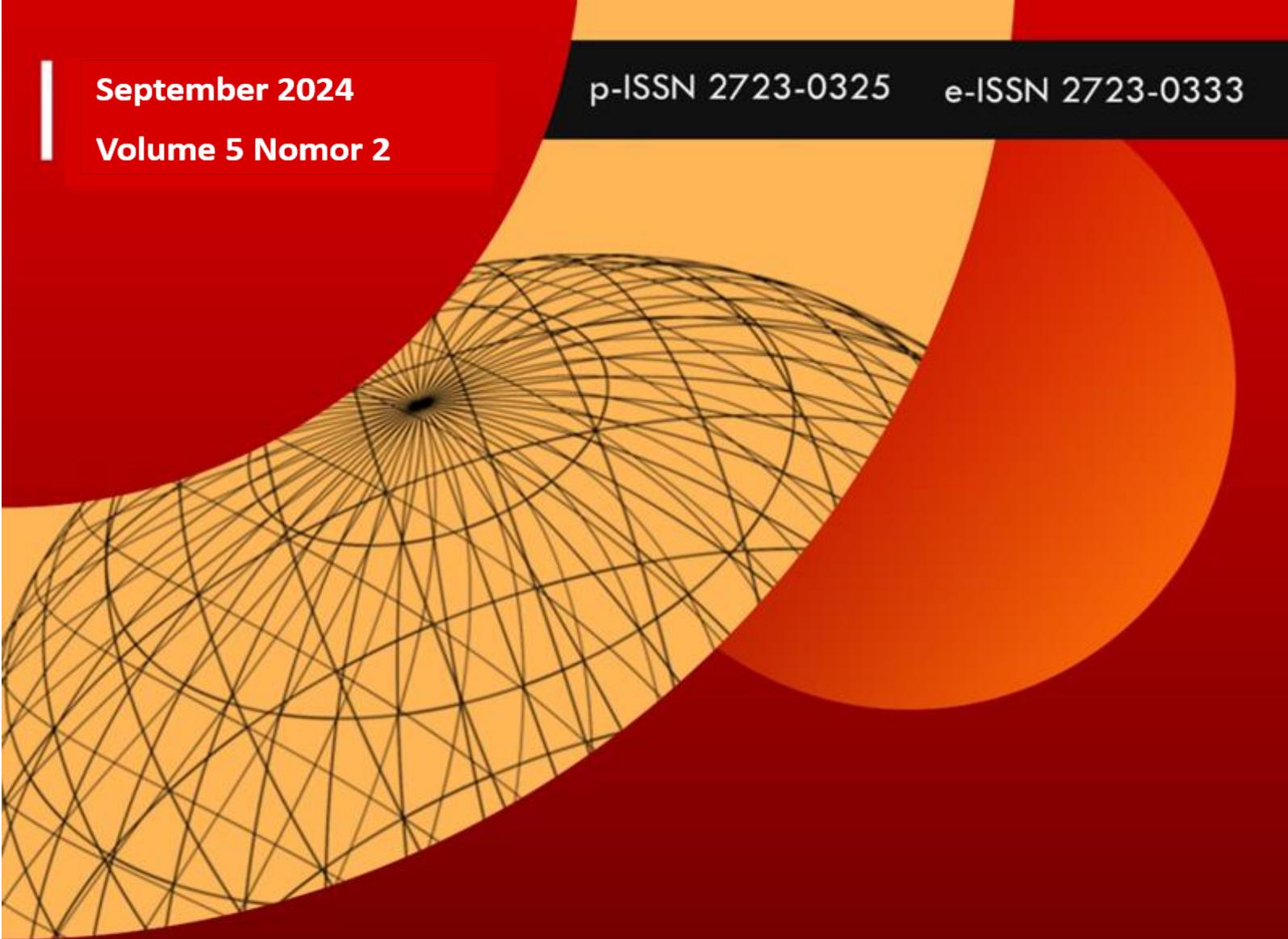


September 2024

Volume 5 Nomor 2

p-ISSN 2723-0325

e-ISSN 2723-0333



# TENSOR

Pure and Applied Mathematics Journal

**PROGRAM STUDI MATEMATIKA**

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS PATTIMURA

# TENSOR

Pure and Applied Mathematics Journal

is an international academic open-access journal that gains a foothold in mathematics, and its applications are issued twice a year. The focus is to publish original research and review articles on all aspects of pure and applied Mathematics. Editorial board members of the Journal and reviewers will review submitted papers. All submitted articles should report original, previously unpublished research results, experimental or theoretical, and will be peer-reviewed. Articles submitted to the journal should meet these criteria and must not be under consideration for publication elsewhere. Manuscripts should follow the journal template and are subject to both review and editing.

