

# KLASIFIKASI NAÏVE BAYES UNTUK MENDIAGNOSIS PENYAKIT PNEUMONIA PADA ANAK BALITA (STUDI KASUS : UPTD PUSKESMAS SUKARAJA SUKABUMI)

Ami Rahmawati<sup>1</sup>, Dede Wintana<sup>2</sup>, Satia Suhada<sup>3</sup>, Gunawan<sup>4</sup>, Hamdun Sulaiman<sup>5</sup>

Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri  
Universitas Bina Sarana Informatika  
Jl. Damai No. 8, Warung Jati Barat (Margasatwa), Pasar Minggu, RT.1/RW.1,  
Ragunan, Pasar Minggu, Jakarta  
Jl. Kamal Raya No.18 Ringroad Barat, Cengkareng, Jakarta Barat  
ami.amv@bsi.ac.id<sup>1</sup>, dede.dwe@bsi.ac.id<sup>2</sup>, satia.shq@bsi.ac.id<sup>3</sup>,  
gunawan.gnz@bsi.ac.id<sup>4</sup>, hamdun.hsl@bsi.ac.id<sup>5</sup>

## Abstract

*Pneumonia is a contagious infectious disease that is the leading cause of death in toddlers in the world. In developed countries, there are 4 million cases each year, totaling 156 million cases of pneumonia every year worldwide. Pneumonia is caused by, among others, bacteria, viruses, fungi, exposure to chemicals or physical damage from the lungs, as well as indirect effects from other diseases. Pneumonia is characterized by symptoms of coughing and / or difficulty breathing such as rapid breathing, and pulling the lower chest wall inward. Therefore, early detection of pneumonia in children under five is very necessary in order to be able to prevent and cope with the disease into a serious stage as the purpose of this study is to diagnose pneumonia in toddlers using data mining classification, the naïve Bayes algorithm. Of the 118 cases consisting of 113 cases of patients diagnosed with pneumonia and 5 cases of patients who were not diagnosed with pneumonia, an accuracy value of 98% was obtained, so it can be interpreted that the naïve bayes algorithm has a good correlation with the attributes contained in the dataset.*

**Keywords:** Naïve Bayes Algorithm, Pneumonia.

## Abstrak

*Pneumonia adalah penyakit infeksi menular yang merupakan penyebab utama kematian pada balita di dunia. Di negara maju terdapat 4 juta kasus setiap tahun hingga total di seluruh dunia ada 156 juta kasus pneumonia anak balita setiap tahun. Pneumonia antara lain disebabkan oleh bakteri, virus, jamur, paparan bahan kimia atau kerusakan fisik dari paru-paru, maupun pengaruh tidak langsung dari penyakit lain. Pneumonia ditandai dengan gejala batuk dan atau kesulitan bernapas seperti napas cepat, dan tarikan dinding dada bagian bawah ke dalam. Oleh Karena itu, deteksi dini penyakit pneumonia pada anak balita sangat diperlukan agar dapat mencegah dan menanggulangi penyakit tersebut kedalam tahap yang serius seperti tujuan penelitian ini yaitu untuk mendiagnosis penyakit pneumonia pada anak balita menggunakan klasifikasi data mining yaitu algoritma naïve*

bayes. Dari 118 kasus yang terdiri dari 113 kasus pasien yang terdiagnosis pneumonia dan 5 kasus pasien yang tidak terdiagnosis pneumonia maka diperoleh nilai akurasi sebesar 98%, sehingga dapat diartikan bahwa algoritma naïve bayes memiliki korelasi yang baik dengan atribut yang terdapat pada dataset.

**Kata kunci :** Algoritma Naïve Bayes, Penyakit Pneumonia.

## 1. PENDAHULUAN

Pneumonia adalah penyakit infeksi menular yang merupakan penyebab utama kematian pada balita di dunia. *World Health Organization (WHO)* memperkirakan insidens pneumonia anak-balita di negara berkembang adalah 0,29 episode per anak-tahun atau 151,8 juta kasus pneumonia per tahun, 8,7% (13,1 juta) di antaranya merupakan pneumonia berat dan perlu rawat-inap. Di negara maju terdapat 4 juta kasus setiap tahun hingga total di seluruh dunia ada 156 juta kasus pneumonia anak balita setiap tahun [1].

