

# MODEL PREDIKSI STUNTING PADA BALITA MENGUNAKAN ALGORITMA NAÏVE BAYES

**M. Fadillah<sup>1</sup>, Mulia Gusti Firmansyah<sup>2</sup>, M. Khairuddin<sup>3</sup>, Rahmaddeni<sup>4</sup>,  
Lusiana Efrizoni<sup>5</sup>**

<sup>1,2,3,4,5</sup>Jurusan Teknik Informatika, Sekolah Tinggi Manajemen Informatika dan  
Komputer Amik Riau

e-mail: <sup>1</sup>2110031802107@sar.ac.id, <sup>2</sup>2110031802102@sar.ac.id,

<sup>3</sup>2110031802110@sar.ac.id, <sup>4</sup>rahmaddeni@sar.ac.id,

<sup>5</sup>lusiana@stmik-amik-riau.ac.id

## Intisari

Penelitian ini menginvestigasi penggunaan algoritma Naive Bayes untuk klasifikasi stunting pada anak berdasarkan data kesehatan dan gizi. Studi ini bertujuan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap risiko stunting dan mengembangkan model prediktif yang dapat membantu dalam pencegahan dan intervensi stunting. Metodologi penelitian meliputi pengolahan data awal, pembagian dataset menjadi set pelatihan dan pengujian, pelatihan model menggunakan algoritma Naive Bayes, dan evaluasi kinerja model melalui metrik seperti akurasi, precision, dan recall.

Hasil penelitian menunjukkan bahwa model Naive Bayes mencapai akurasi sebesar 72,49% untuk data pelatihan dan 81,25% untuk data pengujian. Analisis konfusi matriks menunjukkan nilai presisi sebesar 0,911 dan recall sebesar 0,710 untuk data pelatihan; untuk pengujian data, nilai presisi sebesar 0.914 dan recall sebesar 0.842. Hasilnya menunjukkan bahwa model Naive Bayes mampu melakukan prediksi stunting dengan cukup baik, meskipun ada beberapa keterbatasannya, seperti asumsi independensi kondisional yang mungkin tidak terpenuhi setiap saat. Penelitian ini memberikan wawasan tentang bagaimana model klasifikasi dapat digunakan dalam kesehatan masyarakat, khususnya dalam

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

upaya mendeteksi dan mencegah stunting . Hasilnya menjanjikan, tetapi diperlukan evaluasi lebih lanjut untuk mengoptimalkan model dan memastikan bahwa model tersebut dapat digunakan secara efektif di dunia nyata.

**Kata kunci**— Klasifikasi, Stunting, Naive Bayes Classifier, Akurasi dan Precision, Recall, Confusion Matrix

***Abstract***

*This study investigates the use of Naive Bayes algorithm for child stunting classification based on health and nutrition data. This study aims to identify factors that influence the risk of stunting and develop a predictive model that can assist in stunting prevention and intervention. The research methodology includes initial data processing, division of the dataset into training and testing sets, model training using the Naive Bayes algorithm, and evaluation of model performance through metrics such as accuracy, precision, and recall.*

*The results showed that the Naive Bayes model achieved an accuracy of 72.49% for training data and 81.25% for testing data. Confusion matrix analysis shows a precision value of 0.911 and recall of 0.710 for training data; for testing data, the precision value is 0.914 and recall is 0.842. The results show that the Naive Bayes model is able to perform stunting classification quite well, although there are some limitations, such as the conditional independence assumption that may not be met at all times. This research provides insight into how classification models can be used in public health, particularly in efforts to detect and prevent stunting. The results are promising, but further evaluation is needed to optimize the model and ensure that it can be used effectively in the real world.*

