

## Prediksi Tingkat Kesehatan Lingkungan Masyarakat Dalam Program Sustainable Development Goals Menggunakan Algoritma Naive Bayes.

Zulkipli<sup>1\*</sup>, Kusrini<sup>2</sup>, Sudarmawan<sup>3</sup>

<sup>1,2,3</sup> Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta

\*zulkipli@students.amikom.ac.id

### Abstrak

Indikasi penurunan tingkat kesadaran masyarakat tentang menjaga lingkungan berpengaruh langsung terhadap kebutuhan mendesak dalam perencanaan dan pengelolaan lingkungan secara terpadu, sehingga dampaknya berpengaruh pada aspek lainnya, seperti lingkungan fisik dan sosial ekonomi. Faktanya rusaknya lingkungan berkaitan erat dengan masalah kemiskinan dan tingkat pertumbuhan ekonomi. Dalam menjaga kesehatan lingkungan masyarakat dan menerapkan Program Sustainable Development Goals di Kabupaten Lombok Timur Nusa Tenggara Barat. Universitas Hamzanwadi bekerja sama dengan pemerintah Kabupaten Lombok Timur untuk melakukan sampling kepada masyarakat dengan jumlah data 4624 penduduk di sepuluh kecamatan di wilayah Kabupaten Lombok Timur. Tujuan penelitian ini adalah Untuk membantu memprediksi tingkat kesehatan lingkungan masyarakat di wilayah Lombok Timur, data yang dimiliki akan diklasifikasikan lalu diolah dengan algoritma Naive Bayes dengan metode multinomial naive bayes. Hasil pengujian algoritma naive bayes setelah dilakukan split data sebanyak lima kali pengujian di dapatkan hasil terbaik dengan dataset dibagi data testing 20 % dan data training 80 % di dapatkan nilai prediksi sebesar 93.28% Lingkungan penduduk di sepuluh kecamatan di Kabupaten Lombok Timur tergolong sehat.

**Kata kunci :** Algoritma Naive Bayes, Lingkungan Sehat, SDGs

### Abstract

Indications of a decrease in the level of public awareness about protecting the environment have a direct effect on the urgent need for integrated environmental planning and management, so that the impact affects other aspects, such as the physical and socio-economic environment. The fact is that environmental damage is closely related to poverty and economic growth. In maintaining community environmental health and implementing the Sustainable Development Goals Program in East Lombok Regency, West Nusa Tenggara. Hamzanwadi University is working with the East Lombok Regency government to conduct sampling of the community with a total data of 4624 residents in ten sub-districts in the East Lombok Regency area. The purpose of this study is to help predict the level of public environmental health in the East Lombok region, the data will be classified and then processed using the Naive Bayes algorithm with the multinomial naive Bayes method. The results of testing the naive Bayes algorithm after splitting the data five times, the best results were obtained, with the dataset divided by 20% testing data and 80% training data, an prediction value of 93.28% was obtained. The population environment in ten sub-districts in East Lombok Regency was classified as healthy

**Keywords :** Naive Bayes Algorithm, Healthy Environment, SDGs

### 1. Pendahuluan

Menurunnya tingkat kesadaran masyarakat tentang menjaga lingkungan dan rusaknya fungsi

ekosistem menyebabkan masalah yang serius dan ancaman pada kehidupan manusia. Faktanya sejumlah aktivitas manusia memberikan tekanan

terhadap lingkungan, sehingga berpotensi terjadinya degradasi lingkungan dan rusaknya fungsi ekosistem, baik di daratan dan lautan yang berakibat pada kerugian yang sangat besar dalam kehidupan. Indikasi penurunan tingkat kesadaran masyarakat tentang menjaga lingkungan berpengaruh langsung terhadap kebutuhan mendesak dalam perencanaan dan pengelolaan lingkungan secara terpadu, sehingga dampaknya berpengaruh pada aspek lainnya, seperti lingkungan fisik dan sosial ekonomi. Faktanya rusaknya lingkungan berkaitan erat dengan masalah kemiskinan dan tingkat pertumbuhan ekonomi [1].