Menurut data Riskesdas 2007, *prevalens* pneumonia (berdasarkan pengakuan pernah didiagnosis pneumonia oleh tenaga kesehatan dalam sebulan terakhir sebelum survei) pada bayi di Indonesia adalah 0,76% dengan rentang antar provinsi sebesar 0-13,2%. Prevalensi tertinggi adalah provinsi Gorontalo (13,2%) dan Bali (12,9%), sedangkan provinsi lainnya di bawah 10%, Sedangkan prevalensi pada anak balita (1-4 tahun) adalah 1,00% dengan rentang antar provinsi sebesar 0,1% - 14,8%. Seperti pada bayi, prevalensi tertinggi adalah provinsi Gorontalo (19,9%) dan Bali (13,2%) sedangkan provinsi lainnya di bawah 10% [1].

Pneumonia adalah bentuk parah dari infeksi saluran pernapasan akut bagian bawah yang secara khusus mempengaruhi paru, keadaan dimana alveoli pada salah satu atau kedua paru-paru terisi oleh cairan yang menyebabkan terganggunya pertukaran oksigen yang membuat sulit untuk bernapas [2].

Pneumonia antara lain disebabkan oleh bakteri, virus, jamur, pajanan bahan kimia atau kerusakan fisik dari paru-paru, maupun pengaruh tidak langsung dari penyakit lain. Bakteri yang biasa menyebabkan pneumonia adalah *Streptococcus* dan *Mycoplasma pneumoniae*, sedangkan virus yang menyebabkan pneumonia adalah *adenoviruses*, *rhinovirus*, *influenza virus*, *respiratory syncytialvirus (RSV)* dan para *influenza virus*, Terjadinya pneumonia ditandai dengan gejala batuk dan atau kesulitan bernapas seperti napas cepat, dan tarikan dinding dada bagian bawah ke dalam [3]. Oleh Karena itu, deteksi dini penyakit pneumonia sangat diperlukan, agar dapat mencegah dan menanggulangi penyakit tersebut kedalam tahap yang serius.

Berdasarkan uraian tersebut, penelitian ini dimaksudkan untuk meningkatkan akurasi prediksi dalam mendiagnosis penyakit pneumonia dengan menerapkan metode klasifikasi naïve bayes sebagai upaya pendeteksian dini terhadap penyakit tersebut.

## **2. METODOLOGI PENELITIAN**

### **2.1. Penyakit Pneumonia**

Pneumonia adalah keadaan dimana alveoli pada salah satu atau kedua paru-paru terisi oleh cairan yang menyebabkan terganggunya pertukaran oksigen yang membuat sulit untuk bernapas. Infeksi dapat disebabkan oleh bakteri, virus maupun jamur. Pneumonia juga dapat terjadi akibat kecelakaan karena menghirup cairan atau bahan kimia [2].

### **2.2. Data Mining**

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [4].

### **2.3. Klasifikasi**

Klasifikasi merupakan proses pembelajaran suatu fungsi tujuan (target) yang memetakan tiap himpunan atribut  $x$  ke satu dari label kelas  $y$  yang didefinisikan sebelumnya [5].

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Proses classification biasanya dibagi menjadi dua fase: learning dan test. Pada fase learning, sebagian data yang telah diketahui kelas datanya diumpungkan untuk membentuk model perkiraan. Kemudian pada fase test model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut. Bila akurasinya mencukupi model ini dapat dipakai untuk prediksi kelas data yang belum diketahui [6].

### **2.4. Algoritma Naïve Bayes**

Algoritma Naïve Bayes adalah algoritma yang mampu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya dengan menggunakan metode probabilitas dan statistic [7].