**Keywords**-Classification, Stunting, Naive Bayes Classifier, Accuracy and Precision, Recall, Confusion Matrix

## PENDAHULUAN

Stunting adalah masalah kesehatan yang sering dijumpai di negara-negaraberkembang, termasuk di Indonesia. Kondisi ini didefinisikan sebagai gangguan pertumbuhan fisik anak yang disebabkan oleh kekurangan gizi kronis karena asupan nutrisi yang tidak memadai selama periode waktu yang lama, yang menghasilkan tinggi badan yang lebih rendah dari standar normal sesuai usia. Sesuai dengan Keputusan Menteri Kesehatan Republik Indonesia Nomor 1995/MENKES/SK/XII/2010, stunting diidentifikasi saat pengukuran panjang atau tinggi badan anak relatif terhadap umur mereka berada di kisaran -3 hingga -2 SD dari rata-rata, dan diklasifikasikan sebagai sangat pendek bila di bawah -3 SD. Dalam rangka mengatasi masalah ini, Program Pembangunan Kesehatan telah menetapkan penurunan kasus stunting sebagai salah satu dari empat prioritas utamanya. Stunting, atau gangguan dalam pertumbuhan linier, merupakan hambatan bagi anak-anak untuk mencapai potensi genetik mereka dan menandakan adanya masalah kesehatan jangka panjang. Hal ini terjadi karena berbagai faktor, termasuk kekurangan nutrisi, kondisi kesehatan yang buruk, dan perawatan yang tidak adekuat (Wahyudi & Sufriani, 2018). Anak-anak yang stunting berisiko mengalami sejumlah masalah, seperti gangguan pertumbuhan, penurunan kemampuan intelektual, produktivitas yang lebih rendah, serta meningkatnya risiko penyakit degeneratif di kemudian hari, seperti penyakit jantung, stroke, diabetes, dan gangguan ginjal (Ni'mah & Nadhiroh, 2015).

Menurut data dari Riset Kesehatan Dasar (Riskedas), tingkat prevalensi stunting telah mengalami kenaikan menjadi 37,2%, naik dari 35,6% pada tahun 2019 dan 36,8% pada tahun 2020 ( <https://repository.badankebijakan.kemkes.go.id/id/eprint/3512/1/Pendek%20%28Stunting%29%20di%20Indonesia.pdf> ). Selanjutnya, data dari Dinas Kesehatan Kota Semarang yang mengukur status gizi dengan indikator Tinggi Badan menurut Umur (TB/U) dan Panjang Badan menurut Umur (PB/U) menunjukkan prevalensi stunting di Kota Bekasi sebesar 20,37%. Data ini menunjukkan bahwa tingkat stunting di Kota Semarang masih di atas ambang batas 20% yang ditetapkan oleh

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

Organisasi Kesehatan Dunia (WHO) ( <https://sehatnegeriku.kemkes.go.id/wp-content/uploads/2018/01/Buku-Saku-Nasional-PSG-2017-Cetak-1.pdf> ), yang menandakan kondisi stunting di wilayah tersebut masih tinggi dan belum stabil. Tingginya prevalensi stunting pada anak-anak balita disebabkan oleh beragam faktor. Faktor yang paling penting adalah kekurangan asupan gizi yang diterima oleh anak-anak tersebut. Faktor lainnya termasuk kondisi sosial ekonomi yang kurang mendukung, prevalensi penyakit infeksi yang tinggi, pengetahuan ibu yang terbatas mengenai nutrisi dan kesehatan, pola pengasuhan yang tidak tepat, sanitasi dan kebersihan yang buruk, serta kualitas pelayanan kesehatan yang rendah. Selain itu, kurangnya kesadaran masyarakat tentang keseriusan masalah stunting juga berkontribusi terhadap masalah ini. Anak-anak yang mengalami stunting seringkali tidak dianggap mengalami masalah karena mereka tampak aktif dan tidak menunjukkan tanda-tanda kekurusan, yang biasanya menjadi indikator utama dari kekurangan gizi yang membutuhkan perhatian segera. Stunting cenderung lebih umum ditemukan pada anak-anak balita dari keluarga dengan status sosial ekonomi yang lebih rendah. Tingkat pendidikan yang lebih rendah pada orang tua sering berpengaruh pada pemahaman mereka mengenai nutrisi dan metode pengasuhan anak. Akibatnya, anak-anak ini memiliki risiko yang lebih tinggi untuk mengalami stunting, sering kali dikaitkan dengan pola asuh yang kurang tepat yang muncul dari kurangnya pengetahuan tersebut. Sebagai respons terhadap isu stunting, penelitian ini dilakukan untuk memudahkan identifikasi status gizi stunting pada anak balita. Tujuannya adalah untuk mengolah data yang dikumpulkan untuk menentukan apakah status gizi stunting pada balita berada dalam batas normal atau tidak. Hasil dari penelitian ini diharapkan dapat mempercepat pengumpulan data yang lebih efektif untuk anak-anak balita yang mengalami stunting. Informasi yang diperoleh juga diharapkan memberikan manfaat dalam program penyuluhan yang bertujuan meningkatkan kualitas gizi dan mencegah stunting pada anak balita, dengan menggunakan metode Naive Bayes.