Dalam menjaga kesehatan lingkungan masyarakat dan menerapkan program sustainable development goals di Kabupaten Lombok Timur Nusa Tenggara Barat. Universitas Hamzanwadi bekerja sama dengan pemerintah kabupaten lombok timur untuk melakukan sampling kepada masyarakat dengan jumlah data 4624 penduduk di sepuluh kecamatan di wilayah kabupaten lombok timur. Untuk membantu memprediksi tingkat kesehatan lingkungan masyarakat di wilayah Lombok Timur, data yang dimiliki akan dianalisa dan diprediksi lalu diolah menggunakan algoritma Naive bayes. Analisa dan prediksi dilakukan agar peneliti mendapatkan nilai akurasi tertinggi dengan mengambil variabel seperti pengelolaan sampah, kondisi mata air, jamban, kandang peliharaan dan galian.

Mengapa Naive bayes? Algoritma Naive bayes Naive bayes sendiri adalah sebuah metode pengklasifikasian data yang digunakan untuk memprediksi probabilitas suatu class, algoritma ini sangat sederhana karena tidak memerlukan banyak data proses pengklasifikasian [2]. Data yang digunakan dalam penelitian ini adalah data private yang di peroleh dari wawancara kepada masyarakat di sepuluh wilayah Kecamatan di Kabupaten Lombok Timur di antaranya Kecamatan Selong, Kecamatan Labuhan Haji, Kecamatan Suralaga, Kecamatan Masbagik, Kecamatan Terara, Kecamatan Aikmel, Kecamatan Lenek, Kecamatan Sukamulia, Kecamatan Sikur dan Kecamatan Pringgasele. Penelitian kali ini menggunakan Naive bayes yang bertujuan untuk memprediksi tingkat Kesehatan lingkungan masyarakat di kabupaten Lombok Timur Nusa Tenggara Barat, dan juga untuk mengetahui performance algoritma naive bayes dalam memprediksi sebuah dataset. Metode ini digunakan untuk menganalisis dan memprediksi data yang hasil akhirnya adalah sehat atau tidak sehat lingkungan di wilayah Kabupaten Lombok Timur.

## 2. Tinjauan Pustaka

### 2.1. Penelitian Terkait

Adapun penelitian sebelumnya yang menggunakan algoritma naive bayes sebagai metode dalam pengolahan data sebagai

pendukung penelitian kali ini, antara lain: Annur penelitian yang berjudul “klasifikasi masyarakat miskin menggunakan metode naïve bayes” Berdasarkan hasil pengujian confusion matrix dengan teknik split validasi, penggunaan metode klasifikasi naïve bayes terhadap dataset yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori Good. Sementara nilai Precision sebesar 92% dan Recall sebesar 86% [3].

Penelitian yang dilakukan oleh Yupi Kuspani dkk. yang berjudul “ Analisis Perbandingan Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Kesejahteraan Masyarakat Pada Desa Suralaga Dengan Menggunakan Algoritma Naive Bayes Dan Support Vector Machine (Svm)” Nilai akurasi yang tertinggi ditunjukkan oleh eksperimen menggunakan K-Vold Validation 8 dan K-Vold Validation 10. Sedangkan toleransi yang diberikan pada K-Vold Validation 8 (0.49%) lebih kecil dari K=Vold Validation 10 sebesar (0.58%). Hal ini berarti K-Vold Validation 8 lebih ketat dari K-Vold Validation 10. Sehingga yang paling baik digunakan dalam pengambilan keputusan adalah K-Vold Validation 8 sebesar 99.62% dengan toleransi 0.49%. Hasil pengolahan data menggunakan algoritma Naive Bayes dan Support Vector Machine, sama-sama memberikan gambaran bahwa pengaruh ekonomi terhadap tingkat kesejahteraan masyarakat Desa Suralaga sangat besar dan dapat diambil

kesimpulan bahwa rata-rata masyarakat Desa Suralaga termasuk dalam kategori masyarakat yang tidak Sejahtera [4].

Muhammad Wasil dan Mahfuz dalam penelitiannya yang berjudul “Analisis Persepsi Masyarakat Terhadap Kinerja Aparatur Desa Dalam Meningkatkan Pelayanan Publik Di Desa Lenek Lauk Menggunakan Metode Naive Bayes” Dalam penelitian ini akan dilakukan prediksi mengenai puas atau tidak puas masyarakat terhadap kinerja aparatur desa dalam meningkatkan pelayanan menggunakan metode algoritma Naïve Bayes. Algoritma Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai teorema bayes. Penelitian ini dilakukan dengan menggunakan teknik pengumpulan data berupa angket yang di isi oleh 50 responden 10% dari jumlah populasi, dari populasi yang berjumlah 5.539 orang penduduk tetap yang ada di desa Lenek Lauk. Analisis hasil penelitian dilakukan dengan Metode algoritma naïve bayes. Berfungsi untuk melihat tingkat akurasinya. Pengukuran dengan algoritma naïve bayes menghasilkan akurasi 92.26% dan nilai AUC pada Kurva ROC adalah sebanyak 0.924 [5]. Ahmadi Irmansyah Lubis dan kawan kawan dalam penelitiannya yang berjudul “Komparasi Akurasi