Algoritma Naïve Bayes adalah metode yang membagi permasalahan dalam sebuah kelas-kelas berdasarkan ciri-ciri persamaan dan perbedaan dengan menggunakan statistic yang bisa memprediksi probabilitas sebuah kelas. Langkah-langkah perhitungan dengan menggunakan naïve bayes yaitu [8] :

- a. Menentukan kriteria apa saja yang digunakan, kriteria sangat penting dalam melakukan perhitungan naïve bayes, karena digunakan untuk mengumpulkan data.
- b. Menyiapkan data yang digunakan untuk melakukan perhitungan naïve bayes.
- c. Membuat tabel aturan, hanya dengan batasan aturan dan data pada tabel aturan diambil dari data yang ada.
- d. Membuat tabel probabilitas kemunculan setiap atribut dari semua kriteria yang ada.

- e. Menghitung nilai *likelihood* ya dan *likelihood* tidak yang diambil dari tabel probabilitas kemunculan setiap nilai atribut.
- f. Menghitung nilai probabilitas dimana dapat dihitung dengan melakukan normalisasi terhadap *likelihood* dan dengan menghitung nilai probabilitas ini bias mengetahui hasil akhir dari perhitungan dengan menggunakan metode naïve bayes layak atau tidak layak berdasarkan dari probabilitas ya atau tidak.

Rumus perhitungan yang digunakan pada algoritma Naïve Bayes adalah sebagai berikut [7] :

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)}$$

Di mana:

Y : Data dengan class yang belum diketahui

X : Hipotesis data merupakan suatu class spesifik

P(X|Y) : Probabilitas hipotesis X berdasar kondisi Y (posteriori probabilitas)

P(X) : Probabilitas hipotesis X (prior probabilitas)

P(Y|X) : Probabilitas Y berdasarkan kondisi pada hipotesis X

P(Y) : Probabilitas Y

Untuk mengetahui kelas apa yang cocok bagi parameter yang akan dianalisis untuk melakukan proses klasifikasi maka penyesuaian dilakukan terhadap Algoritma Naïve Bayes seperti:

$$P(C|F1...Fn) = \frac{P(C|F1...Fn)}{P(C|F1...Fn)}$$

Variabel C pada persamaan diatas dapat digunakan untuk mewakili kelas, sedangkan variabel F1 ... Fn digunakan untuk mewakili karakteristik petunjuk yang merupakan bagian penting guna melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*).

$$Posterior = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

Naïve bayes mempunyai kelebihan dan kekurangan. Kelebihan dan kekurangannya yaitu sebagai berikut [9] :

Kelebihan Naive Bayes:

- a. Menangani kuantitatif dan data diskrit.

- b. Mengeksplor batasan-batasan keoptimalan algoritma klasifikasi Naive Bayes dan mencoba menjelaskan mengenai alasan mengapa algoritma klasifikasi Naive Bayes berfungsi dengan baik pada berbagai jenis data.
- c. Manfaat lain dari algoritma Naive Bayes adalah metode ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter (rata-rata dan variansi dari variabel) yang diperlukan dalam proses pengklasifikasian.

Kekurangan Naive Bayes:

- a. Metode Naive Bayes hanya bisa digunakan untuk persoalan klasifikasi dengan supervised learning dan data-data kategorikal.
- b. Metode Naive Bayes memerlukan pengetahuan awal untuk dapat mengambil suatu keputusan. Tingkat keberhasilan metode ini sangat tergantung pada pengetahuan awal yang diberikan.
- c. Tidak berlaku jika probabilitas kondisionalnya adalah nol. Apabila nol maka probabilitas prediksi akan bernilai nol juga.

## 2.5. *Confusion Matrix (Accuracy)*

Confusion matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Rumus ini melakukan perhitungan dengan 3 keluaran, yaitu : *recall*, *precision*, dan *accuracy*, ditunjukkan pada tabel 1 [9].