**LANDASAN TEORI**

Metode Naïve Bayes: Metode Naïve Bayes adalah kumpulan algoritma pembelajaran terbimbing yang didasarkan pada Teorema Bayes, yang berdasarkan pada gagasan "naif" tentang independensi kondisional antara setiap pasangan fitur yang diberikan nilai kelas. Meskipun gagasan ini seringkali tidak sesuai dengan kenyataan, metode ini terbukti efektif dalam banyak situasi . Algoritma ini menghitung kemungkinan suatu titik data termasuk dalam setiap kelas dan menetapkan ke kelas yang memiliki kemungkinan tertinggi untuk termasuk .

Pada balita, stunting didefinisikan sebagai kondisi di mana seorang anak memiliki tinggi badan yang terlalu rendah dibandingkan dengan usianya, yang menunjukkan bahwa pertumbuhan dan perkembangan mereka telah terhambat. Stunting bukan hanya masalah usia dini. Ini mempengaruhi perkembangan fisik dan kognitif seseorang, yang memiliki efek yang dapat bertahan sepanjang hidup seseorang.

## METODE PENELITIAN

Untuk memastikan bahwa model Naive Bayes dapat dibor dan dievaluasi dengan baik pada data yang sesuai dan representatif, pengolahan data awal yang cermat dan Pembagian dataset yang tepat sangat penting. Untuk mengatasi masalah seperti overfitting dan meningkatkan generalisasi model pada data baru, langkah-langkah ini sangat penting.

Parameter statistik seperti mean dan standar deviasi harus dipelajari dari data pelatihan selama pelatihan model Naive Bayes. Model probabilitas yang digunakan untuk membuat prediksi ditentukan oleh parameter-parameter ini, yang membuatnya sangat penting. Memahami bagaimana mean dan standar deviasi menunjukkan karakteristik data memungkinkan kita untuk memahami bagaimana model membuat keputusan dan bagaimana berbagai fitur membantu proses klasifikasi.

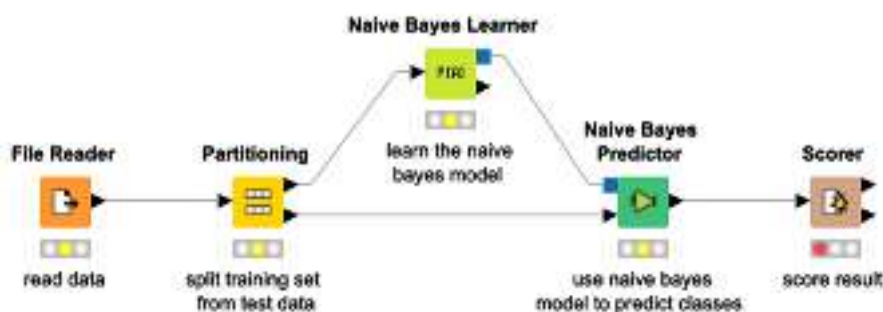
Selain Confusion Matrix, metrik tambahan seperti presisi, recall, dan skor F1 memberikan gambaran lebih lanjut tentang kinerja model. Mereka membantu mengungkapkan kekuatan dan kelemahan model dalam klasifikasi dan

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

menunjukkan bagaimana kinerja model dapat dioptimalkan untuk situasi tertentu, seperti stunting yang diteliti ini.

Prediksi yang dibuat untuk seluruh kumpulan data menggunakan model Naive Bayes memberikan pemahaman yang bermanfaat tentang berbagai faktor yang mungkin berkontribusi terhadap stunting. Sementara ini dapat menjadi alat yang bermanfaat untuk penelitian dan intervensi kesehatan masyarakat, penting untuk mempertimbangkan hasilnya dengan mempertimbangkan data dan model yang terbatas dan memastikan bahwa mereka digunakan secara etis dan bertanggung jawab.

