. Untan*, vol. 06, no. 03, pp. 205–214, 2018.
- [3] Iis Dewi Ratih, S M Retnaningsih, and V M Dewi, "Klasifikasi Kualitas Tanah Menggunakan Metode Naive Bayes Classifier," *J. Apl. Mat. dan Stat.*, vol. 1, no. 1, pp. 11–20, 2022, doi: 10.53625/jams.v1i1.4427.
- [4] I. V. Onibala and A. S. Purnomo, "Sistem Pakar untuk Diagnosa Penyakit Infeksi Saluran Pernapasan Akut ( ISPA ) Menggunakan Metode Naive Bayes," vol. 4, pp. 14666–14685, 2024.
- [5] N. L. W. S. R. Ginantra, I. G. A. D. Indradewi, and E. Hartono, "Machine learning approach for Acute Respiratory Infections (ISPA) prediction: Case study Indonesia," *J. Phys. Conf. Ser.*, vol. 1469, no. 1, 2020, doi: 10.1088/1742-6596/1469/1/012044.
- [6] R. Pradana, "Penerapan Konsep Non-deterministic Finite Automata Dalam Diagnosa Penyakit ISPA," *TeKa*, vol. 13, no. 02, pp. 121–130, 2023, doi: 10.36342/teika.v13i02.3143.
- [7] D. Syifani and A. Dores, "Aplikasi Sistem Rekam Medis Di Puskesmas Kelurahan Gunung," *Teknol. Inform. dan Komput.*, vol. 9, no. 1, 2018.
- [8] W. Hidayatullah, "Sistem Pakar Diagnosis Penyakit Ispa Menggunakan Metode Naive Expert System Diagnosis of Ari Disease Using Naive Bayes Method Based on Web Based Puskesmas Teratak," vol. 2, no. 1, pp. 32–42, 2023.
- [9] E. D. Sikumbang, "Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *J. Tek. Komput. AMIK BSI*, vol. Vol 4, No., no. September, pp. 1–4, 2018.
- [10] A. Hafiyyan and E. Seniwati, "Implementation of The Naive Bayes Algorithm For Diagnosing Lung Disesase In An Expert System Concept," *Int. J. Inf. Syst. Technol. Akreditasi*, vol. 7, no. 1, pp. 23–29, 2023.
- [11] B. Betrisandi, R. Sulaehani, and I. C. R. Drajana, "Klasifikasi Nasabah Dalam Pengelolaan Resiko Kredit Menggunakan Metode Naive Bayes," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 6, no. 6, pp. 804–808, 2024, doi: 10.32672/jnkti.v6i6.7246.
- [12] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [13] A. Y. Simanjuntak, I. S. S. Simatupang, and Anita, "Implementasi Data Mining Menggunakan Metode Naive Bayes Classifier Untuk Data Kenaikan Pangkat Dinas," *J. Sci. Soc. Res.*, vol. 4307, no. 1, pp. 85–91, 2022.
- [14] N. T. Rahman, F. I. Komputer, U. Darwan, and A. Sampit, "Analisa algoritma," vol. 10, no. 2, pp. 144–151, 2020.
- [15] M. Jannah, "Perbandingan Metode Naive Bayes Dan K-Nearest Neighbor Dalam Mengklasifikasi Status Pertumbuhan Anak Stunting (Studi Kasus: Posyandu Cemara)," vol. 14, no. 1, pp. 250–255, 2022, [Online]. Available: [https://repository.unilak.ac.id/eprint/2807%0Ahttps://repository.unilak.ac.id/2807/1/1855201057\\_BAB-I\\_VI\\_DAFTAR-PUSTAKA.pdf](https://repository.unilak.ac.id/eprint/2807%0Ahttps://repository.unilak.ac.id/2807/1/1855201057_BAB-I_VI_DAFTAR-PUSTAKA.pdf).