Tabel 1. Model *Confusion matrix*

<b>Class</b>	<b>Class Positif</b>	<b>Class Negatif</b>
<b>Prediksi Positif</b>	TP	FN
<b>Prediksi Negatif</b>	FP	TN

Keterangan :

- a. Jika hasil prediksi negatif dan data sebenarnya negatif (FN).
- b. Jika hasil prediksi positif sedangkan nilai sebenarnya negatif (TN).
- c. Jika hasil prediksi negatif sedangkan nilai sebenarnya positif (FP).
- d. Jika hasil prediksi positif dan nilai sebenarnya positif (TP).

Data yang digunakan pada penelitian ini menggunakan data primer yang diperoleh dari UPTD Puskesmas Sukaraja Sukabumi, dengan jumlah 128 record yang terdiri dari 113 record yang menderita pneumonia dan 15 record yang tidak menderita pneumonia.

Dalam penelitian ini digunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang terdiri dari enam tahap, yaitu *business*

*understanding, data understanding, data preparation, modeling, evaluation, dan deployment.*

a. Tahap *Business Understanding*

Berdasarkan hasil pemeriksaan pasien penyakit pneumonia pada balita dengan jumlah 128 orang di UPTD Puskesmas Sukaraja Sukabumi terdeteksi 113 orang positif pneumonia, untuk mengurangi jumlah resiko penyakit pneumonia positif maka harus dilakukan pemeriksaan terhadap pasien agar dapat diketahui penyakit pneumonia secara dini. Pada penelitian ini, untuk dapat menghasilkan tingkat prediksi dan nilai akurasi yang tinggi dilakukan Klasifikasi Naïve Bayes Untuk Mendiagnosis Penyakit Pneumonia Pada Anak Balita.

b. Tahap *Data Understanding*

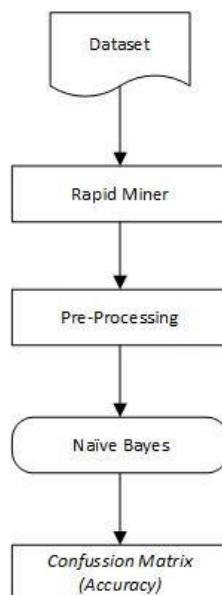
Data yang digunakan adalah data primer yang diperoleh dari UPTD Puskesmas Sukaraja Sukabumi. Data tersebut terdiri dari 7 atribut *predictor* dan 1 atribut hasil yang dapat diketahui status pasien yang positif ataupun negatif pneumonia.

c. Tahap *Data Preparation*

Pada tahap ini dilakukan proses pembersihan data (*data cleaning*). Pekerjaan yang dilakukan pada proses ini meliputi; mencoba melakukan pengisian *missing value*, melancarkan *noise*, mengidentifikasi *outlier* dan membenarkan *inconsistencies* dalam dataset menggunakan tools RapidMiner.

Akhir dari proses tersebut diperoleh hasil preprocessing data yang didapat dari dataset Pneumonia UPTD Puskesmas Sukaraja Sukabumi sebanyak 128 record direduksi menjadi 118 record yang terdiri dari 113 record positif Pneumonia dan 5 negatif Pneumonia.

d. Tahap *Modelling*



Gambar 1. Model yang Diusulkan

Pada penelitian ini data yang digunakan yaitu data training yang telah disiapkan.

e. Tahap *Evaluation*

Tahap ini melakukan evaluasi terhadap model-model yang telah terbentuk dengan menggunakan metode *Confusion Matrix* sehingga mendapatkan informasi model yang akurat.

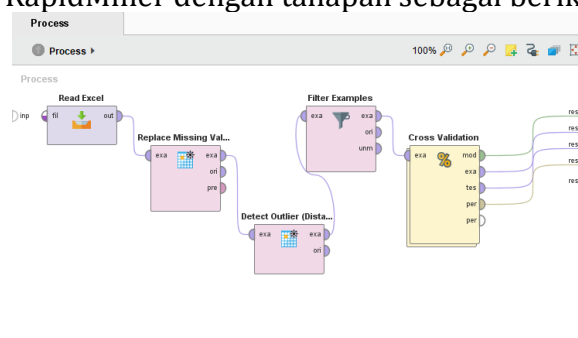
f. Tahap *Devoyment*

Dari model yang telah dihasilkan maka perlu diuji dengan menggunakan data baru dan dilakukan kembali evaluasi untuk keakuratan data.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Dataset