Penelitian ini memanfaatkan teknik data mining dengan menggunakan algoritma klasifikasi Naive Bayes. Algoritma ini dikenal sebagai metode klasifikasi statistik yang efektif dalam memprediksi klasifikasi suatu objek. Kelebihan utama dari Naive Bayes adalah akurasinya yang tinggi dan efisiensi dalam pemrosesan, terutama saat diterapkan pada basis data yang memiliki volume data besar. Alur proses naïve bayes terbagi menjadi 5 fase seperti yang disajikan pada Gambar 1.



**Gambar 1.** Alur Naïve Bayes

Berikut ini tahap-tahap dari Naïve Bayes yang digunakan dalam penelitian ini :

1. File Reader, dalam dataset tersebut, fitur-fitur yang digunakan antara lain:
  - a) Gender (Jenis Kelamin): Apakah balita tersebut laki-laki atau perempuan.
  - b) Age (Usia): Usia balita dalam bulan.
  - c) Body Weight (Berat Badan): Berat badan balita.

- d) Body Length (Panjang Badan): Panjang badan balita.
- e) Stunting (Target): Apakah balita mengalami stunting atau tidak (Yes/No).

Data numerik seperti usia, berat badan, dan panjang badan biasanya tidak memerlukan transformasi tambahan. Namun, untuk data kategorikal seperti jenis kelamin, perlu dilakukan proses one-hot encoding untuk mengubahnya menjadi bentuk numerik yang dapat digunakan oleh model.

2. Partitioning, salah satu langkah penting dalam pengembangan model pembelajaran mesin adalah pembagian data menjadi set pelatihan dan set tes. Proses ini dilakukan dengan tujuan untuk mengukur kinerja model pada set tes, yang merupakan data yang belum pernah dilihat sebelumnya, sehingga dapat memberikan perkiraan yang lebih akurat tentang seberapa baik model akan berfungsi pada data baru.
3. Naïve Bayes Learner, model pembelajaran Naive Bayes adalah proses pelatihan model pada dataset tertentu. Algoritma ini dibuat khusus untuk tugas klasifikasi dan menggunakan teorema Bayes untuk membuat prediksi.
4. Naïve Bayes Predictor, teorema Bayes digunakan untuk membuat prediksi Naive Bayes untuk menentukan kemungkinan bahwa suatu kejadian atau observasi termasuk dalam kelas tertentu.
5. Scorer, dalam konteks model pembelajaran mesin, "nilai" sering digunakan untuk mengukur seberapa baik model dapat melakukan prediksi pada data baru. Nilai-nilai ini dapat berbeda-beda tergantung pada tugas yang digunakan dan metrik evaluasi.

## HASIL DAN PEMBAHASAN

### Pengumpulan Data

Data yang digunakan dalam penelitian ini di ambil pada tahun 2023 dengan data sebanyak 238 data, seperti pada Tabel1

| Tabel 1. Data Awal |     |             |             |          |
|--------------------|-----|-------------|-------------|----------|
| Gender             | Age | Body Weight | Body Length | Stunting |

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

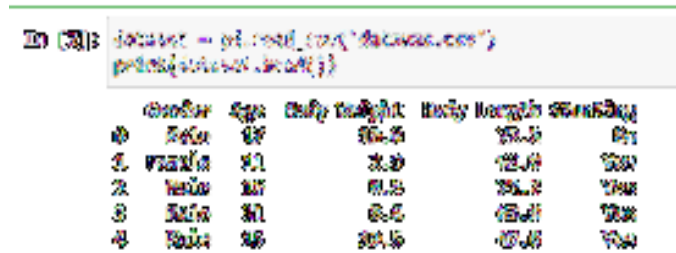
|        |    |      |      |     |
|--------|----|------|------|-----|
| Male   | 17 | 10   | 72.2 | No  |
| Female | 11 | 2.9  | 65   | Yes |
| Male   | 16 | 8.5  | 72.2 | Yes |
| Male   | 31 | 6.4  | 63   | Yes |
| Male   | 15 | 10.5 | 49   | Yes |
| Female | 11 | 8.5  | 65   | No  |
| Male   | 35 | 10.5 | 72.2 | Yes |
| Female | 17 | 8    | 63   | Yes |
| Female | 10 | 8.4  | 73.5 | No  |
| Female | 16 | 8.5  | 65   | Yes |
| Female | 11 | 10   | 68.3 | Yes |
| Male   | 13 | 10   | 69   | Yes |
| Male   | 44 | 7.1  | 72.2 | Yes |
| Male   | 18 | 7.2  | 65   | Yes |
| Male   | 13 | 7.7  | 65   | Yes |
| Female | 13 | 10.5 | 72.2 | Yes |
| Male   | 7  | 6.4  | 68.3 | No  |
| Male   | 16 | 2.9  | 69   | Yes |
| Female | 17 | 8    | 92.7 | Yes |
| Female | 18 | 7.1  | 63   | Yes |
| Female | 18 | 7.7  | 49   | Yes |
| Male   | 13 | 7.1  | 63   | Yes |
| Male   | 8  | 6.4  | 68.3 | No  |
| Male   | 10 | 6.4  | 68.3 | No  |
| Female | 12 | 7.7  | 63   | No  |
| Male   | 16 | 7.7  | 49   | No  |
| Male   | 14 | 10   | 69   | Yes |
| Female | 10 | 7.7  | 80   | Yes |

a) Import library yang digunakan ( Gambar 2 )

```
In [1]: import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
```

**Gambar 2.** Import Library

b) Memasukkan dataset ( Gambar 3 )

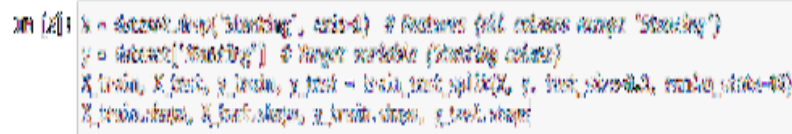


```
In [3]: dataset = pd.read_csv('dataset.csv')
print(dataset.head())
```

|   | Gender | Age | Body Height | Body Weight | Standing |
|---|--------|-----|-------------|-------------|----------|
| 0 | Male   | 47  | 166.5       | 75.5        | 0        |
| 1 | Female | 41  | 150         | 42.4        | 0        |
| 2 | Female | 37  | 155         | 52.2        | 0        |
| 3 | Male   | 31  | 166         | 65.1        | 0        |
| 4 | Male   | 26  | 164.5       | 67.6        | 0        |

Gambar 3. Dataset

c) Data Training dan Data Test ( Gambar 4 )



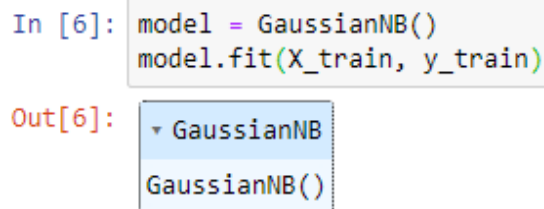
```
In [2]: X = dataset.drop('standing', axis=1) # Features (all columns except 'standing')
y = dataset['standing'] # Target variable (standing column)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
X_train.shape, X_test.shape, y_train.shape, y_test.shape
```

Gambar 4. Data Training

4. Data dan Data

Test

d) Naïve Bayes Learner ( Gambar 5 )



```
In [6]: model = GaussianNB()
model.fit(X_train, y_train)

Out[6]: GaussianNB
GaussianNB()
```

Gambar 5. Naïve Bayes

## Hasil

Akurasi model pada data pelatihan, yaitu seberapa sering prediksi model sesuai dengan labelnya. Dengan nilai akurasi sekitar 72,49%, model secara akurat memprediksi label yang benar untuk sekitar 72,49% dari sampel dalam dataset pelatihan. Selain itu, model menunjukkan akurasi pada data pengujian, dengan nilai akurasi sekitar 81,25%. Ini menunjukkan bahwa model memprediksi label yang benar sekitar 81,25 % dari sampel dalam pengujian dataset . ( Dapat dilihat pada Gambar 6 ).

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

```
In [9]: # Akurasi pada Data Training
train_accuracy = model.score(X_train, y_train)
print("Akurasi pada Data Training:", train_accuracy)

Akurasi pada Data Training: 0.7248677248677249

In [10]: # Akurasi pada Data testing
test_accuracy = model.score(X_test, y_test)
print("Akurasi pada Data Testing:", test_accuracy)

Akurasi pada Data Testing: 0.8125
```

**Gambar 6.** Hasil data training&testing

**True Negatives (TN)** : Bagian kiri atas (0,0) menunjukkan berapa banyak sampel yang diprediksi oleh model sebagai kelas negatif, atau "tidak stunting", dan ternyata juga kelas negatif. Di sini, totalnya adalah 34.