Pada Naïve Bayes Dan Random Forest Dalam Klasifikasi Penyakit Liver” Berdasarkan hasil pengujian yang dilakukan dalam memperoleh nilai akurasi perhitungan klasifikasi menggunakan Confusion Matrix, maka metode Random Forest memperoleh hasil yang terbaik yaitu dengan peroleh akurasi sebesar 70.60 % bila dibandingkan dengan Naïve Bayes yang hanya memperoleh akurasi sebesar 55.80 %. Sehingga Random Forest memiliki performa kinerja yang lebih unggul dalam perolehan akurasi yang dihasilkan dalam klasifikasi penyakit liver [6]. Penelitian yang dilakukan oleh Dwi Normawati dan Surya Allit Prayogi yang berjudul “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter” Data yang diteliti sebagai studi kasus sederhana hanya menggunakan sebanyak 8 data cuitan yang dibagi menjadi 5 data latih dan 3 data uji. Data tersebut diolah dengan tahap preprocessing, lalu diklasifikasi menggunakan metode NBC, perhitungan performa menggunakan teknik confusion matrix. Penelitian ini menghasilkan pemaparan yang terstruktur pada proses dan hasil implementasi NBC dan pengujian performa menggunakan confusion matrix yang didapatkan akurasi sebesar 82%, presisi 93%, dan recall sebesar 52% [7].

Penelitian yang dilakukan oleh Veronica Retno Sari dkk. dengan judul “perbandingan prediksi

kualitas kopi arabika dengan menggunakan algoritma sgd, random forest dan naïve bayes,” Penelitian ini bertujuan untuk mengetahui algoritma mana yang paling efektif untuk memprediksi kualitas kopi pada dataset. Kemudian hasil prediksi akan diujikan dengan metode *K-Fold Cross Validation* dan *Area Under the Curve* (AUC). Hasil penelitian menunjukkan bahwa algoritma *Stochastic Gradient Descent* mendapatkan hasil akurasi paling baik dibandingkan dua metode lainnya dengan nilai akurasi sebesar 98% dan meningkat menjadi 99% dengan pengujian menggunakan metode *K-fold Cross Validation* dan AUC. [8]

Yunita Sartika Sari dalam penelitiannya yang berjudul “Penerapan Metode Naïve Bayes Untuk Mengetahui Kualitas Air Di Jakarta” Pada penelitian ini menerapkan algoritma Naïve Bayes untuk mengetahui kualitas air di wilayah Jakarta dan menghasilkan tingkat akurasi dari hasil klasifikasi air di wilayah dki jakarta sebesar 50,6% [9].

Penelitian yang di lakukan oleh Yahya dan Hariman Bahtiar berjudul “Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Kesejahteraan Masyarakat Kecamatan Selong Kabupaten Lombok Timur – Nusa Tenggara Barat Menggunakan Algoritma Naive Bayes. Dalam penelitian ini salah satu yang menjadi fokus penelitian adalah komponen ekonomi. Data yang diperoleh di Kecamatan Selong,

khususnya komponen ekonomi akan dikelola dan diolah menggunakan algoritma Naive Bayes. Setelah dilakukan pengolahan data menggunakan algoritma Naive Bayes, diperoleh nilai tingkat akurasi kedekatan dengan keadaan yang sebenarnya sebesar 93,45%. Dari data yang diperoleh 93,45% atau  $0,9345 \times$  jumlah data ( $kk$ ) =  $0,9345 \times 1130$   $kk = 1056$   $kk$  yang menunjukkan masyarakatnya sejahtera dan  $6,55\% \times 1130 = 74$   $kk$  yang menyatakan masyarakat yang tidak sejahtera dan dapat dijadikan sebagai acuan dalam pengentasan kemiskinan melalui program yang sudah dicanangkan oleh pemerintah [10].