---

**Published by:**

**Department of Mathematics,  
Faculty of Mathematics and Natural Sciences,  
Pattimura University.  
Ambon  
2024**

**Copyright© Program Studi Matematika FMIPA UNPATTI 2024**

# TENSOR

Pure and Applied Mathematics Journal

Volume 5 Number 2 | September 2024

## Person In Charge

Head of Undergraduate Program in Mathematics,  
Faculty of Mathematics and Natural Sciences, Pattimura University

## Editor in Chief

Dr. H. Batkunde, S.Si, M.Si

## Editors

M. I. Tilukay, S.Si, M.Si (Managing and Section Editor)  
L. Bakarbessy, S.Si, M.Si (Managing and Section Editor)  
Z. A. Leleury, S.Si., M.Si (Copy and Production Editor)  
B. P. Tomasouw, S.Si, M.Si (Copy and Production Editor)  
Dr. L. K. Beay, S.Pd., M.Si (Proofreader)  
N. Dahoklory (Proofreader)

## Secretariat and Financial Officer

M. E. Rijoly, S.Si, M.Sc

## Graphic Design

V. Y. I. Ilwaru, S.Si, M.Si

## Expert Editorial Boards

Prof. Dr. Basuki Widodo, M.Sc (Institut Teknologi Sepuluh November Surabaya, Indonesia)  
Prof. Dr. M. Salman A. N, M.Si (Institut Teknologi Bandung, Indonesia)  
Dr. H. J. Wattimanela, S.Si., M.Si (Universitas Pattimura, Indonesia)  
Dr. Al Azhary Masta, S.Si., M.Si (Universitas Pendidikan Indonesia, Indonesia)  
Dr. Muh. Nur, S.Si., M.Si (Universitas Hasanudin, Indonesia)  
Dr. Meta Kallista, S.Si., M.Si (Universitas Telkom, Indonesia)  
Dr. Teguh Herlambang, S.Si., M.Si (Universitas Nahdlatul Ulama Surabaya, Indonesia)  
Asst. Prof. Dr. Anurak Thanyacharoen (Muban Chombueng Rajabhat University, Ratchaburi, Thailand)

## Publisher

Department of Mathematics,  
Faculty of Mathematics and Natural Sciences,  
Pattimura University, Ambon, Indonesia

## Editorial Address

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Pattimura  
Jln. Ir. M. Putuhena, Kampus Unpatti, Poka - Ambon 97233, Provinsi Maluku, Indonesia  
Contact : +62 82397854220  
Email : [tensormathematics@gmail.com](mailto:tensormathematics@gmail.com)

|   |   |         |
|---|---|---------|
| Mapping of The Transportation Sector in Maluku Province Using Biplot Analysis   | Zeth A. Leleury<br>Jefri E. T. Radjabaycolle<br>Venn Y. I. Ilwaru<br>Lexy J. Sinay                                    | 57-66   |
| Prediction of Divorce Data in Pamekasan District Based on Comparison of Exponential Smoothing and Moving Average                                  | Ira Yudistira<br>Siti Romlah<br>Tony Yulianto<br>Faisol<br>M.Fariz Fadillah M   | 67-78   |
| Exploring the Lazy Witness Complex for Efficient Persistent Homology in Large-Scale Data  | Mst Zinia Afroz Liza<br>Md. Al-Imran<br>Md. Morshed Bin Shiraj<br>Tozam Hossain<br>Md. Masum Murshed<br>Nasima Akhter | 79-92   |
| Penyelesaian <i>Unit Commitment Problem</i> (UCP) Menggunakan Algoritma Genetika  | Aisyah Fadhilah Whardhana<br>Asri Bekti Pratiwi<br>Edi Winarko  | 93-104  |
| The Total Disjoint Irregularity Strength of Double and Triple Star Graphs   | Tasya I. Titawanno<br>Meilin I. Tilukay<br>Zeth A. Leleury<br>Pranaya D. M. Taihuttu<br>Luvita Loves                  | 105-110 |
| An Application of the Naïve Bayes Algorithm Method for Classification of Families at Risk of Stunting (Case Study: Waeapo District, Buru Regency) | Siti Adnan Rumanama<br>M. S. Noya Van Delsen<br>N. S. Laamena   | 111-118 |

# An Application of the Naïve Bayes Algorithm Method for Classification of Families at Risk of Stunting (Case Study: Waeapo District, Buru Regency)

Siti Adnan Rumanama<sup>1</sup>, M. S. Noya Van Delsen<sup>2\*</sup>, N. S. Laamena<sup>3</sup>