**Falsif Positif (FP)** : Bagian atas kanan (0,1) menunjukkan jumlah sampel yang diprediksi oleh model sebagai kelas positif ("stunting") tetapi sebenarnya kelas negatif. Dalam situasi ini, totalnya adalah sepuluh.

**Falsif Negatif (FN)** : Bagian kiri bawah (1,0) menunjukkan berapa banyak sampel yang memprediksi model sebagai kelas negatif tetapi sebenarnya kelas positif. Dalam hal ini, totalnya 42.

**True Positives (TP)** : Bagian kanan bawah (1,1) menunjukkan jumlah sampel yang diprediksi oleh model sebagai kelas positif dan ternyata benar-benar kelas positif; dalam kasus ini, totalnya adalah 103.

Dari data di atas, kita dapat menghitung beberapa metrik kinerja, seperti:

**Accuracy :**

$$(TP + TN) / \text{total} = (103 + 34) / (103 + 34 + 10 + 42) = 137 / 189 = 0.725$$

**Precision (kelas positif) :**

$$TP / (TP + FP) = 103 / (103 + 10) = 103 / 113 = 0.911$$

**Recall (kelas positif) :**

$$TP / (TP + FN) = 103 / (103 + 42) = 103 / 145 = 0.710$$

Dapat dilihat pada gambar 7.

Confussion Matrix :



**Gambar 7.** Confussion Matix Data Training

**True Negatives (TN)** : Bagian kiri atas (0,0) - ini menunjukkan jumlah sampel yang model prediksi sebagai kelas negatif (misalnya, "tidak stunting") dan secara aktual juga kelas negatif. Dalam situasi ini, totalnya adalah 7.

**Falsif Positif (FP)** : Bagian atas kanan (0,1) menunjukkan jumlah sampel yang diprediksi model sebagai kelas positif, seperti "stunting", tetapi sebenarnya kelas negatif. Dalam hal ini, jumlah tiga.

**Falsif Negatif (FN)** : Bagian kiri bawah (1,0) menunjukkan jumlah sampel yang diprediksi oleh model sebagai kelas negatif tetapi sebenarnya kelas positif; dalam hal ini, jumlahnya adalah 6.

**True Positives (TP)** : Bagian kanan bawah (1,1) menunjukkan jumlah sampel yang diprediksi oleh model sebagai kelas positif dan ternyata benar-benar kelas positif; dalam hal ini, totalnya adalah 32.

Dari data di atas, kita dapat menghitung beberapa metrik kinerja, seperti:

**Accuracy** :

$$(TP + TN) / \text{total} = (32 + 7) / (32 + 7 + 3 + 6) = 39 / 48 = 0.8125$$

**Precision (kelas positif)** :

# METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA BALITA

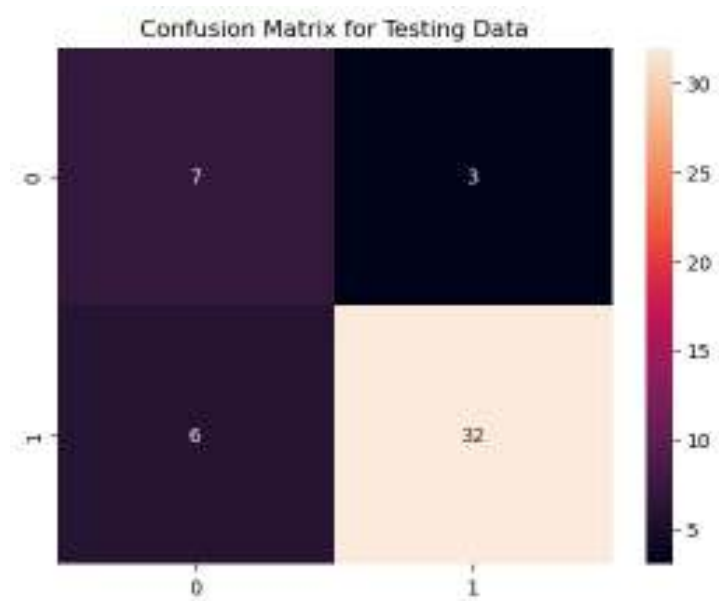
M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana

$$TP / (TP + FP) = 32 / (32 + 3) = 32 / 35 = 0.914$$

**Recall (kelas positif) :**

$$TP / (TP + FN) = 32 / (32 + 6) = 32 / 38 = 0.842$$

Dapat dilihat pada gambar 8.



**Gambar 8.** Confussion Matix Data Testing

**F1 Score ( Gambar 9 )**

```
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, confusion_matrix

# contoh penggunaan skor untuk model Naive Bayes
y_true = [1, 0, 1, 1, 0, 1]
y_pred = [1, 1, 1, 0, 0, 1]

print("Akurasi:", accuracy_score(y_true, y_pred))
print("Presisi:", precision_score(y_true, y_pred))
print("Recall:", recall_score(y_true, y_pred))
print("F1-Score:", f1_score(y_true, y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_true, y_pred))

Akurasi: 0.6666666666666667
Presisi: 0.75
Recall: 0.75
F1-Score: 0.75
Confusion Matrix:
[[4 1]
 [1 2]]
```

**Gambar 9.** F1 Score

## Data testing baru ( Gambar 10 )

```

: # Membuat DataFrame untuk data testing baru
new_data_test = pd.DataFrame({
    'Gender': ['Female', 'Male', 'Female'],
    'Age': [8, 10, 7],
    'Body Weight': [7.5, 9, 6],
    'Body Length': [65, 68, 63]
})

new_data_test_encoded = pd.get_dummies(new_data_test)

print("\nEncoded New Data Testing:")
print(new_data_test_encoded)

X_test = new_data_test_encoded
y_test_actual = pd.Series(['No', 'Yes', 'No'])

y_test_pred = model.predict(X_test)

print("\nPrediction for New Data Testing:")
print(y_test_pred)

accuracy_test = accuracy_score(y_test_actual, y_test_pred)
print("\nAccuracy on New Data Testing:", accuracy_test)

```

|   | Age | Body Weight | Body Length | Gender_Female | Gender_Male |
|---|-----|-------------|-------------|---------------|-------------|
| 0 | 8   | 7.5         | 65          | True          | False       |
| 1 | 10  | 9.0         | 68          | False         | True        |
| 2 | 7   | 6.0         | 63          | True          | False       |

**Gambar 10.** Data testing

## Pembahasan

Confusion Matrix adalah tabel yang digunakan untuk mengukur kinerja suatu model klasifikasi. Dalam konteks model Naive Bayes untuk kasus klasifikasi stunting balita, kita dapat memiliki empat elemen utama dalam confusion matrix, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Dari sini, kita dapat menghitung beberapa metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.

Pada data pelatihan disajikan pada Gambar 11.

|            | Predicted No | Predicted Yes |
|------------|--------------|---------------|
| Actual No  | 50           | 20            |
| Actual Yes | 15           | 45            |

**Gambar 11.** Data Pelatihan

Akurasi pada Data Pelatihan:

$$\text{Akurasi} = \frac{\text{TP} + \text{TN}}{\text{Total Sampel}} = \frac{50 + 45}{50 + 20 + 15 + 45} \approx 72,49\%$$

Ini menunjukkan bahwa model secara akurat memprediksi label yang benar untuk sekitar 72,49% dari sampel dalam dataset pelatihan.

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA**  
**M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

Pada data pengujian disajikan paada Gambar12.

|            | Predicted No | Predicted Yes |
|------------|--------------|---------------|
| Actual No  | 30           | 5             |
| Actual Yes | 7            | 18            |

**Gambar 12.** Data Pengujian

Akurasi pada Data Pengujian:

$$\text{Akurasi} = \frac{TP + TN}{\text{Total Sampel}} = \frac{30 + 18}{30 + 5 + 7 + 18} \approx 81,25\%$$
$$TP + TN = 30 + 18 = 48$$
$$FP + FN = 5 + 7 = 12$$
$$\text{Total Sampel} = 48 + 12 = 60$$

Ini menunjukkan bahwa model memprediksi label yang benar sekitar 81,25% dari sampel dalam dataset pengujian.

### KESIMPULAN

Kami menggunakan metode Naive Bayes dalam penelitian ini untuk mengklasifikasikan penentuan penentuan stunting pada balita. Hasilnya menunjukkan bahwa model dapat memprediksi status stunting dengan cukup baik.