## 2.2. Landasan Teori

### 1. Naive bayes

Algoritma Naive Bayes adalah salah satu algoritma yang terdapat pada teknik data mining klasifikasi. Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris yaitu Thomas Bayes, Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya, sehingga dikenal dengan Teorema Bayes. Teorema tersebut dikombinasikan dengan naive dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. [11]

### 2. Klasifikasi

Klasifikasi adalah sebuah metode dari data mining yang biasa digunakan untuk suatu bidang pembelajaran dimana proses Analisa data dilakukan sehingga menciptakan model model untuk menggambarkan kelas-kelas dari data yang sedang dianalisa[12].

Metode klasifikasi yang digunakan yaitu metode Multinomial Naive Bayes. Multinomial Naive Bayes merupakan salah satu metode pada Probabilistic Reasoning. Algoritma Multinomial Naive Bayes bertujuan untuk melakukan klasifikasi data pada kelas tertentu [13].

### 3. Data Mining

Data mining merupakan proses penggunaan teknik statistik, matematika, artificial intelligence (kecerdasan buatan) dan machine learning yang digunakan untuk mengekstrak serta mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [13].

### 4. Confussion Matrix

Confusion Matrix adalah sebuah metode yang digunakan untuk melihat hasil kinerja dari metode klasifikasi, dan biasanya menggunakan 4 istilah dari hasil proses klasifikasi yaitu True Positif (TP), True Negatif (TN), False Positif (FP) dan false Negatif (FN). [14]

Confusion matrix melakukan pengujian untuk memperkirakan obyek yang benar dan salah [13].

Tabel Confusion Matrix dapat di lihat pada Tabel 1.

Tabel 1. Confusion Matrix

		Actual Class	
		False Posittif	True Positif
Predicat Class	False Posittif		
	True Negatif		

Untuk permasalahan dalam Multy classification, kriteria evaluasi yang biasa digunakan adalah prediction, Accuray, recall, dan F1. Score.

Precision

$$Precision = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (TP_i + FP_i)} \times 100\%$$

Recall

$$Recall = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (TP_i + FN_i)} \times 100\%$$

Accuracy

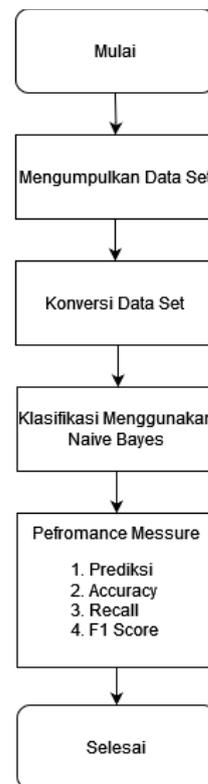
$$Accuracy = \frac{\sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \times 100\%}{l}$$

F<sub>1</sub>

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

## 2.2 Tahapan Penelitian

Adapun tahapan penelitian yang dikerjakan dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

Tahapan pertama dalam penelitian ini adalah

1. Mengumpulkan data set yang berjumlah 4624 penduduk di 10 Kecamatan di Kabupaten Lombok Timur yang diperoleh dari kegiatan KKN tematik desa gemilang yang dilakukan oleh univeritas hamzanwadi.
2. Tahap kedua yaitu konversi data set, dalam hal ini yang dilakukan adalah merubah data yang berjenis string atau text menjadi angka. Hal ini bertujuan supaya program yang digunakan bisa membaca dataset tersebut. [15] Konversi data dilakukan dengan cara merubah karakter menjadi kode ASCII kemudian menjumlahkannya. Hasil penjumlahan

tersebut yang digunakan sebagai nilai dari atribut yang akan digunakan.

3. Tahapan ketiga melakukan klasifikasi menggunakan algoritma naïve bayes untuk melihat prediksi, akurasi, recall, dan F1 Score, Terhadap masing-masing pengujian.

### 3. Metode Penelitian

#### 3.1. Teknik Pengumpulan Data

Pada penelitian ini data diperoleh dari kegiatan KKN Tematik Universitas Hamzanwadi dan termasuk data private, sedangkan teknik pengumpulan data dalam kegiatan tersebut menggunakan teknik observasi, wawancara dan penyebaran angket kepada penduduk di sepuluh Kecamatan di Kabupaten Lombok Timur, Nusa Tenggara Barat.