Eksperimen dan pengujian dalam penelitian dilakukan menggunakan prediksi naïve bayes. Eksperimen ini dilakukan terhadap dataset yang telah divalidasi berdasarkan hasil data *preprocessing* yang telah dilakukan menggunakan *tools* RapidMiner dengan tahapan sebagai berikut:



Gambar 2. Tahapan Validasi Data Menggunakan *Tools* RapidMiner

#### 3.2. Hasil Pemodelan dengan Algoritma Naïve Bayes

Untuk menghitung *Accuracy* penyakit pneumonia pada anak balita dengan menggunakan algoritma naïve bayes maka langkah pertama yang dilakukan adalah membaca data training penyakit pneumonia. Adapun data yang digunakan sebagai berikut:

Tabel 2. Data Pasien Pneumonia Puskesmas Sukaraja

Data	Jenis Kelamin	Usia	Demam	Batuk	Nafas Cepat	Sesak Napas	Sakit Kepala	Hasil
1	Laki-Laki	2 tahun s/d kurang dari 3 tahun	Tidak	Ya	Ya	Ya	Tidak	Tidak
2	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Tidak	Ya	Ya	Ya	Ya	Ya

Data	Jenis Kelamin	Usia	Demam	Batuk	Nafas Cepat	Sesak Napas	Sakit Kepala	Hasil
3	Perempuan	3 tahun s/d kurang dari 4 tahun	Ya	Ya	Ya	Ya	Ya	Ya
4	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Ya	Ya	Ya	Ya	Ya	Ya
5	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
6	Laki-Laki	6 bulan s/d kurang dari 9 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
7	Laki-Laki	6 bulan s/d kurang dari 9 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
8	Perempuan	3 tahun s/d kurang dari 4 tahun	Ya	Ya	Ya	Tidak	Ya	Ya
9	Perempuan	2 tahun s/d kurang dari 3 tahun	Ya	Ya	Ya	Ya	Ya	Ya
10	Perempuan	2 tahun s/d kurang dari 3 tahun	Tidak	Ya	Ya	Tidak	Ya	Ya
11	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Ya	Ya	Ya	Ya	Ya	Ya
12	Laki-Laki	3 bulan s/d kurang dari 6 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
13	Perempuan	2 tahun s/d kurang dari 3	Tidak	Ya	Ya	Ya	Ya	Ya



Data	Jenis Kelamin	Usia	Demam	Batuk	Nafas Cepat	Sesak Napas	Sakit Kepala	Hasil
		tahun						
14	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Ya	Ya	Ya	Ya	Ya	Ya
15	Perempuan	6 bulan s/d kurang dari 9 bulan	Tidak	Ya	Ya	Ya	Ya	Ya
16	Perempuan	9 bulan s/d kurang dari 12 bulan	Tidak	Ya	Ya	Ya	Ya	Ya
17	Perempuan	18 bulan s/d kurang dari 24 bulan	Tidak	Ya	Ya	Tidak	Ya	Ya
18	Laki-Laki	18 bulan s/d kurang dari 24 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
19	Laki-Laki	18 bulan s/d kurang dari 24 bulan	Ya	Ya	Ya	Ya	Ya	Ya
20	Laki-Laki	2 tahun s/d kurang dari 3 tahun	Tidak	Ya	Ya	Tidak	Ya	Ya
21	Perempuan	3 tahun s/d kurang dari 4 tahun	Ya	Ya	Ya	Ya	Ya	Ya
22	Perempuan	2 tahun s/d kurang dari 3 tahun	Ya	Ya	Ya	Ya	Ya	Ya
23	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Tidak	Ya	Ya	Ya	Ya	Ya

