

Klasifikasi Pemenuhan Pilar Sanitasi Puskesmas Menggunakan Metode *Naive Bayes*

Muhammad Farhan Syam^{1*}, Lilis Nur Hayati², Lukman Syafie³

^{1,3})Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Muslim Indonesia
Jl. Urip Sumoharjo KM. 5, Makassar, Indonesia 90231

²) Program Sistem Informasi, Fakultas Ilmu Komputer, Universitas Muslim Indonesia
Jl. Urip Sumoharjo KM. 5, Makassar, Indonesia 90231

*email: 13020190005@umi.ac.id

(Naskah masuk: 6 Juli 2023; diterima untuk diterbitkan: 26 September 2023)

ABSTRAK – Sanitasi merupakan usaha dalam menjaga kebersihan serta kondisi lingkungan sekitar. Dalam memenuhi pilar sanitasi di setiap wilayah pastinya kita juga membutuhkan peran instansi kesehatan untuk melakukan pemecuan dan pemberian edukasi. Pada desa area lingkup Puskesmas Bontomangape diketahui bahwa pemenuhan pilar sanitasi masih belum merata. Berdasarkan hal tersebut, penulis bermaksud untuk melakukan klasifikasi pemenuhan pilar sanitasi puskesmas menggunakan metode *Naive Bayes* sehingga hasil klasifikasi tersebut dapat digunakan sebagai tolak ukur desa yang perlu diprioritaskan oleh petugas sanitasi. Adapun hasil klasifikasi diperoleh sebanyak 56 terlaksana dan 19 tidak terlaksana untuk desa Bontomangape, 70 terlaksana dan 5 tidak terlaksana untuk desa Campagaya, 60 terlaksana dan 15 tidak terlaksana untuk desa Kalenna, 45 terlaksana dan 30 tidak terlaksana untuk desa Parambambe, 52 terlaksana dan 23 tidak terlaksana untuk desa Parangmata, 64 terlaksana dan 11 tidak terlaksana untuk desa Parasangangberu, serta 57 terlaksana dan 18 terlaksana untuk desa Pattinoang. Adapun hasil klasifikasi memperoleh rata-rata nilai akurasi sebesar 96,01%, nilai presisi sebesar 95,03% dan nilai recall sebesar 100%.

Kata Kunci – Sanitasi; Kesehatan; Puskesmas; Klasifikas; *Naive Bayes*.

Classification of Fulfillment of Health Center Sanitation Pillars Using the Naive Bayes Method

ABSTRACT – Sanitation is an attempt to maintain the cleanliness and condition of the surrounding environment. In fulfilling the sanitation pillar in each region, of course we also need the role of health agencies to trigger and provide education. In the village where the scope of the Bontomangape Health Center is located, it is known that the fulfillment of the sanitation pillar is still uneven. Based on this, the author intends to classify the fulfillment of the sanitation pillars of the Bontomangape Health Center using the *Naive Bayes* method so that the results of this classification can be used as a benchmark for villages that need to be prioritized by sanitation workers. The classification results obtained were 56 implemented and 19 not implemented for Bontomangape village, 70 implemented and 5 not implemented for Campagaya village, 60 implemented and 15 not implemented for Kalenna village, 45 implemented and 30 not implemented for Parambambe village, 52 implemented and 23 not implemented implemented for Parangmata village, 64 implemented and 11 not implemented for Parasangangberu village, and 57 implemented and 18 implemented for Pattinoang village. The classification results obtained an average accuracy value of 96,01%, a precision value of 95,03% and a recall value of 100%.

Keywords – Sanitation; Health; Puskesmas; Classification; *Naive Bayes*.

1. PENDAHULUAN

Kesehatan merupakan suatu hal yang sangat penting bagi seluruh warga di Indonesia. Dimana pada saat ini telah banyak program-program yang telah dilakukan oleh pemerintah khususnya dalam bidang kesehatan untuk meningkatkan kepedulian warga terhadap lingkungan tempat tinggalnya salah satunya yaitu sanitasi total berbasis masyarakat (STBM) [1]. Sanitasi merupakan perilaku hidup sehat dengan cara melakukan usaha dalam menjaga kebersihan serta kondisi lingkungan sekitar [2]. Adapun terdapat lima pilar sanitasi yang seharusnya dipenuhi oleh setiap masyarakat yaitu stop buang air besar disembarangan tempat, cuci tangan menggunakan sabun, pengelolaan air minum dan makanan, pengelolaan sampah serta pengelolaan limbah cair [3]. Apabila sanitasi dalam lingkungan masyarakat tidak diperhatikan, maka dapat memungkinkan warga terkena penyakit seperti diare, scabies, infeksi saluran pernapasan atas (ISPA) maupun penyakit yang mengganggu sistem pernapasan. Oleh karena itu, diperlukan kontribusi masyarakat dalam memenuhi pilar sanitasi ditempat tinggalnya untuk mengurangi kemungkinan terkena penyakit yang diakibatkan oleh tidak terpenuhinya salah satu dari lima pilar sanitasi tersebut [4].