Dalam pengolahan data perangkat lunak yang digunakan adalah python dengan metode Klasifikasi dan menggunakan algoritma Naïve Bayes dengan metode multinomial naïve bayes.

##### 1. Teknik Observasi

Teknik observasi dilakukan dengan melakukan pengamatan langsung ke lokasi penelitian yaitu di kecamatan Kecamatan Selong, Kecamatan Labuhan Haji, Kecamatan Suralaga, Kecamatan Masbagik, Kecamatan Terara, Kecamatan Aikmel, Kecamatan Lenek, Kecamatan Sukamulia, Kecamatan Sikur dan Kecamatan Pringgasela, di

Kabupaten Lombok Timur, Nusa Tenggara Barat.

##### 2. Teknik Wawancara

Teknik wawancara digunakan untuk mendapatkan data langsung dari penduduk di lokasi penelitian.

##### 3. Angket

Teknik angket ini dilakukan untuk mempermudah saat melakukan wawancara dan juga sebagai alat bukti untuk penginputan data.

#### 3.2 Teknik Analisis Data

Teknik analisis data untuk penelitian kali ini adalah kuantitatif, metode ini digunakan karena sesuai dengan penelitian yang sedang dilakukan. Data yang digunakan adalah data private yang belum pernah digunakan pada penelitian-penelitian sebelumnya. Pada penelitian ini instrument penelitian untuk pengumpulan data adalah observasi langsung ke lokasi, wawancara dan angket.

#### 3.3. Model yang di usulkan

Penelitian ini mengusulkan menggunakan metode klasifikasi dengan algoritma Naïve bayes. Pengujian dilakukan dengan Bahasa pemrograman python dengan metode multinomial naïve bayes dan untuk evaluasi menggunakan confusion matrix dengan hasil akhir nilai akurasi

#### 4. Hasil dan Pembahasan

##### 4.1. Hasil Penelitian

Proses pengolahan data pada penelitian ini menggunakan tools jupyter notebook dengan metode klasifikasi dan menggunakan algoritma naïve bayes untuk menganalisis dan memprediksi tingkat kesehatan lingkungan sepuluh Kecamatan di Kabupaten Lombok Timur NTB. Jumlah data yang terkumpul adalah 4624 data dengan atribut yang digunakan sebanyak 5 (lima) atribut antara lain: Tempat Membuang Sampah, Mempunyai Jamban, Pembuangan Air Rumah Tangga, Kondisi Mata Air dan Penambangan Galian C. Dari lima atribut tersebut dapat di gunakan sebagai komponen penentu keputusan akhir dari kondisi lingkungan masyarakat di Kabupaten Lombok Timur.

Adapun tabel atribut yang digunakan dapat di lihat tabel Tabel 2.

Tabel.2 Atribut yang di gunakan

No.	Atribut
1	Tempat Membuang Sampah
2	Mempunyai Jamban
3	Pembuangan Air Rumah Tangga
4	Kondisi Mata Air
5	Penambangan Galian C

Berikut adalah hasil pengolahan data menggunakan tools jupyter notebook dengan

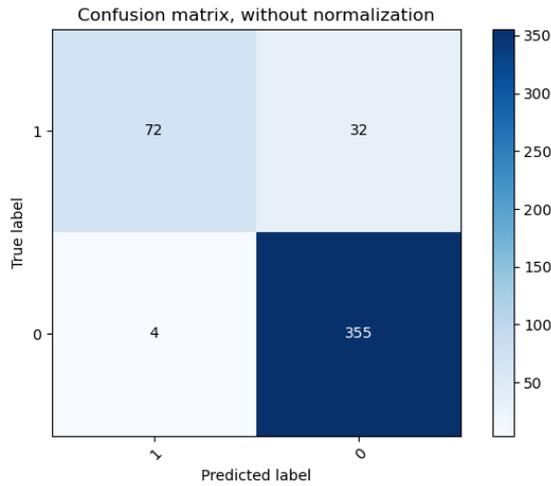
metode klasifikasi dan algoritma naïve bayes dengan metode Multinomial Naïve Bayes.