<sup>1,2,3</sup> Program Studi Statistika, FST Universitas Pattimura

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

Manuscript submitted : August 2024

Accepted for publication : September 2024

doi: <https://doi.org/10.30598/tensorvol5iss2pp67-74>

---

**Abstract :** Classification is a job of assessing data objects to put them into a certain class from a number of available classes. One algorithm that can be used for classification is the Naïve Bayes Classifier. Naïve Bayes Classifier is a probability concept that can be used to determine class groups of text documents and can process large amounts of data with high accuracy results. The aim of this research is to determine the results of the classification of families at risk of stunting in Waeapo District, Buru Regency and to determine the level of accuracy of three data proportions, namely 70:30, 80:20 and 90:10. The sample in this study was 2290 families. Based on the known level of accuracy, the best accuracy value is a data proportion of 90:10 with an accuracy value of 93.9%.

**Keywords:** Stunting, Data Mining, Naïve Bayes Classifier

---

## 1. Pendahuluan

Balita memerlukan asupan gizi yang lebih tinggi, baik dari segi kuantitas maupun kualitas, karena mereka masih dalam tahap belajar dan biasanya banyak melakukan aktivitas fisik sepanjang tahun-tahun awal kehidupan mereka [1]. Tiga sasaran utama pembangunan kesehatan Indonesia dari tahun 2015 hingga 2020 adalah terhindar dari penyakit menular dan tidak menular, menurunkan angka stunting, dan menurunkan angka kematian ibu dan bayi baru lahir. Salah satu prioritas pembangunan nasional yang tercantum dalam tujuan utama rencana pembangunan jangka menengah adalah meningkatkan status gizi masyarakat, yang meliputi penurunan prevalensi stunting atau kekurangan gizi pada balita.

Gizi buruk yang kronis sepanjang masa tumbuh kembang sejak usia dini disebut dengan istilah stunting. Angka z-score usia (TB/U) tinggi badan menandakan kondisi ini apabila kurang dari plus atau minus dua simpangan baku (SD) menurut kriteria pertumbuhan WHO [2]. Stunting berisiko bagi kesehatan manusia. Hal ini dikarenakan stunting dapat mengakibatkan tinggi badan rendah dan menghambat ginjal, jantung, otak, serta organ tubuh lainnya untuk tumbuh dan berkembang secara normal [3]. Oleh karena itu, stunting akan mengakibatkan berkurangnya fungsi organ tubuh, masalah psikomotorik, IQ, serta kapasitas motorik dan neurosensorik pada anak-anak. Lebih jauh lagi, stunting dapat memberikan dampak yang merugikan baik secara langsung maupun jangka panjang. Stunting berdampak jangka pendek pada kecerdasan anak, pertumbuhan fisik, metabolisme, serta pencernaan abnormal, sedangkan dampak jangka panjangnya antara lain berkurangnya kemampuan kognitif dan prestasi belajar, menurunnya kekebalan tubuh sehingga lebih

rentan terhadap penyakit, gangguan kesehatan reproduksi, risiko tinggi terkena penyakit diabetes, kanker, stroke, jantung dan pembuluh darah, serta mengalami disabilitas di usia lanjut [4].

Survei Status Gizi Indonesia (SSGI) melaporkan bahwa prevalensi stunting di Indonesia menurun dari 24,4% pada tahun 2021 menjadi 21,6% pada tahun 2022. Meskipun terjadi penurunan, angka tersebut masih lebih tinggi dari ambang batas yang ditetapkan WHO yaitu kurang dari 20% [5]. Maluku, yang menempati peringkat ke-13 di Indonesia, memiliki angka kejadian stunting sebesar 26,1%. Dibandingkan dengan tahun 2021, ketika angka prevalensi sebesar 28,7%, angka tahun ini lebih rendah. Sebaliknya, 23,3% balita di Kabupaten Buru mengalami stunting pada SSGI 2022 [6]. Oleh karena itu, pada penelitian kali ini akan dilakukan klasifikasi keluarga berisiko *stunting* terkhususnya di Kecamatan Waeapo Kabupaten Buru.

Metode data mining yang digunakan dalam penelitian ini adalah algoritma klasifikasi *Naïve Bayes Classifier*. *Naïve Bayes Classifier* merupakan salah satu alat klasifikasi statistik yang dapat digunakan untuk menentukan kemungkinan keanggotaan suatu kelas tertentu. Jika digunakan pada basis data dengan jumlah data yang besar, *Naïve Bayes Classifier* memiliki akurasi dan kecepatan yang sangat tinggi. *Naïve Bayes Classifier* merupakan salah satu algoritma yang mampu menurunkan ambang batas dibandingkan dengan semua klasifikasi lainnya [7].