Data	Jenis Kelamin	Usia	Demam	Batuk	Nafas Cepat	Sesak Napas	Sakit Kepala	Hasil
24	Perempuan	2 tahun s/d kurang dari 3 tahun	Ya	Ya	Ya	Tidak	Ya	Ya
25	Laki-Laki	2 tahun s/d kurang dari 3 tahun	Tidak	Ya	Ya	Ya	Ya	Ya
26	Laki-Laki	18 bulan s/d kurang dari 24 bulan	Ya	Ya	Ya	Ya	Ya	Ya
27	Laki-Laki	18 bulan s/d kurang dari 24 bulan	Tidak	Ya	Ya	Tidak	Ya	Ya
28	Perempuan	6 bulan s/d kurang dari 9 bulan	Ya	Ya	Ya	Ya	Ya	Ya
29	Laki-Laki	6 bulan s/d kurang dari 9 bulan	Ya	Ya	Ya	Tidak	Ya	Ya
30	Laki-Laki	9 bulan s/d kurang dari 12 bulan	Ya	Ya	Ya	Tidak	Ya	Ya

Data pasien pneumonia pada anak balita di puskesmas sukaraja hanya ditampilkan 42 data seperti di atas.

Tabel 3. Probabilitas Kelas

Kelas			
	Ya	Tidak	
	Ya	113	Tidak 5
<b>P(Ya) =</b>	<b>0,96</b>	<b>P(Tidak) =</b>	<b>0,04</b>
<b>113/118</b>		<b>5/118</b>	

Tabel diatas merupakan nilai probabilitas untuk setiap kelas berdasarkan data pasien pneumonia di puskesmas sukaraja. Mencari probabilitas hipotesis untuk masing-masing kelas  $P(H)$  dilakukan dalam membuat sebuah model naïve

bayes. Hipotesis yang ada yaitu pasien balita yang menderita penyakit pneumonia dan pasien balita yang tidak menderita penyakit pneumonia. Perhitungan probabilitas yaitu seperti dibawah ini:

$$P(\text{Ya}) = \frac{113}{118} = 0,96$$

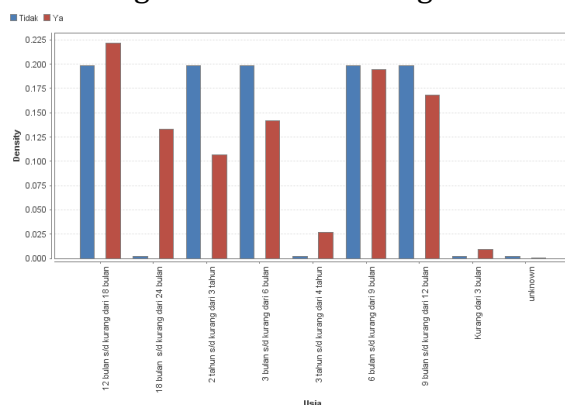
$$P(\text{Tidak}) = \frac{5}{118} = 0,04$$

Selanjutnya menghitung probabilitas kondisi tertentu (probabilitas X) berdasarkan probabilitas hipotesis (probabilitas H) atau dinamakan *probabilitas prior*. Berdasarkan data penyakit pneumonia pada anak balita di tabel 2 maka dihitung Probabilitas Prior yang dapat dilihat pada tabel 4.

Tabel 4. Probabilitas Prior

Atribut	Jumlah Kasus			P(X   Ci)	
	Total (S)	Ya (Si)	Tidak (Si)	Ya	Tidak
<b>Total</b>	118	113	5	0.957627118	0.0423
<b>Jenis Kelamin</b>					
Laki-laki	56	54	2	0.477876106	0.4
Perempuan	62	59	3	0.522123894	0.6
<b>Usia</b>					
Kurang dari 3 bulan	1	1	0	0.008849558	0
3 bulan s/d kurang dari 6 bulan	17	16	1	0.14159292	0.2
6 bulan s/d kurang dari 9 bulan	23	22	1	0.194690265	0.2
9 bulan s/d kurang dari 12 bulan	20	19	1	0.168141593	0.2
12 bulan s/d kurang dari 18 bulan	26	25	1	0.221238938	0.2
18 bulan s/d kurang dari 24 bulan	15	15	0	0.132743363	0
2 tahun s/d kurang dari 3 tahun	13	12	1	0.10619469	0.2
3 tahun s/d kurang dari 4 tahun	3	3	0	0.026548673	0
<b>Demam</b>					
Ya	63	63	0	0.557522124	0
Tidak	55	50	5	0.442477876	1
<b>Batuk</b>					
Ya	115	113	2	1	0.4
Tidak	3	0	3	0	0.6
<b>Nafas Cepat</b>					
Ya	115	113	2	1	0.4
Tidak	3	0	3	0	0.6
<b>Sesak Nafas</b>					
Ya	67	65	2	0.575221239	5
Tidak	51	48	3	0.424778761	0.6
<b>Sakit Kepala</b>					
Ya	113	113	0	1	0
Tidak	5	0	5	0	1