Tabel 3. Hasil Pengolahan Data Menggunakan Algoritma Naïve Bayes

Hasil Pengolahan Algoritma Naïve Bayes				
Dataset	Presisi	Akurasi	Recall	F1 Score
	%	%	%	%
10:90	91.73	92.22	98.88	95.17
20:80	93.28	93.83	99.31	96.80
30:70	92.12	92.86	99.35	95.60
40:60	91.91	92.75	99.44	95.53
50:50	92.17	93.03	99.49	95.69

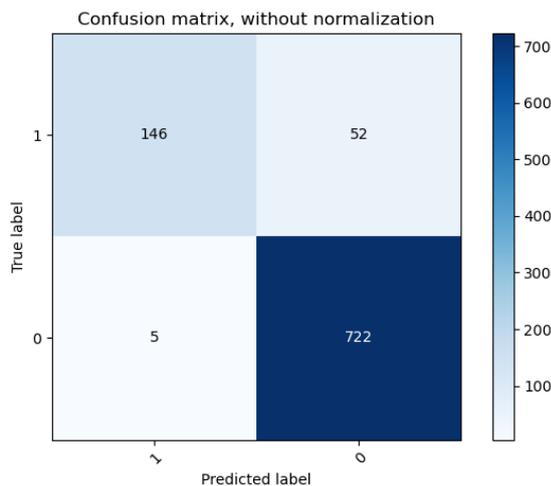
Pada percobaan yang dilakukan sebanyak 5 kali pada dataset nilai presisi tertinggi Ketika dicoba dengan data test 20 % data training 80 dengan presisi 93.28 % sedangkan akurasi tertinggi di dapatkan pada dataset dibagi menjadi 20 data test dan 80 data training. Adapun hasil nilai akuarsinya yaitu 93.83 %. Untuk nilai recall nilai tertinggi di dapatkan ketika nilai dibagi menjadi data test dibagi menjadi 50 dan data training dibagi menjadi 50 dengan nilai recall 99.49 %. Untuk nilai F1 Score di dapatkan nilai tertinggi Ketika data set dibagi menjadi data test 20 dan data training 80 dengan nilai f1 score 96.80 %. Dari pengujian yang dilakukan di atas menggunakan metode multinomial naïve bayes. Dapat di buktikan bahwa algoritma naïve bayes sangat baik untuk mengklasifikasikan tingkat Kesehatan lingkungan di sepuluh Kecamatan di Kabupaten Lombok Timur NTB.

Adapun hasil untuk confusion matrik untuk dataset yang yang di lakukan dengan pembagian data test 10 % dan data training 90 % dapat dilihat pada gambar 2.



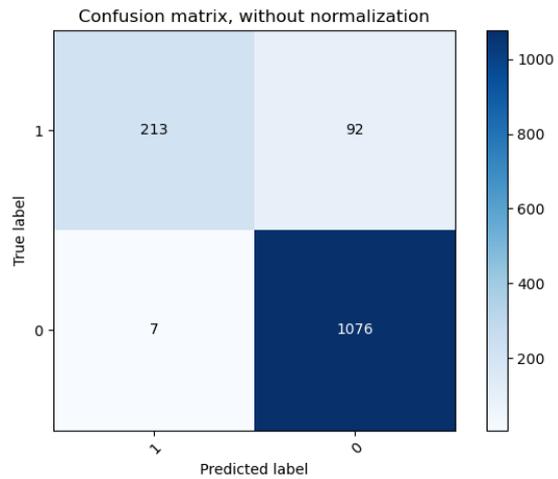
Gambar 2. Confusion Matrik 10:90

Hasil pengujian data yang dibagi data test 20 % data test dan 80 % data training dapat dilihat dalam bentuk confusion matrix pada gambar 3.



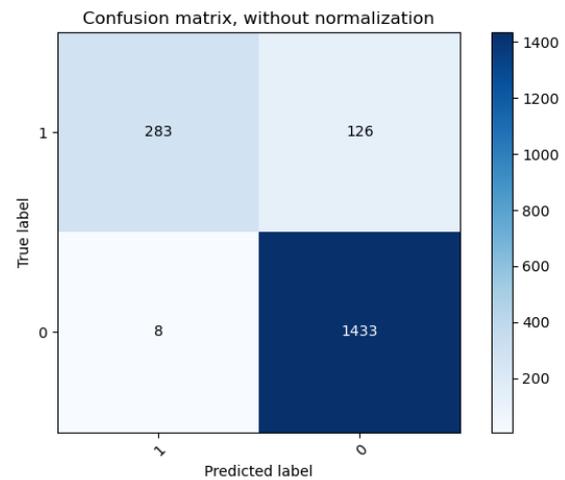
Gambar 3. Confusion Matrix 20:80

Confusion Matrix Selanjutnya dengan pembagian data 30 % data test dan 70 % dilihat pada gambar 4.



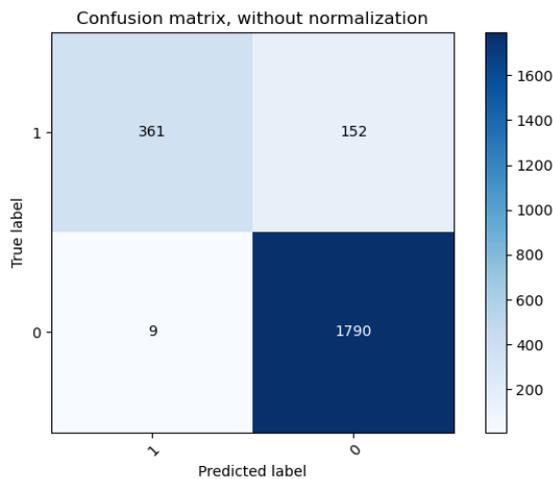
Gambar 4. Confusion Matrik 30:70

Untuk pembagian dataset 40 % data test dan 60 % data training digambarkan dalam bentuk confusion matrik yang dapat dilihat pada gambar 5.



Gambar 5. Confusion Matrik 40:60

Selanjutnya dataset yang dibagi menjadi 50 % data test dan 50 % data training dapat dilihat pada gambar 6.



Gambar 6. Confusion Matrik 50:50

Dari keterangan tabel dan gambar confusion matrix di atas dapat dilihat hasil nilai terbaik dapat di lihat pada pengujian kedua yang pembagian data test 20 % dan data training 80 % dengan nilai akurasi 93.83 %.

#### 4.2. Pembahasan

Hasil pengolahan data yang digambarkan pada tabel dan confusion matrix di atas dapat dijelaskan sebagai berikut :

- Nilai akurasi pengujian pertama dengan data testing 10% data training 90% memberikan nilai akurasi 92.22%. Artinya jumlah Masyarakat yang lingkungannya tergolong bersih adalah  $0.922 \times 4624 = 4264$  KK dan  $0.778 \times 4624 = 360$  KK. Lingkungannya Tergolong tidak bersih.
- Nilai akurasi pengujian kedua dengan data testing 20% dan data training 80% memberikan nilai akurasi 93.83%. Artinya jumlah Masyarakat yang lingkungannya tergolong bersih  $0.9383 \times$

$4624 = 4338$  KK. dan  $0.0617 \times 4624 = 285$  KK. Lingkungannya tidak sehat.

- Nilai akurasi pengujian ketiga dengan data testing 30% dan data training 70% memberikan nilai akurasi 92.86%. Artinya jumlah Masyarakat yang lingkungannya tergolong bersih adalah  $0.9286 \times 4624 = 4294$  KK. dan  $0.0714 \times 4624 = 330$  KK. lingkungannya tidak sehat
- Nilai akurasi pengujian ke empat dengan data testing 40% dan data training 60% memberikan nilai akurasi 92.75%. Artinya jumlah masyarakat yang lingkungannya bersih adalah  $0.9275 \times 4624 = 4289$  KK. dan  $0.725 \times 4624 = 335$  KK. lingkungannya tidak sehat.
- Nilai akurasi pengujian ke lima dengan data testing 50% data testing dan 50% data training memberikan nilai 93.03%. Artinya jumlah Masyarakat yang lingkungannya bersih adalah  $0.9303 \times 4624 = 4301$  KK. dan  $0,0697 \times 4624 = 322$  KK. Lingkungannya tidak sehat.

#### 5. Kesimpulan

Dari hasil penelitian yang telah dilakukan, metode klasifikasi dengan algoritma naïve bayes dalam memprediksi tingkat kesehatan lingkungan di sepuluh Kecamatan di Kabupaten Lombok Timur NTB, dengan lima kali percobaan didapatkan nilai akurasi terbaik dengan pengujian data test 20 % dan 80 % data training mendapatkan nilai akurasi sebesar 93.83% lingkungan Masyarakat tergolong bersih dan 6,17% lingkungan tidak

sehat. Dari nilai akurasi yang diperoleh dapat disimpulkan untuk naïve bayes sangat akurat digunakan untuk menganalisis serta memprediksi tingkat kesehatan lingkungan di Lombok Timur Nusa Tenggara Barat. Sebagai acuan melalui program yang sudah dicanangkan oleh pemerintah