Dalam memenuhi pilar sanitasi di setiap wilayah tertentu pastinya kita juga membutuhkan peran instansi kesehatan, seperti petugas sanitasi puskesmas. Terdapat peran petugas sanitasi puskesmas yaitu membantu keberhasilan program STBM dengan cara melaksanakan seluruh kegiatan terkait pemenuhan pilar sanitasi di tempat tinggal warga yang termasuk area lingkup wilayahnya [5]. Adapun kegiatan yang akan dilakukan petugas sanitasi puskesmas yaitu seperti melakukan observasi, pencatatan data, penginputan data, klasifikasi data, pelaporan, evaluasi serta edukasi dan pemecuan berdasarkan data warga yang telah diperoleh. Keseluruhan data warga sebuah desa yang telah diklasifikasikan antara telah atau belum melaksanakan pemenuhan pilar sanitasi nantinya dapat menjadi tolak ukur bahwa warga desa telah menjalankan program STBM serta menentukan desa mana yang harus diprioritaskan dalam melakukan edukasi dan pemecuan STBM. Dimana diketahui pada saat ini kondisi pemenuhan pilar sanitasi warga tiap desa belum merata khususnya desa yang merupakan area lingkup Puskesmas Bontomangape.

Klasifikasi adalah suatu kegiatan yang bertujuan untuk mengelompokkan data berdasarkan ciri khas suatu data sehingga dapat dilakukan pemberian kategori pada data tersebut [6]. Dalam melakukan klasifikasi data tentunya kita perlu menggunakan metode yang sesuai dengan data yang akan kita olah serta memiliki keunggulan dibandingkan metode

klasifikasi lainnya. Adapun metode *Naive Bayes* merupakan suatu metode yang dapat digunakan dalam mengklasifikasikan suatu variabel tertentu sebab *Naive Bayes* dapat melakukan klasifikasi dengan jumlah dataset yang banyak maupun dengan dataset yang sedikit. Oleh karena itu, dalam beberapa penelitian mengenai klasifikasi lebih banyak menggunakan *Naive Bayes* dalam melakukan klasifikasi suatu data karena *Naive Bayes* dapat memberikan keuntungan lebih serta dapat bekerja lebih baik dibandingkan dengan algoritma *classifier* lainnya [7].

Berdasarkan penelitian yang dilakukan oleh Vebby Regista dkk dengan judul "Gambaran Pelaksanaan 5 Pilar Sanitasi Total Berbasis Masyarakat di Kelurahan Kasturian Kota Ternate" memperoleh hasil analisis pemenuhan pilar 1 sebesar 20,9%, pilar 2 sebesar 83,7%, pilar 3 sebesar 100%, pilar 4 sebesar 62,8% dan pilar 5 sebesar 27,9% [8].

Penelitian yang dilakukan oleh Alfanz Zainal Macfud dkk dengan judul "Analisis Algoritma *Naive Bayes Classifier* (NBC) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell" didapatkan hasil klasifikasi 23 barang kategori minat rendah, 5 barang kategori minat sedang, dan 1 barang kategori minat tinggi dengan akurasi sebesar 82,76% [9].

Penelitian yang dilakukan oleh May Siti Syarah dkk dengan judul "Klasifikasi Penderita ISPA Menggunakan Metode *Naive Bayes Classifier*" didapatkan hasil pasien ISPA terbanyak termasuk kedalam jenis berat dengan jumlah 18 orang dan nilai akurasi diperoleh sebesar 93,33% [10].

Penelitian yang dilakukan oleh Hakam Febtadianrano Putro dkk dengan judul "Penerapan Metode *Naive Bayes* Untuk Klasifikasi Pelanggan" dengan hasil yaitu klasifikasi 21 pelanggan berpotensi dan 2 tidak berpotensi dengan akurasi sebesar 92%, *precision* 100%, dan *recall* 91% [11].

Penelitian yang dilakukan oleh Eka Fitriani dengan judul "Implementasi Metode *Naive Bayes* Dalam Penyeleksian Karyawan Untuk Penempatan Bagian Pemasaran" diperoleh hasil bahwa algoritma ini tergolong baik untuk seleksi karyawan dan diperoleh akurasi sebesar 87,22% [12].

2. METODE DAN BAHAN

2.1 Deskripsi Data

Data yang digunakan dalam penelitian ini merupakan dataset hasil observasi pemenuhan pilar sanitasi puskesmas pada warga desa area lingkup Puskesmas Bontomangape. Adapun dataset diperoleh langsung dari petugas sanitasi Puskesmas Bontomangape. Dataset yang diperoleh berjumlah 1750 data yang telah diletakkan pada *file excel* dengan masing-masing file terbagi atas 250 data disetiap satu *file excel* untuk 7 desa berbeda-beda. Adapun salah satu contoh *file excel dataset* suatu desa dapat kita

lihat pada gambar 1 berikut ini.

	Nama Warga	Pilar 1	Pilar 2	Pilar 3	Pilar 4	Pilar 5	Keterangan
0	Herman dg. Tika	Baik	Baik	Baik	Buruk	Buruk	Terlaksana
1	Kaharuddin	Buruk	Baik	Buruk	Buruk	Buruk	Tidak terlaksana
2	Ruppa dg. Nassa	Baik	Baik	Baik	Buruk	Buruk	Terlaksana
3	Muhajir	Baik	Baik	Baik	Buruk	Buruk	Terlaksana
4	Saleng dg. Sutte	Baik	Baik	Baik	Buruk	Buruk	Terlaksana
...
245	Bahrir dg. Nompo	Baik	Baik	Buruk	Baik	Buruk	Terlaksana
246	Rannu dg. Te'ne	Baik	Buruk	Baik	Buruk	Buruk	Terlaksana
247	Jamaluddin dg. Situru	Baik	Baik	Baik	Buruk	Buruk	Terlaksana
248	Rosmawati	Baik	Buruk	Buruk	Baik	Buruk	Terlaksana
249	H. Muh. Arif tontong	Baik	Baik	Buruk	Baik	Buruk	Terlaksana