Model yang dihasilkan dari pengolahan data penyakit pneumonia pada anak balita dengan algoritma naïve bayes ditunjukkan oleh *distribution table* dengan atribut hasil *class* Ya sebesar 0,96 dan *class* Tidak 0,04. Model lain juga dapat dilihat di grafik distribusi dengan atribut usia sebagai berikut:



Gambar 3. Grafik Distribusi dengan Atribut Usia

Selanjutnya proses pengolahan data dengan algoritma naïve bayes pada penelitian ini, melakukan evaluasi efektifitas algoritma terhadap proses ekstraksi data berdasarkan tingkat akurasi dengan *Confusion Matrix*. Dari dataset (data pasien pneumonia di puskesmas sukaraja) digunakan 118 tupel dengan rincian jumlah *True Positive (TP)* 113, *False Positive (FP)* 2, *True Negative (TN)* 0 dan *False Negative (FN)* 3. Maka dihitung nilai *accuracy* sebagai berikut:

$$Accuracy = \frac{113+3}{113+2+0+3} = 0,983$$

#### 4. SIMPULAN

Dari hasil penelitian yang diperoleh, akurasi dari penggunaan algoritma naïve bayes untuk mendiagnosis penyakit pneumonia pada anak balita sebesar 98%, sehingga dengan nilai tersebut dapat diartikan bahwa algoritma naïve bayes memiliki korelasi yang baik dengan atribut yang terdapat pada dataset.

#### DAFTAR PUSTAKA

- [1] Departemen Kesehatan, "Pneumonia Balita," *www.depkes.go.id*, 2014. [Online]. Available: <http://www.depkes.go.id/article/view/13010200020/pneumonia-balita.html>.
- [2] M. S. Aulina, M. Rahardjo, and Nurjazuli, "POLA SEBARAN KEJADIAN PENYAKIT PNEUMONIA PADA BALITA DI KECAMATAN BERGAS, KABUPATEN SEMARANG," vol. 5, 2017.
- [3] A. Anwar, I. Dharmayanti, P. Teknologi, I. Kesehatan, M. Badan, and P. Kesehatan, "Pneumonia pada Anak Balita di Indonesia Pneumonia among Children Under Five Years of Age in Indonesia," no. 29, pp. 359–365, 2013.

- [4] D. Nofriansyah, *Konsep Data Mining VS Sistem Pendukung Keputusan*. Yogyakarta: Depublisher, 2014.
- [5] R. Wajhillah, "OPTIMASI ALGORITMA KLASIFIKASI C4 . 5 BERBASIS PARTICLE SWARM," no. September, 2014.
- [6] Aradea, S. A, A. Z, and Y. A, "Penerapan Decision Tree untuk Penentuan Pola Data Penerimaan Mahasiswa Baru," no. November, 2014.
- [7] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," vol. 2, no. 3, pp. 207–217, 2015.
- [8] H. Pratiwi, *Buku Ajar Sistem Pendukung Keputusan*. Yogyakarta: Deepublish, 2016.
- [9] T. Setiadi, "Penerapan Klasifikasi Bayes Untuk Memprediksi Jenis Latihan Siswa Pencak Silat ( Studi Kasus Pencak Silat PSHT )," vol. 7, pp. 60–65, 2018.