## 6. Daftar Pustaka

- [1] E. Khairina, E. P. Purnomo, and A. D. Malawnai, "Sustainable Development Goals: Kebijakan Berwawasan Lingkungan Guna Menjaga Ketahanan Lingkungan Di Kabupaten Bantul Daerah Istimewa Yogyakarta," *J. Ketahanan Nas.*, vol. 26, no. 2, p. 155, Sep. 2020, doi: 10.22146/jkn.52969.
- [2] M. Ridho Handoko, "SISTEM PAKAR DIAGNOSA PENYAKIT SELAMA KEHAMILAN MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 50–58, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [3] H. Annur, "KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES," 2018.
- [4] Y. Kuspandi Putra, "Analisis Perbandingan Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Kesejahteraan Masyarakat Pada Desa Suralaga Dengan Menggunakan Algoritma Naive Bayes Dan Support Vector Machine (Svm)," *J. Inform. dan Teknol.*, vol. 4, no. 1, 2021, doi: 10.29408/jit.v4i1.2961.
- [5] M. Wasil, "Infotek : Jurnal Informatika dan Teknologi Analisis Persepsi Masyarakat Terhadap Kinerja Aparatur Desa Dalam Meningkatkan Pelayanan Publik Di Desa Lenek Lauk Menggunakan Metode Naive Bayes Pelayanan publik adalah kegiatan atau rangkaian kegiatan dalam ran," vol. 3, no. 2, pp. 140–146, 2020.
- [6] A. I. Lubis, U. Erdiansyah, and R. Siregar, "Komparasi Akurasi Pada Naïve Bayes Dan Random Forest Dalam Klasifikasi Penyakit Liver Comparison of Accuracy in Naïve Bayes and Random Forests in Classification of Liver Disease." [Online]. Available: [www.jurnal.unimed.ac.id](http://www.jurnal.unimed.ac.id)
- [7] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," 2021.
- [8] V. Sari, F. Firdausi, and Y. Azhar, "Perbandingan Prediksi Kualitas Kopi Arabika dengan Menggunakan Algoritma SGD, Random Forest dan Naive Bayes," *Edumatic J. Pendidik. Inform.*, vol. 4, no. 2, pp. 1–9, 2020, doi: 10.29408/edumatic.v4i2.2202.
- [9] Y. S. Sari, "Penerapan Metode Naïve Bayes Untuk Mengetahui Kualitas Air Di Jakarta," *J. Ilm. FIFO*, vol. 13, no. 2, p. 222, 2021, doi: 10.22441/fifo.2021.v13i2.010.
- [10] Y. Yahya and H. Bahtiar, "Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Kesejahteraan Masyarakat Kecamatan Selong Kabupaten Lombok Timur – Nusa Tenggara Barat Menggunakan Algoritma Naive Bayes," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 20–28, 2021, doi: 10.29408/jit.v4i1.2981.
- [11] W. P. Nurmayanti, "Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak," *Geodika J. Kaji. Ilmu dan Pendidik. Geogr.*, vol. 5, no. 1, pp. 123–132, 2021, doi: 10.29408/geodika.v5i1.3430.
- [12] M. Julkarnain and K. R. Ananda, "Sistem Informasi Pengolahan Data Ternak Unit Pelaksana," *Jinteks*, vol. 2, no. 1, pp. 32–39, 2020.
- [13] M. T. H. Bunga, B. S. Djahi, and Y. Y. Nabuasa, "Multinomial Naive Bayes Untuk Klasifikasi Status Kredit Mitra Binaan Di Pt . Angkasa Pura I Program Kemitraan," *J-lcon*, vol. 6, no. 2, pp. 30–34, 2018, [Online]. Available:

- <https://media.neliti.com/media/publications/292504-multinomial-naive-bayes-untuk-klasifikas-2536567f.pdf>
- [14] Karsito and Susanti Santi, "Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah Dengan Algoritma Naive Bayes Di Perumahan Azzura Residence," *J. Teknol. Pelita Bangsa*, vol. 9, pp. 43–48, 2019.
- [15] M. Vishwakarma and N. Kesswani, "A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection," *Decis. Anal. J.*, vol. 7, no. April, p. 100233, 2023, doi: 10.1016/j.dajour.2023.100233.