250 rows x 7 columns

Gambar 1. Contoh file excel dataset

2.2 Naive Bayes

Naive Bayes adalah salah satu algoritma yang diperkenalkan oleh seorang ilmuwan yang berasal dari Inggris yaitu Thomas Bayes yang dapat digunakan untuk mengklasifikasikan suatu variabel tertentu dengan menggunakan metode probabilitas dan statistik [13]. Algoritma ini juga sering kali digunakan untuk hal seperti penghilang *spam*, sistem rekomendasi, dan teks klasifikasi. Oleh karena itu *Naive Bayes* sering digunakan dalam melakukan klasifikasi sebab memiliki algoritma yang cukup sederhana yaitu dengan menghitung peluang kondisi tertentu [14]. Adapun bentuk umum *teorema bayes* yaitu sebagai berikut [15].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

Keterangan:

$P(A|B)$: Probabilitas A apabila B benar

$P(B|A)$: Probabilitas B apabila A benar

$P(A)$: Probabilitas A

$P(B)$: Probabilitas B

Adapun tahapan dalam menggunakan metode *Naive Bayes* yaitu sebagai berikut [16].

- 1) Mengumpulkan dan membaca data *training*.
- 2) Menghitung jumlah dan probabilitas kelas.
- 3) Dapatkan nilai probabilitas data tiap kelas pada kategori yang sama.
- 4) Menghitung probabilitas akhir setiap kelas yaitu dengan mengalikan probabilitas akhir fitur berdasarkan kelas yang sama.

$$P(X|Kelas) = P(V_1|Kelas) \times P(V_2|Kelas) \times \dots \times P(V_n|Kelas) \quad (2)$$

Keterangan:

$P(X|Kelas)$: Probabilitas x berdasarkan kelas

$P(V|Kelas)$: Probabilitas v berdasarkan kelas

- 5) Menghitung probabilitas akhir ke dalam rumus *Naive Bayes*

$$P(Kelas|X) = P(Kelas) \times P(X) \quad (3)$$

Keterangan:

$P(Kelas|X)$: Probabilitas akhir kelas

$P(Kelas)$: Probabilitas kelas

$P(X)$: Probabilitas X

- 6) Melakukan normalisasi menggunakan rumus berikut ini

$$P(Kelas) = \frac{P(Kelas|X)}{(P(X|Kelas) + P(X|Kelas))} \quad (4)$$

Keterangan:

$P(Kelas|X)$: Probabilitas akhir kelas

$P(X|Kelas)$: Probabilitas akhir seluruh kategori

- 7) Melakukan pengujian performa klasifikasi berdasarkan seluruh data menggunakan rumus *confusion matrix* berikut ini [17].

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

$$Precision = \frac{TP}{FP + TP} \quad (6)$$

$$Recall = \frac{TP}{FN + TP} \quad (7)$$

Keterangan:

TP (*True Positive*) : Data positif yang terklasifikasi positif

TN (*True Negative*) : Data negatif yang terklasifikasi positif

FP (*False Positive*) : Data positif yang terklasifikasi negatif

FN (*False Negative*) : Data negatif yang terklasifikasi negatif

2.3 Laplace Smoothing

Laplace Smoothing merupakan salah satu teknik yang dapat digunakan untuk mengatasi nol frekuensi pada suatu data. Teknik ini biasa digunakan pada proses klasifikasi *Naive Bayes* apabila terdapat probabilitas nol pada suatu fitur. Adapun rumus *laplace smoothing* yaitu berikut ini [18].

$$P(X_k|C_i) = \frac{\sum C_k|C_i + K}{\sum C_i + K|X|} \quad (8)$$

Keterangan:

$P(X_k|C_i)$: Probabilitas fitur X_k dengan label pada kelas C_i

$\sum X_k|C_i$: Jumlah data fitur X_k dengan label pada kelas C_i

$\sum C_i$: Jumlah data dengan label pada kelas C_i

K : Parameter *smoothing*, $K=1$

$|X|$: Jumlah kelas pada sampel

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Dalam melakukan klasifikasi menggunakan *Naive Bayes* dengan bantuan pemrograman *python* tentunya melalui beberapa langkah-langkah yang diperlukan agar sistem dapat melakukan proses klasifikasi data dengan baik dan sesuai dengan model yang digunakan. Adapun langkah-langkah dalam melakukan klasifikasi menggunakan *Naive Bayes* dengan bantuan pemrograman *python* yaitu sebagai berikut:

- 1) Persiapan *Library Python* di *Google Collaboratory*
 Dalam Bahasa pemrograman, *library* merupakan kumpulan modul, fungsi dan metode yang dapat

digunakan untuk memperluas fungsionalitas program oleh para pengembang perangkat lunak. Adapun *library* yang dipersiapkan pada *Google Collaboratory* dapat kita lihat pada gambar 2 berikut.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import OrdinalEncoder
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import CategoricalNB
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