## 2. Metode Penelitian

### 2.1. Variabel Penelitian

Data yang digunakan dalam penelitian ini terdiri dari 1 variabel dependen ( $Y$ ) dan 7 variabel independen ( $X$ ), dengan seluruh variabel bertipe kategorik. Penjelasan terkait masing-masing variabel diperlihatkan pada tabel berikut [8]:

Tabel 1. Variabel Penelitian

| Variabel | Keterangan   | Kategori   |
|----------|--|--|
| Y        | Kategori keluarga berisiko <i>stunting</i>                 | 0 Keluarga memenuhi kategori keluarga sasaran tersebut       |
|          |  | 1 Keluarga tidak memenuhi kategori keluarga sasaran tersebut |
| $X_1$    | Keluarga tidak mempunyai sumber air minum utama yang layak | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_2$    | Keluarga tidak mempunyai jamban yang layak                 | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_3$    | Umur istri terlalu muda ( Umur istri < 20 tahun)           | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_4$    | Umur istri terlalu tua (Umur istri 35 sd 40 tahun)         | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_5$    | Jarak kelahiran terlalu dekat (< 2 tahun)                  | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_6$    | Jumlah anak terlalu banyak ( $\geq 3$ anak)                | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |
| $X_7$    | Bukan peserta KB modern                                    | 0 Keluarga terpenuhi untuk variabel tersebut                 |
|          |  | 1 Keluarga tidak terpenuhi untuk variabel tersebut           |

### 2.2. Tahapan Penelitian

Tahapan analisis yang akan dilakukan dalam penelitian ini diuraikan sebagai berikut:

- a. Menganalisis data secara deskriptif untuk mengetahui gambaran umum mengenai data survei BKKBN Provinsi Maluku pada Kecamatan Waeapo Kabupaten Buru Tahun 2022. serta menentukan yang diperlukan dalam proses klasifikasi.

- b. Membagi data menjadi data *training* dan data *testing*. Proporsi data data *training* dan data *testing* yang digunakan, yaitu: 70:30, 80:20, 90:10.
- c. Pembuatan model klasifikasi dari data *training* dengan menggunakan algoritma *Naïve Bayes Classifier*, dengan tahapan-tahapan sebagai berikut:
  1. Menghitung probabilitas *prior* pada data *training*, yaitu probabilitas kategori keluarga berisiko stunting dan tidak berisiko stunting.
  2. Menghitung probabilitas setiap variabel pada data *training* terhadap masing-masing kelas.
- d. Menghitung ketepatan klasifikasi dengan cara menghitung akurasi, *precision* dan *recall* dalam proses pengklasifikasian.

### 3. Hasil dan Pembahasan

#### 3.1. Karakteristik Variabel

Data yang digunakan dalam penelitian ini berjumlah 2325 data. Namun setelah dilakukan *preprocessing* data terdapat 35 data yang *missing*, sehingga untuk proses analisis hanya menggunakan 2290 data. Untuk melihat gambaran umum dari data yang digunakan, maka karakteristik dari setiap variabel diperlihatkan pada tabel berikut:

Tabel 2. Karakteristik Variabel

| Variabel   | Kategori | Frekuensi | Persentase |
|--|----------|-----------|------------|
| Kategori keluarga berisiko <i>stunting</i>                 | 0        | 665       | 29,04      |
|  | 1        | 1625      | 70,96      |
| Jumlah   |          | 2290      | 100,00     |
| Keluarga tidak mempunyai sumber air minum utama yang layak | 0        | 43        | 1,88       |
|  | 1        | 2247      | 98,12      |
| Jumlah   |          | 2290      | 100,00     |
| Keluarga tidak mempunyai jamban yang layak                 | 0        | 404       | 17,64      |
|  | 1        | 1886      | 82,36      |
| Jumlah   |          | 2290      | 100,00     |
| Umur istri terlalu muda ( Umur istri < 20 tahun)           | 0        | 17        | 0,74       |
|  | 1        | 2273      | 99,26      |
| Jumlah   |          | 2290      | 100,00     |
| Umur istri terlalu tua (Umur istri 35 sd 40 tahun)         | 0        | 577       | 25,20      |
|  | 1        | 1713      | 74,80      |
| Jumlah   |          | 2290      | 100,00     |
| Jarak kelahiran terlalu dekat (< 2 tahun)                  | 0        | 10        | 0,44       |
|  | 1        | 2280      | 99,56      |
| Jumlah   |          | 2290      | 100,00     |
| Jumlah anak terlalu banyak ( $\geq 3$ anak)                | 0        | 542       | 23,67      |
|  | 1        | 1748      | 76,33      |
| Jumlah   |          | 2290      | 100,00     |
| Bukan peserta KB modern                                    | 0        | 787       | 34,37      |
|  | 1        | 1503      | 65,63      |
| Jumlah   |          | 2290      | 100,00     |