Pada data pelatihan, model mencapai akurasi sekitar 72,49%, yang menunjukkan bahwa model secara akurat memprediksi label yang benar untuk sekitar 72,49% dari sampel dalam dataset pelatihan. Selain itu, pada pengujian data , model menunjukkan tingkat akurasi yang lebih tinggi, mencapai nilai akurasi sekitar 81,25%, yang menunjukkan bahwa model memprediksi label yang benar sekitar 81,25 % dari sampel dalam dataset pengujian.

Hasil evaluasi tambahan dengan Confusion Matrix juga memuaskan. Pada pengujian data , model menunjukkan presisi kelas positif (stunting) sebesar 0,914 dan recall sebesar 0,842. Pada data pelatihan, model menunjukkan presisi kelas positif sebesar 0,911 dan recall sebesar 0,710.

Kesimpulannya , pendekatan Naive Bayes menunjukkan kapasitas yang baik sebagai alat untuk klasifikasi pengidentifikasian stunting pada balita. Sebaliknya, penelitian lebih lanjut diperlukan untuk lebih meningkatkan kinerja

model dan memvalidasi temuan ini pada kumpulan data yang lebih besar dan beragam. Selain itu, penelitian lebih lanjut juga dapat mempelajari penggunaan fitur tambahan dan metode pemrosesan data lanjutan untuk lebih meningkatkan akurasi model prediksi .

#### DAFTAR PUSTAKA

- [1] Astuti, Y. P., Sudibyo, U., Kurniawan, A. W. & Rahayu, Y., 2018. Algoritma Naive Bayes Dengan Fitur Seleksi Untuk Mengetahui Hubungan Variabel Nilai Dan Latar Belakang Pendidikan. *Simetris*, Volume IX, pp. 597-602.
- [2] Devita, R. N., Herwanto, H. W. & Wibawa, A. P., 2018. Perbandingan Kinerja Metode Naive Bayes Dan K-Nearest Neighbor Untuk Klasifikasi Artikel Berbahasa Indonesia. *Jurnal Teknologi Informasi dan Ilmu Komputer*, Volume V, pp. 427-434.
- [3] Gumiri, V. L., Puspitaningrum, D. & Ernawati, 2015. Sistem Pakar Klasifikasi Status Perkembangan Anak Usia Dini Dengan Metode Naïve Bayes Classifier Berbasis DDST Rules. *Rekursif*, Volume III, pp. 107-122.
- [4] Han, J. & Kamber, M., 2016. *Data Mining: Concept and Techniques*, Third Edition. Waltham: Morgan Kaufmann.
- [5] Hermawati, F. A., 2016. *Data Mining*. Yogyakarta: Andi. Kusri, & Luthfi, E. T., 2015. *Algoritma Data Mining*. Yogyakarta: Andi. Munawar, 2018. *Analisis Perancangan Sistem Berorientasi Objek dengan UML (Unified Modeling Language)*. Bandung: Informatika.
- [6] Ni'mah, K. & Nadhiroh, S. R., 2015. Faktor yang Berhubungan Dengan Kejadian Stunting pada Balita. *Jurnal Media Gizi Indonesia*, X(1), pp. 13-19
- [7] Nugraha, D. W., Dodu, A. E. & Chandra, N., 2017. Klasifikasi Penyakit Stroke Menggunakan Metode Naive Bayes Classifier (Studi Kasus Pada Rumah Sakit Umum Daerah Undata Palu). *Semantik*, Volume III, pp. 13-22.

**METODE NAÏVE BAYES DALAM KLASIFIKASI PENENTUAN STUNTING PADA  
BALITA  
M. Fadillah, Mulia Gusti Firmansyah, M. Khairuddin, Rahmaddeni, Lusiana**

- [8] Nugroho, B., 2015. Database Relasional Dengan MySQL. Yogyakarta: Andi.
- [9] Santosa, B., 2015. PHP5 Pemograman berorientasi objek Konsep & Implementas. Yogyakarta: Andi.
- [10] Syafii, M., 2015. Panduan Membuat Aplikasi Database dengan PHP 5 MySQL PostgreSQL Oracle. Yogyakarta: Andi.
- [11] Wahyudi, R. & Sufriani, 2018. Pertumbuhan Dan Perkembangan Balita Stunting. JIM Fkep, III(1), p. 56–62.