Gambar 2. *Library* pada *Google Collaboratory*

2) Persiapan Data

Proses selanjutnya yaitu mempersiapkan data pemenuhan pilar sanitasi yang telah disimpan pada *file excel* dengan nama “nama desa.xls” yang telah dilakukan proses upload ke *Google Drive* terlebih dahulu. Adapun data yang telah disiapkan berjumlah 250 sampel data disetiap satu file desa dikarenakan pada penelitian ini ingin dilakukan perbandingan pemenuhan pilar sanitasi untuk 7 desa yang memiliki jumlah populasi yang berbeda-beda sehingga apabila memutuskan jumlah sampel menggunakan suatu rumus tentunya akan menghasilkan ketidakseimbangan jumlah sampel yang akan diuji. Berdasarkan hal tersebut, dari 7 desa terdapat desa yang memiliki jumlah populasi sebanyak 260 dibandingkan yang lainnya lebih dari 300 populasi. Dengan demikian, keputusan jumlah sampel dilakukan dengan pertimbangan jumlah populasi desa yang jumlah populasinya kecil sehingga kita putuskan masing-masing sampel tiap desanya sebanyak 250 sampel.

3) Proses *Ordinal Encoder*

Proses selanjutnya yaitu melakukan *ordinal encoder* pada data pilar sanitasi setiap kolom yang bertujuan agar data dapat diproses lebih lanjut oleh model klasifikasi. Adapun data pada awalnya berbentuk kategorik yaitu “Baik” dan “Buruk” yang kemudian diubah oleh *ordinal encoder* seperti yang dapat kita lihat pada gambar 3 berikut ini.

```
[ ] dataset
```

	Nama Warga	Pilar 1	Pilar 2	Pilar 3	Pilar 4	Pilar 5	Keterangan
0	Herman dg Tika	1	1	1	0	0	Terlaksana
1	Kaharuddin	0	1	0	0	0	Tidak terlaksana
2	Ruppa dg Nassa	1	1	1	0	0	Terlaksana
3	Muhajir	1	1	1	0	0	Terlaksana
4	Saleng dg. Sutte	1	1	1	0	0	Terlaksana
...
245	Bahrir dg Nompo	1	1	0	1	0	Terlaksana
246	Rannu dg. Te'ne	1	0	1	0	0	Terlaksana
247	Jamaluddin dg. Situru	1	1	1	0	0	Terlaksana
248	Rosmawati	1	0	0	1	0	Terlaksana
249	H. Muh. Arif tontong	1	1	0	1	0	Terlaksana

250 rows x 7 columns

Gambar 3. Hasil pengubahan data menjadi numerik

4) Penentuan Nilai X Dan Y Pada Setiap Kolom

Pada tahap ini dilakukan penentuan baris dan kolom yang ingin diambil untuk perhitungan pada sistem nantinya. Dimana dalam pemodelan dapat kita kita ketahui bahwa variabel x = variabel untuk

menyimpan data fitur dalam sebuah *dataset*, sedangkan y = variabel untuk menyimpan data output seperti yang dapat kita lihat pada gambar 4 berikut ini.

```
[ ] x = dataset.iloc[:, [1, 2, 3, 4, 5]].values
y = dataset.iloc[:, 6].values
```

Gambar 4. *Source Code* penentuan nilai x dan y

Berdasarkan *source code* diatas maka ditentukan lokasi data yang nilainya akan disimpan pada variabel x yaitu terletak pada kolom 1, 2, 3, 4, dan 5 yang merupakan pilar 1, pilar 2, pilar 3, pilar 4, dan pilar 5, sedangkan nilai variabel y yaitu terletak pada kolom 6 yang merupakan keterangan.

5) *Split Dataset*

Adapun proses selanjutnya yaitu melakukan *split dataset* yang bertujuan untuk membagi *dataset* sehingga 250 data terbagi 2 menjadi data *training* dan data *testing* dengan ukuran 70:30.

6) Proses Klasifikasi Menggunakan *Naive Bayes*

Dalam proses klasifikasi diawali dengan melakukan perhitungan probabilitas kelas serta probabilitas fitur pada data *training*. Adapun data *training* yang digunakan sebagai contoh merupakan data *training* yang terdapat pada *file excel* desa Bontomangape seperti pada gambar 5 berikut ini.

No	Nama	Pilar 1	Pilar 2	Pilar 3	Pilar 4	Pilar 5	Keterangan
1	Warga 1	1	1	1	0	0	Terlaksana
2	Warga 2	0	1	0	0	0	Tidak terlaksana
3	Warga 3	1	1	1	0	0	Terlaksana
4	Warga 4	1	1	1	0	0	Terlaksana
5	Warga 5	1	1	1	0	0	Terlaksana
..
88	Warga 88	1	1	1	1	0	Terlaksana
89	Warga 89	1	1	1	1	0	Terlaksana
90	Warga 90	1	0	0	0	0	Tidak terlaksana
91	Warga 91	0	0	0	0	0	Tidak terlaksana
92	Warga 92	1	0	0	0	0	Tidak terlaksana
93	Warga 93	0	0	0	0	0	Tidak terlaksana
..
171	Warga 171	1	1	1	1	1	Terlaksana
172	Warga 172	1	1	1	0	0	Terlaksana
173	Warga 173	1	1	1	0	0	Terlaksana
174	Warga 174	1	1	1	0	0	Terlaksana
175	Warga 175	1	1	1	0	0	Terlaksana