Tabel 2 memberikan informasi bahwa terdapat 665 keluarga berisiko stunting (29,04%) dan 1625 keluarga tidak berisiko stunting (70,96%). Selain itu juga diketahui bahwa jumlah keluarga yang tidak mempunyai sumber air minum utama yang layak sebesar 1,88% atau terdapat 43 keluarga dan sebesar 98,12% keluarga mempunyai sumber air minum utama yang layak atau terdapat 2247 keluarga. Sedangkan untuk keluarga tidak mempunyai jamban yang layak sebesar 17,64% atau terdapat 404 keluarga dan sebesar 82,36% keluarga mempunyai jamban layak atau terdapat 1886 keluarga. Untuk karakteristik dari Umur istri

terlalu muda diperoleh sebesar 0,74% keluarga yang memiliki istri terlalu muda atau terdapat 17 keluarga dan sebesar 99,26% keluarga yang memiliki istri tidak terlalu muda atau sebanyak 2273 keluarga. Sedangkan untuk karakteristik Umur istri terlalu tua diperoleh 577 keluarga yang memiliki istri berumur 35 sd 40 tahun dan 1713 keluarga yang memiliki istri tidak terlalu tua.

Berdasarkan Tabel 2 juga diperoleh informasi terkait karakteristik dari Jarak kelahiran terlalu dekat, yaitu terdapat 10 keluarga yang memiliki anak yang jarak usianya < 2 tahun dan 2280 keluarga yang memiliki anak yang usianya tidak terlalu dekat. Informasi lebih lanjut juga diperoleh terkait karakteristik dari Jumlah anak terlalu banyak, yaitu sebesar 23,67% atau terdapat 542 keluarga yang memiliki jumlah anak terlalu banyak dan sebesar 76,33% atau terdapat 1748 keluarga yang tidak memiliki anak terlalu banyak. Sedangkan untuk karakteristik Bukan peserta KB modern terdapat 787 keluarga bukan peserta KB modern dan 1503 keluarga merupakan peserta KB modern.

### 3.2. Algoritma Naïve Bayes

#### 3.2.1. Perhitungan Nilai Probabilitas *Prior* Dari Kategori Beresiko Stunting

Pada proporsi 70% data *training* dan 30% data *testing* diperoleh jumlah data *training* sebanyak 1603 keluarga dan data *testing* sebanyak 687 keluarga. Pada data *training* diperoleh 1138 keluarga memiliki status kategori keluarga berisiko stunting dan 465 keluarga memiliki status kategori keluarga tidak berisiko stunting. Sehingga nilai probabilitas *prior* keluarga berisiko stunting adalah 0,7099189 dan nilai probabilitas *prior* keluarga tidak berisiko stunting adalah 0,2900811.

Jika proporsi data yang digunakan 80% : 20%, maka pada data *training* diperoleh 1832 data yang terbagi menjadi 1300 keluarga berisiko stunting dan 532 keluarga tidak berisiko stunting. Sehingga nilai probabilitas dari keluarga berisiko dan tidak berisiko stunting dari data *training* pada proporsi ini adalah 0,709607 dan 0,290393. Sedangkan jika proporsi data yang digunakan 90%:10%, maka pada data *training* diperoleh 2060 data yang terbagi menjadi 1462 keluarga berisiko stunting dan 598 keluarga tidak berisiko stunting. Jadi nilai probabilitas *prior* keluarga berisiko stunting adalah 0,7097087 dan nilai probabilitas *prior* keluarga tidak berisiko stunting adalah 0,2902913.

#### 3.2.2. Perhitungan Nilai Probabilitas Setiap Variabel Independen Terhadap Setiap Kelas Variabel Dependen

Nilai probabilitas dari masing-masing variabel independen berdasarkan kategori keluarga berisiko stunting digunakan untuk melihat peluang kejadian setiap kategori variabel independen berdasarkan kategori variabel dependennya. Nilai-nilai probabilitasnya diperlihatkan pada tabel berikut ini:

Tabel 3. Nilai Probabilitas Setiap Variabel Independen Terhadap Setiap Kelas Variabel Dependen

| Variabel   | Kategori | Kategori keluarga berisiko <i>stunting</i> |       |
|--|----------|--|-------|
|  |          | 0  | 1     |
| Keluarga tidak mempunyai sumber air minum utama yang layak | 0        | 0,935                                      | 0,999 |
|  | 1        | 0,065                                      | 0,001 |
| Keluarga tidak mempunyai jamban yang layak                 | 0        | 0,400                                      | 0,999 |
|  | 1        | 0,600                                      | 0,001 |
| Umur istri terlalu muda ( Umur istri < 20 tahun)           | 0        | 0,987                                      | 0,993 |
|  | 1        | 0,013                                      | 0,007 |
| Umur istri terlalu tua (Umur istri 35 sd 40 tahun)         | 0        | 0,632                                      | 0,790 |
|  | 1        | 0,368                                      | 0,210 |
| Jarak kelahiran terlalu dekat (< 2 tahun)                  | 0        | 0,989                                      | 0,998 |
|  | 1        | 0,011                                      | 0,002 |
| Jumlah anak terlalu banyak ( $\geq 3$ anak)                | 0        | 0,665                                      | 0,816 |
|  | 1        | 0,335                                      | 0,184 |
| Bukan peserta KB modern                                    | 0        | 0,447                                      | 0,743 |

|   |       |       |
|---|-------|-------|
| 1 | 0,553 | 0,257 |
|---|-------|-------|

Berdasarkan Tabel 3 dapat dilihat bahwa Keluarga tidak mempunyai sumber air minum utama yang layak memiliki peluang menjadi keluarga beresiko stunting sebesar 0,935; jauh lebih besar dari keluarga yang mempunyai sumber air minum utama yang layak. Hal serupa juga terjadi untuk kategori keluarga tidak beresiko stunting, jika dilihat berdasarkan variabel keluarga tidak mempunyai sumber air minum utama yang layak. Pada kategori keluarga tidak mempunyai jamban yang layak, peluang keluarga beresiko stunting terbesar terdapat pada keluarga yang mempunyai jamban yang layak dengan nilai peluang 0,600. Sedangkan peluang keluarga tidak beresiko stunting terbesar terdapat pada keluarga yang tidak mempunyai jamban yang layak dengan nilai peluang sebesar 0,999.

Tabel 3 juga memperlihatkan peluang keluarga beresiko stunting terbesar terdapat pada keluarga yang memiliki istri terlalu muda dengan nilai peluang 0,987; dibandingkan dengan keluarga yang memiliki istri tidak terlalu muda dengan nilai peluang sebesar 0,013. Selain itu peluang keluarga tidak beresiko stunting terbesar juga dimiliki oleh keluarga yang memiliki istri terlalu muda (0,993). Pada kategori umur istri terlalu tua memiliki peluang beresiko stunting (0,632) lebih besar dari umur istri tidak terlalu tua (0,368). Untuk variabel yang sama, peluang keluarga tidak beresiko stunting terbesar juga terdapat pada umur istri terlalu tua (0,790).

Selanjutnya pada Tabel 3 juga dapat dilihat peluang keluarga beresiko stunting terbesar terdapat pada kategori jumlah kelahiran terlalu dekat (0,989), dibandingkan dengan kategori jumlah kelahiran tidak terlalu dekat (0,011). Sama halnya untuk keluarga tidak beresiko stunting, peluang terbesar terdapat pada kategori jumlah kelahiran terlalu dekat (0,989). Untuk kategori jumlah anak terlalu banyak, keluarga yang memiliki jumlah anak terlalu banyak berpeluang beresiko stunting lebih besar dibanding keluarga yang memiliki jumlah anak tidak terlalu banyak (0,665>0,335). Sedangkan peluang keluarga tidak beresiko stunting terbesar terdapat pada kategori jumlah anak terlalu banyak (0,816), jika dibandingkan dengan kategori jumlah anak tidak terlalu banyak (0,184). Berikutnya untuk variabel bukan peserta KB modern, peluang keluarga beresiko stunting terbesar terdapat pada kategori keluarga peserta KB modern (0,553) dibandingkan dengan kategori keluarga bukan peserta KB modern. Berbeda untuk keluarga tidak beresiko stunting, peluang terbesar terdapat pada kategori bukan peserta KB modern, dengan nilai peluang sebesar 0,743.

### 3.3. Ketepatan Klasifikasi Naïve Bayes

Suatu model klasifikasi ditetapkan sebagai model terbaik berdasarkan kriteria akurasi, sensitivitas dan spesifitas yang tertinggi. Hasil dari jumlah observasi yang benar diklasifikasikan ataupun yang misklasifikasi dapat disusun dalam sebuah matriks konfusi [9].

Sesuai dengan proporsi pembagian data *training* dan *testing* yang disebutkan pada tahapan analisis maka hasil ketepatan klasifikasi dari metode algoritma *Naïve Bayes* akan diperoleh 3 hasil klasifikasi, yaitu hasil klasifikasi untuk proporsi 70%:30%; 80%:20%; dan 90%:10%. Pembentukan matrik konfusi merupakan hasil klasifikasi dari data *testing* dari setiap proporsi pembagian data. Tabel 4 berikut ini memberikan informasi hasil klasifikasi keluarga beresiko stunting dengan menggunakan metode algoritma naïve bayes untuk proporsi 70%:30%; yang termuat dalam sebuah matriks konfusi.

Tabel 4. Matriks Konfusi Kategori Keluarga Beresiko Stunting

| Aktual                                    | Prediksi                            |   |
|---|-------------------------------------|---|
|   | Kategori Keluarga Beresiko Stunting | Kategori Keluarga Tidak Beresiko Stunting |
| Kategori Keluarga Beresiko Stunting       | 486                                 | 59  |
| Kategori Keluarga Tidak Beresiko Stunting | 1                                   | 141                                       |

Berdasarkan Tabel 4 diperoleh hasil klasifikasi kategori keluarga berisiko stunting menggunakan metode *Naive Bayes Classifier* adalah sebanyak 486 keluarga berisiko stunting terklasifikasi benar pada kategori keluarga berisiko stunting (Nilai *True Positive (TP)*), 59 keluarga berisiko stunting terklasifikasi salah pada kategori keluarga tidak berisiko stunting (Nilai *False Negative (FN)*), sebanyak 1 keluarga tidak berisiko stunting terklasifikasi salah pada kategori keluarga berisiko stunting (Nilai *False Positive (FP)*), dan sebanyak 141 keluarga tidak berisiko stunting terklasifikasi benar pada kategori keluarga tidak berisiko stunting (Nilai *True Negative (TN)*).

Selanjutnya, berdasarkan matriks konfusi pada Tabel 3 akan dihitung nilai-nilai yang digunakan untuk mengukur ketepatan hasil klasifikasi, yaitu nilai akurasi, sensitivitas dan spesifitas. Hasil perhitungan ketiga nilai tersebut diuraikan sebagai berikut.

Nilai akurasi digunakan untuk mengetahui hasil ketepatan klasifikasi yang diperoleh dari perbandingan jumlah data yang terklasifikasikan dengan benar dengan jumlah seluruh data [10].

$$\text{akurasi} = \frac{TN + TP}{FN + FP + TN + TP} = \frac{141 + 486}{59 + 1 + 141 + 486} = 0,913$$

Nilai 0,913 berarti, model klasifikasi yang diperoleh mempunyai tingkat ketepatan akurasi sebesar 91,3% untuk mengklasifikasikan data dengan benar.

Nilai sensitivitas berguna untuk mengetahui hasil ketepatan klasifikasi oleh metode algoritma Naïve Bayes pada kategori positif dengan benar [11]. Nilai sensitivitas diperoleh dengan cara berikut:

$$\text{sensitivitas} = \frac{TP}{FP + TP} = \frac{486}{1 + 486} = 0,998$$

Nilai sensitivitas yang sebesar 0,998 memiliki arti, metode algoritma Naïve Bayes dapat dengan tepat mengklasifikasikan kategori positif (keluarga berisiko stunting) dengan benar sebesar 99,8%.

Berikutnya nilai spesifitas berguna untuk mengetahui hasil ketepatan klasifikasi pada kategori negatif dengan benar [12]. Berikut hasil untuk nilai spesifitas:

$$\text{spesifitas} = \frac{TN}{FN + TN} = \frac{141}{59 + 141} = 0,705$$

Nilai spesifitas yang diperoleh sebesar 0,705; bermakna metode algoritma Naïve Bayes dapat dengan tepat mengklasifikasikan kategori negatif (keluarga tidak berisiko stunting) sebesar 70,5%.

Selanjutnya dengan menggunakan proporsi data 80%:20% dan 90%:10%; maka hasil akurasi dari setiap proporsi data, baik untuk proporsi data 70%:30%, 80%:20%, dan 90%:10% dapat diperlihatkan pada Tabel 5 berikut ini.

Tabel 5 Nilai Akurasi Ketepatan Klasifikasi *Naïve Bayes Classifier* Dari Masing-Masing Proporsi Data

| Proporsi Data | Akurasi | Sensitivitas | Spesifitas |
|---------------|---------|--------------|------------|
| 70:30         | 91,3%   | 99,8%        | 70,5%      |
| 80:20         | 91%     | 99,7%        | 69,9%      |
| 90:10         | 93,9%   | 99,4%        | 80,6%      |

Berdasarkan Tabel 4.18 diperoleh nilai akurasi klasifikasi terbaik dengan metode *Naïve Bayes* terdapat pada proporsi data 90%:10%, yaitu sebesar 93,9%. Sedangkan untuk nilai sensitivitas dan spesifitas terbaik masing-masing terdapat pada proporsi data 70%:30%, yaitu sebesar 99,8%; dan 90%:10%, yaitu sebesar 80,6%. Pada penelitian ini proporsi data terbaik ditentukan menggunakan nilai akurasi, karena nilai akurasi diperoleh dari persentasi jumlah keseluruhan data yang terklasifikasi benar dibandingkan keseluruhan data. Sehingga nilai akurasi lebih merepresentasikan seluruh hasil klasifikasi dibandingkan dengan nilai sensitivitas maupun spesifitas.