Gambar 5. Data *training file* desa Bontomangape

Tabel 1. Perhitungan probabilitas kelas

Kelas	Jumlah kejadian	Probabilitas
Terlaksana	137/175	0,782857143
Tidak terlaksana	38/175	0,217142857

Tabel 2. Perhitungan probabilitas pada data *training*

	Fitur	Jumlah Kejadian	
		Terlaksana	Tidak terlaksana
Pilar 1	Baik	126/139	27/40
	Buruk	13/139	13/40
Pilar 2	Baik	111/139	12/40
	Buruk	28/139	28/40
Pilar 3	Baik	126/139	1/40
	Buruk	13/139	39/40
Pilar 4	Baik	70/139	1/40
	Buruk	69/139	39/40
Pilar 5	Baik	36/139	2/40
	Buruk	103/139	38/40

Berdasarkan hasil perhitungan pada tabel 2, diatas dapat kita lihat bahwa *laplace smoothing* akan digunakan ketika terdapat suatu fitur yang memiliki nilai probabilitas 0 misalnya pada fitur "Buruk" pada pilar 3 dan 4. Ketika terdapat probabilitas 0 pada suatu fitur maka otomatis akan dilakukan *laplace smoothing* dengan nilai parameter *laplace* sebesar 1 pada seluruh fitur. Adapun cara *laplace smoothing* pada dapat kita lihat berikut ini.

$$P(X_k|C_i) = \frac{0 + 1}{38 + 1 * (2)} = \frac{1}{40}$$

Tahap selanjutnya yaitu melakukan klasifikasi menggunakan rumus akhir *Naive Bayes* berdasarkan kategori yang sama pada keseluruhan data *testing*. Misalkan pada data *testing* terdapat seorang warga desa dengan pemenuhan pilar sanitasi yaitu pilar 1 = "Baik", pilar 2 = "Baik", pilar 3 = "Buruk", pilar 4 = "Baik", dan pilar 5 = "Buruk". Berdasarkan data *testing* tersebut dapat kita lakukan klasifikasi dengan cara berikut ini.

a. Kelas Terlaksana

$$P(X | \text{Terlaksana}) = P(\text{Baik} | \text{Terlaksana}) \times P(\text{Baik} | \text{Terlaksana}) \times P(\text{Buruk} | \text{Terlaksana}) \times P(\text{Baik} | \text{Terlaksana}) \times P(\text{Buruk} | \text{Terlaksana})$$

$$P(X | \text{Terlaksana}) = 0,782 \times 0,906 \times 0,798 \times 0,093 \times 0,5035 \times 0,741 = 0,019$$

Dengan demikian, diperoleh probabilitas akhir untuk kelas terlaksana yaitu sebesar 0,019.

b. Kelas Tidak terlaksana

$$P(X | \text{Tidakterlaksana}) = P(\text{Baik} | \text{Tidakterlaksana}) \times P(\text{Baik} | \text{Tidakterlaksana}) \times P(\text{Buruk} | \text{Tidakterlaksana}) \times P(\text{Baik} | \text{Tidakterlaksana}) \times P(\text{Buruk} | \text{Tidakterlaksana})$$

$$P(X | \text{Tidakterlaksana}) = 0,217 \times 0,675 \times 0,3 \times 0,975 \times 0,25 \times 0,95 = 0,001$$

Dengan demikian, diperoleh probabilitas akhir kelas tidak terlaksana yaitu sebesar 0,001.

Ketika probabilitas akhir telah didapatkan maka proses terakhir yang akan dilakukan

selanjutnya yaitu melakukan normalisasi dengan cara berikut ini.

$$P(\text{Terlaksana}) = \frac{0,019}{0,019 + 0,001}$$

$$P(\text{Terlaksana}) = 9,5 \times 10^{-1}$$

$$P(\text{Tidak Terlaksana}) = \frac{0,001}{0,001 + 0,019}$$

$$P(\text{Tidak Terlaksana}) = 5 \times 10^{-2}$$

Berdasarkan perhitungan diatas maka dapat disimpulkan bahwa data *testing* tersebut merupakan kelas terlaksana karena memiliki probabilitas terlaksana lebih tinggi dibandingkan probabilitas tidak terlaksana. Tahapan ini juga dapat digunakan pada data *testing* lainnya sehingga keseluruhan data *testing* diketahui hasil klasifikasinya.

7) Proses Perhitungan Akurasi Menggunakan *Confusion Matrix*

Tahap akhir dalam perhitungan ini yaitu melakukan perhitungan akurasi, presisi serta *recall* menggunakan *confusion matrix* yang bertujuan untuk melihat seberapa baik model dalam melakukan klasifikasi. Adapun contoh perhitungan *confusion matrix* dalam sistem dapat kita lihat pada gambar 11 ini.

```
[ ] cm = confusion_matrix(y_test,y_pred)
print (cm)

[[56  1]
 [ 0 18]]

[ ] accuracy = (cm[0][0] + cm[1][1]) / (cm[0][0] + cm[0][1] + cm[1][0] + cm[1][1])
print("Accuracy: {:.2%}".format(accuracy))

Accuracy: 98.67%

[ ] precision = cm[0][0] / (cm[0][1] + cm[0][0])
print("Precision : {:.2%}".format(precision))

Precision : 98.25%

[ ] recall = cm[0][0] / (cm[0][0] + cm[1][0])
print("Recall: {:.2%}".format(recall))

Recall: 100.00%
```

Gambar 6. Hasil pengujian berdasarkan *confusion matrix*

1) Pengujian Menggunakan *Naive Bayes*

Pengujian klasifikasi pada penelitian ini dilakukan menggunakan sistem dengan bahasa pemrograman *python* yang telah dibuat pada *google collaboratory*. Adapun pengujian dilakukan pada 7 *file* desa yang berbeda-beda secara bertahap. Berikut ini hasil klasifikasi masing-masing desa dengan jumlah data *training* sebanyak 175 dan data *testing* sebanyak 75 data untuk masing-masing 1 *file* desa yang dapat kita lihat pada tabel 3.

Tabel 3. Hasil pengujian klasifikasi 7 desa berbeda

Desa	Data Aktual		Data Hasil Klasifikasi	
	Terlaksana	Tidak terlaksana	Terlaksana	Tidak terlaksana
Bontomangape	57	18	54	19
Campagaya	73	2	70	5
Kalenna	64	11	60	15
Parambambe	48	27	45	30

Parangmata	56	19	52	23
Parasangangberu	68	7	64	11
Pattinoang	59	16	57	18
Total	425	100	403	122

Berdasarkan hasil klasifikasi pemenuhan pilar sanitasi 7 desa pada tabel 3 maka diperoleh desa yang memiliki warga dengan pilar sanitasi yang terklasifikasi tidak terlaksana terbanyak. Berdasarkan hal tersebut dapat kita urutkan saran untuk desa mana yang perlu diprioritaskan untuk fokus pemecuan dan pemberian edukasi pilar sanitasi oleh petugas sanitasi puskesmas. Adapun prioritas desa berdasarkan hasil analisis yaitu pada tabel 4.

Tabel 4. Prioritas desa berdasarkan hasil klasifikasi

Desa	Urutan Prioritas
Parambambe	1
Parangmata	2
Bontomangape	3
Pattinoang	4
Kalenna	5
Parasangangberu	6
Campagaya	7

2) Evaluasi Model Menggunakan Confusion Matrix

Berdasarkan hasil pengujian klasifikasi pemenuhan pilar sanitasi warga pada 7 desa area lingkup Puskesmas Bontomangape maka didapatkan nilai *confusion matrix* beserta nilai akurasi, presisi dan *recall* yang dapat kita lihat pada tabel 5 dan tabel 6 berikut ini.

Tabel 5. Perhitungan *confusion matrix*

Desa	TP	FP	FN	TN
Bontomangape	56	1	0	18
Campagaya	70	3	0	2
Kalenna	60	4	0	11
Parambambe	45	3	0	27
Parangmata	52	4	0	19
Parasangangberu	64	4	0	7
Pattinoang	57	2	0	16

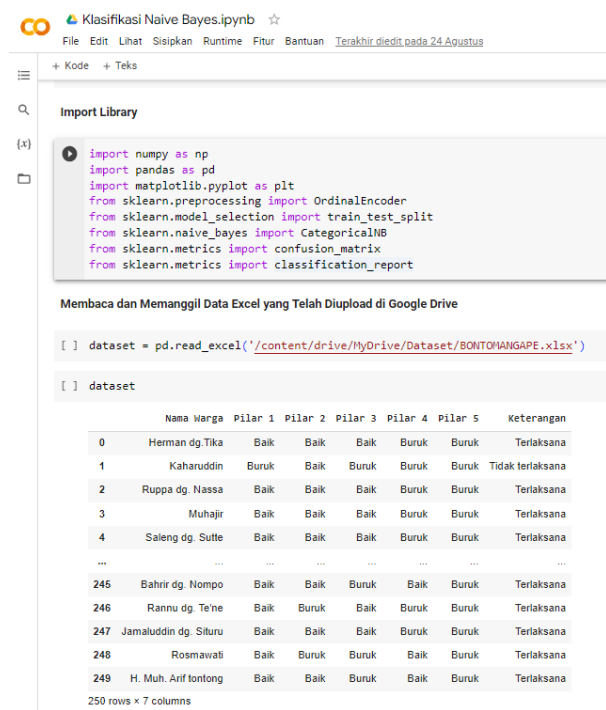
Tabel 6. Hasil evaluasi klasifikasi pemenuhan pilar sanitasi 7 desa area lingkup Puskesmas Bontomangape

Desa	Akurasi	Presisi	Recall
Bontomangape	98,67%	98,25%	100%
Campagaya	96%	95,89%	100%
Kalenna	94,67%	93,75%	100%
Parambambe	96%	93,75%	100%
Parangmata	94,67%	92,86%	100%
Parasangangberu	94,67%	94,12%	100%
Pattinoang	97,33%	96,61%	100%
Rata-rata	95,81%	94,78%	100%

Berdasarkan Tabel 6 dapat diambil beberapa kesimpulan sebagai berikut.

- Penggunaan metode *Naive Bayes* dalam melakukan klasifikasi dapat dikatakan sangat baik sebagaimana tingkat akurasi yang diperoleh setiap pengujian cukup tinggi yaitu diatas 90%.
- Pengujian metode *Naive Bayes* dalam melakukan klasifikasi memperoleh tingkat akurasi, presisi dan recall yang sangat baik pada pengujian untuk desa Bontomangape.
- Metode *Naive Bayes* juga mudah melakukan klasifikasi data dengan secara tepat apabila fitur pada *dataset* kurang bervariasi sehingga nantinya dapat memperoleh nilai akurasi yang tinggi.

Selain itu, dapat juga kita lihat *interface* pada *Google Collaboratory* yang terdapat pada beberapa gambar berikut ini.



Gambar 7. Tampilan membaca dan memanggil data pada *google collaboratory*

Pada gambar 12 dapat kita lihat proses pembacaan dan pemanggilan dataset yang telah diletakkan pada *Google Drive*.

```
[ ] clf=CategoricalNB(alpha=1)
clf.fit(x_train,y_train)

CategoricalNB
CategoricalNB(alpha=1)

Menghitung dan Menampilkan Probabilitas Kelas
• Dimana karena menggunakan perulangan sehingga kelas terlaksana diberi label 0 dan tidak terlaksana diberi label 1

[ ] class_probabilities = np.exp(clf.class_log_prior_)
for i, class_prob in enumerate(class_probabilities):
    print("Probabilitas kelas", i, ":", class_prob)

Probabilitas kelas 0 : 0.7028571428571427
Probabilitas kelas 1 : 0.21714285714285703

Menghitung dan Menampilkan probabilitas fitur

[ ] feature_probabilities = np.exp(clf.feature_log_prob_)
for i, feature_probs in enumerate(feature_probabilities):
    print("Probabilitas fitur untuk kelas", i)
    for j, feature_prob in enumerate(feature_probs):
        print("Fitur", j, ":", feature_prob)
    print()

Probabilitas fitur untuk kelas 0
Fitur 0 : [0.09352518 0.90647482]
Fitur 1 : [0.325 0.675]

Probabilitas fitur untuk kelas 1
Fitur 0 : [0.20143885 0.79856115]
Fitur 1 : [0.7 0.3]

Probabilitas fitur untuk kelas 2
Fitur 0 : [0.09352518 0.90647482]
Fitur 1 : [0.975 0.025]

Probabilitas fitur untuk kelas 3
Fitur 0 : [0.40640228 0.59359772]
Fitur 1 : [0.975 0.025]

Probabilitas fitur untuk kelas 4
Fitur 0 : [0.74190719 0.25809281]
Fitur 1 : [0.95 0.05]
```

Gambar 8. Tampilan proses perhitungan probabilitas kelas dan probabilitas fitur

Pada gambar 13 dapat kita lihat proses perhitungan probabilitas kelas dan fitur dilakukan terlebih dahulu sehingga nantinya dapat digunakan dalam melakukan klasifikasi pada data *testing*.

4. KESIMPULAN

Berdasarkan hasil pengujian model klasifikasi menggunakan *Naive Bayes* dengan sampel data *training* sebanyak 175 dan data *testing* sebanyak 75 sampel data dalam tiap satu *file* excel untuk tiap desa menghasilkan hasil klasifikasi yang diperoleh disetiap pengujian yaitu 54 terlaksana dan 19 tidak terlaksana untuk desa Bontomangape, 70 terlaksana dan 5 tidak terlaksana untuk desa Campagaya, 60 terlaksana dan 15 tidak terlaksana untuk desa Kalenna, 45 terlaksana dan 30 tidak terlaksana untuk desa Parambambe, 52 terlaksana dan 23 tidak terlaksana untuk desa Parangmata, 64 terlaksana dan 11 tidak terlaksana untuk desa Parasangangberu, serta 57 terlaksana dan 18 tidak terlaksana untuk desa Pattinoang. Dengan kata lain, total warga terlaksana sebanyak 403 dan total warga tidak terlaksana 121 dari 525 data *testing*. Adapun hasil klasifikasi yang telah dilakukan memperoleh rata-rata nilai akurasi sebesar 96,01%, nilai presisi sebesar 95,03% dan nilai *recall* sebesar 100%. Berdasarkan hal tersebut penggunaan algoritma *Naive Bayes* untuk melakukan analisis klasifikasi pemenuhan pilar sanitasi warga dapat dikatakan sangat baik, dilihat dari tingkat akurasi yang diperoleh dalam penelitian ini cukup tinggi maka dapat disimpulkan bahwa model dapat mempelajari data *training* dengan baik sehingga memberikan pengaruh pada tingkat akurasi yang dihasilkan. Berdasarkan hasil klasifikasi data pada pengujian, maka hasil tersebut dapat dijadikan tolak ukur oleh petugas sanitasi untuk menentukan desa mana yang perlu diprioritaskan

dalam hal pemberian edukasi dan pemecuan serta bantuan dalam pemenuhan pilar sanitasi sehingga pemenuhan pilar sanitasi warga tiap desa dapat merata untuk seluruh desa area lingkup Puskesmas Bontomangape. Berdasarkan hasil klasifikasi maka desa yang perlu diprioritaskan secara berurut yaitu desa Parambambe, desa Parangmata, desa Bontomangape, desa Pattinoang, desa Kalenna, desa Parasangangberu, dan yang terakhir adalah desa Campagaya.

DAFTAR PUSTAKA

- [1] S. Manton, M. Kanan, H. Herawati, F. S. Lanyumba, and S. N. Sakati, "Gambaran Pencapaian Lima Pilar Sanitasi Total Berbasis Masyarakat di Desa Lokotoy Kabupaten Banggai Laut," *Jurnal Kesmas Untika Luwuk : Public Health Journal*, vol. 12, no. 2, 2021, doi: 10.51888/phj.v12i2.67.
- [2] I. Maliga, R. Rafi'ah, H. Hasifah, and N. A. Sholihah, "Penyuluhan Sanitasi Total Berbasis Masyarakat sebagai Upaya Pencegahan Penyakit Diare di Dusun Batu Bangka Kecamatan Moyo Hilir Kabupaten Sumbawa," *Jurnal Abdidas*, vol. 3, no. 1, pp. 1-9, Feb. 2022, doi: 10.31004/abdidas.v3i1.519.
- [3] R. Hasibuan, M. Nanda, and R. Artikel, "Pemicuan Sanitasi Total Berbasis Masyarakat (STBM) Pilar BABS di Kota Sibolga," *Jurnal Pengabdian Kesehatan Masyarakat*, vol. 1, no. 1, pp. 1-7, 2021.
- [4] A. M. Fahham, "Sanitasi dan Dampaknya bagi Kesehatan: Studi dari Pesantren," *Aspirasi: Jurnal Masalah-masalah Sosial*, vol. 10, no. 1, 2019, doi: 10.46807/aspirasi.v10i1.1230.
- [5] W. Wulandari, D. A. Kartikasari, and L. P. Ratri, "Peran Petugas Kesehatan Terhadap Sanitasi Rumah di Kelurahan Semanggi Kota Surakarta," *Jurnal Kesehatan*, vol. 12, no. 1, 2019, doi: 10.23917/jk.v12i1.8939.
- [6] N. Widiastuti, A. Hermawan, and D. Avianto, "Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Minat Pencari Kerja," *Jurnal Teknoinfo*, vol. 17, no. 1, pp. 1-9, 2023.
- [7] I. Loelianto, Moh. S. S. Thayf, and H. Angriani, "Implementasi Teori *Naive Bayes* Dalam Klasifikasi Calon Mahasiswa Baru Stmik Kharisma Makassar," *Jurnal Science And Information Technology (SINTECH)*, vol. 3, no. 2, pp. 1-8, 2020.
- [8] V. Regista, A. Fitria, and M. A. Kadir, "Gambaran Pelaksanaan 5 Pilar Sanitasi Total Berbasis Masyarakat di Kelurahan Kasturian Kota Ternate," *Svasta Harena: Jurnal Ilmiah Gizi*, vol. 1, no. 2, 2021, doi: 10.33860/shjig.v2i1.478.

- [9] A. Zainal Macfud, A. Pandu Kusuma, and W. Dwi Puspitasari, "Analisis Algoritma *Naive Bayes Classifier* (NBC) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell," *Jurnal Mahasiswa Teknik Informatika (JATI)*, vol. 7, no. 1, pp. 1-8, 2023.
- [10] M. S. Syarah, M. Wati, and N. Puspitasari, "Klasifikasi Penderita ISPA Menggunakan Metode *Naive Bayes Classifier*," *Innovation in Research of Informatics (INNOVATICS)*, vol. 4, no. 1, pp. 1-8, 2022.
- [11] H. F. Putro, R. T. Vlandari, and W. L. Y. Saptomo, "Penerapan Metode *Naive Bayes* Untuk Klasifikasi Pelanggan," *Jurnal Teknologi Informasi dan Komunikasi (TIKOMSiN)*, vol. 8, no. 2, Oct. 2020, doi: 10.30646/tikomsin.v8i2.500.
- [12] E. Fitriani, A. Saepudin, D. Ardiansyah, and R. Aryanti, "Implementasi Metode *Naive Bayes* Dalam Penyeleksian Karyawan Untuk Penempatan Bagian Pemasaran," *Jurnal Teknik Komputer AMIK BSI*, vol. 8, no. 2, pp. 1-7, 2022, doi: 10.31294/jtk.v4i2.
- [13] Apif Supriadi and Fatmasari, "Implementasi Metode Klasifikasi *Naive Bayes* Pada Sistem Analisis Opini Pengguna *Twitter* Berbasis Web," *Jurnal Sistem Informasi*, vol. 10, no. 1, pp. 1-9, 2021, doi: 10.51998/jsi.v10i1.356.
- [14] L. Syafie, F. Umar, A. Mude, H. Darwis, Herman, and Harlinda, "Missing Data Handling Using The *Naive Bayes Logarithm* (NBL) Formula," *IEEE Xplore*, pp. 1-4, 2018.
- [15] Herman et al., "Comparison of Artificial Neural Network and Gaussian *Naive Bayes* in Recognition of Hand-Writing Number," *IEEE Xplore*, pp. 1-4, 2018.
- [16] T. Imandasari, E. Irawan, A. P. Windarto, and A. Wanto, "Algoritma *Naive Bayes* Dalam Klasifikasi Lokasi Pembangunan Sumber Air," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 1-12, 2019, doi: 10.30645/senaris.v1i0.81.
- [17] I. Widhi Saputro and B. Wulan Sari, "Uji Performa Algoritma *Naive Bayes* untuk Prediksi Masa Studi Mahasiswa," *Citec Journal*, vol. 6, no. 1, 2019.
- [18] D. Utami and P. Aisyiyah Rakhma Devi, "Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan (PKH) Menggunakan Metode *Weighted Naive Bayes* Dengan *Laplace Smoothing*," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 4, pp. 1-12, 2022.