Sehingga hasil klasifikasi terbaik dari keluarga beresiko stunting di Kecamatan Waeapo, Kabupaten Buru, Provinsi Maluku Tahun 2022 dengan menggunakan metode algoritma Naïve Bayes terdapat pada proporsi data 90%:10%. Selain nilai akurasi dapat diperoleh juga nilai tingkat kesalahan suatu metode klasifikasi, atau yang disebut nilai APPER (*Apparent Error Rate*). Nilai APPER diperoleh dengan rumus berikut [13]:

$$APPER = \frac{FN + FP}{FN + FP + TN + TP} = \frac{13 + 1}{13 + 1 + 54 + 162} = 0,061$$

Nilai APPER yang diperoleh sebesar 0,061 atau 6,1%; yang artinya metode algoritma Naïve Bayes memiliki tingkat kesalahan sebesar 6,1% dalam mengklasifikasikan keluarga beresiko stunting di Kecamatan Waeapo, Kabupaten Buru, Provinsi Maluku Tahun 2022.

#### 4. Kesimpulan

Hasil klasifikasi kategori keluarga berisiko stunting di Kecamatan Waeapo Kabupaten Buru dengan menggunakan metode algoritma Naïve Bayes dengan pembagian data 70%:30%, 80%:20%, dan 90%:10%, diperoleh persentase keakuratan tertinggi yaitu pada pembagian data 90%:10% dengan tingkat akurasi sebesar 93,9% dan tingkat kesalahan sebesar 6,1%.

#### DAFTAR PUSTAKA

- [1] Ni'mah, K, & Nadhiroh, S. R. (2015, Januari-Juni). Faktor Yang Berhubungan Dengan Kejadian Stunting pada Balita. *Jurnal Media Gizi Indonesia*, 10(1), 13-19.
- [2] Kemenkes RI. (2020). Pedoman Gizi Seimbang. *Direktorat Bina Gizi dan KIA. Jakarta*.
- [3] Shauma, D. G., Nabila, U., & Purbaningrum. (2022). Implementasi Kebijakan Percepatan Pencegahan Stunting. *J. Kebijakan*, 13, 97-104.
- [4] Saputri, R. A. & J. Tumangger. (2019). Hulu-Hilir Penanggulangan Stunting Di Indonesia. *J. Polit. Issues*, 1 no 1, 1-9. doi:10.33019/jpi.v1i1.2
- [5] Rokom. (2023). Prevalensi Stunting Di Indonesia Turun ke 21,6% dari 24,4%. *Sehat Negeriku*. <https://sehatnegeriku.kemkes.go.id/baca/rilismedia/20230125/3142280/prevalensi-stunting-di-indonesia-turun-ke-216-dari-244/>
- [6] Annur, C. M. (2023). Prevalensi Balita Stunting Provinsi Maluku Menurut Kabupaten/Kota (2022). <https://databoks.katadata.co.id/datapublish/2023/02/07/prevalensi-balita-stunting-di-maluku-capai-261-pada-2022-ini-sebaran-wilayahnya>
- [7] Kusri, Luthfi, E. T., & Prabawati, T. A. (2009). Algoritma Data Mining. (Theresia Ari Prabawati(ed)). ANDI.
- [8] BKKBN. (2022). Prevalensi Balita Stunting Kabupaten Buru Tahun 2022. Ambon: BKKBN Wilayah Maluku.
- [9] M. Fadly, R., D., D., & Nurma, I. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN). *Jurnal Informatika*, 11(1).
- [10] Anam, S., Sugiman, S., & Sunarmi, S. (2017). Ketepatan Klasifikasi Dengan Menggunakan Metode Multivariate Adaptive Regression Spline (MARS) Pada Data Kelompok Rumah Tangga Kabupaten Cilacap. *Jurnal Matematika*, 6(1).
- [11] Fiqih, S., Zamharini, Z., & Apun Syaripudin, M. (2020). Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah dan Ilmu Komunikasi UIN Raden Intan Lampung. *Jurnal Ilmiah Matrik*, 22(1).

- [12] Eka Wulansari, F. (2015). Analisa Data Mining Untuk Prediksi Penyakit Hepatitis dengan Menggunakan Metode Naive Bayes dan Support Vector Machine. *Jurnal Khatulistiwa Informatika*, 3(1).
- [13] Johnson, R. A. dan Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall.