

Implementasi Algortima *Naïve Bayes* Untuk Klasifikasi Penerima Bantuan Langsung Tunai

Nur Azise ¹, Ahmad Gunawan ², Achmad Baijuri ³
^{1,2,3} Universitas Ibrahimy, Indonesia

Info Artikel	ABSTRAK
<u>Riwayat Artikel</u>	Prosedur pengajuan dan penetapan penerima Bantuan Langsung Tunai (BLT) di Desa Pesanggrahan masih dilakukan secara manual sehingga sering terjadi ketidakupasan sasaran. Kondisi ini menimbulkan potensi ketidakadilan di tengah-tengah masyarakat, sehingga diperlukan pendekatan berbasis teknologi yang dapat meningkatkan objektivitas dan akurasi dalam seleksi penerima bantuan sesuai kriteria yang telah ditetapkan, seperti keluarga miskin, tidak menerima bantuan lain, dan lansia. Penelitian ini mengimplementasikan algoritma klasifikasi <i>Naïve Bayes Gaussian</i> dengan tahapan data <i>preparation</i> , <i>labelling</i> , pembagian data latih dan uji menggunakan rasio 80:20, serta pengolahan atribut dengan <i>TF-IDF Vectorizer</i> . Evaluasi performa dilakukan melalui <i>confusion matrix</i> , <i>classification report</i> dengan hasil akurasi, <i>precision</i> , <i>recall</i> , dan <i>f1-score</i> sebesar 100%. Hasil penelitian ini menunjukkan efektivitas algoritma <i>Naïve Bayes</i> dalam mengklasifikasi penerima Bantuan Langsung Tunai (BLT), dan penentuan penerima BLT dilakukan secara akurat dan tepat sasaran.
Diterima: 07-09-2025	
Disetujui: 12-12-2025	
<u>Kata Kunci</u>	
Data Mining	
<i>Naïve Bayes</i>	
Klasifikasi	
Bantuan Langsung Tunai	
nuriza3010@gmail.com*	

1. PENDAHULUAN

Saat ini, kemiskinan masih menjadi masalah yang berkelanjutan dan menjadi perhatian penuh dari pemerintah, dan hal ini menghambat perkembangan ekonomi baik di tingkat nasional maupun regional. Ketidakmampuan pada penyediaan dasar termasuk makanan, pakaian, perumahan, Pendidikan, hingga Kesehatan, dan kurangnya sumber daya menjadi kontribusi meningkatnya kemiskinan. Mengatasi permasalahan kemiskinan menjadi tanggung jawab pemerintah sesuai dengan ketentuan dan perundang-undangan yang berlaku [1][2]. Permasalahan kemiskinan di Indonesia menjadi fokus pemerintah dengan melakukan beberapa upaya untuk menangani kasus kemiskinan [3]. Dengan bekerja sama dengan Dinas Sosial (Dinsos), pemerintah berupaya meningkatkan kesejahteraan masyarakat di berbagai wilayah melalui beberapa program bantuan sosial, salah satunya adalah Bantuan Langsung Tunai (BLT) yang dialokasikan dari dana desa [1][4].

BLT merupakan sebuah inisiatif pemerintah yang bertujuan untuk meningkatkan jaringan pengaman sosial (JPS) melalui pemanfaatan dana desa yang diberikan dalam bentuk uang tunai. Inisiatif pemberian Bantuan Langsung Tunai (BLT) bertujuan untuk mengatasi berbagai masalah yang dihadapi oleh semua penduduk, khususnya yang berkaitan dengan masalah ekonomi [5][6]. Kebijakan ini berdasarkan instruksi Presiden (InPres) Nomor 4 Tahun 2022, yang bertujuan untuk mempercepat penghapusan kemiskinan di kabupaten dan kota, dan menetapkan data keluarga miskin berdasarkan hasil musyawarah di tingkat desa/ kelurahan [7]. Program penyaluran dana BLT harus dilaksanakan dengan benar, transparan dan teratur agar

diberikan dan diterima oleh masyarakat miskin yang berhak dan membutuhkan [6].

Saat ini, proses pemilihan penerima BLT di desa Pesanggrahan masih yang dilakukan secara manual dan tidak sepenuhnya bergantung pada sistem yang terorganisir, dimana proses pemilihan tersebut berdasarkan saran atau pendapat dari warga tertentu, hal tersebut mengakibatkan terjadinya kesalahan dalam penyaluran BLT. Proses pendataan penerima BLT yang dianggap tidak merata menyebabkan potensi ketidakadilan dan kecemburuhan sosial di kalangan masyarakat, serta menyebabkan penyaluran BLT tidak tepat sasaran. Selain itu, pendataan juga mengakibatkan proses penetapan penerima BLT membutuhkan waktu lama [8]. Masalah tersebut dapat diselesaikan dengan penerapan metode data *mining* sebagai pendekatan yang lebih objektif, relevan dan terukur dengan melakukan klasifikasi yang dapat membantu mengidentifikasi penduduk yang berhak menerima BLT berdasarkan kriteria yang sesuai. Metode ini juga dapat mengurangi kelasalahan dalam menentukan penduduk yang layak menerima BLT [9][10].

Data mining merupakan proses pengumpulan data dalam jumlah besar, mengekstraksi data tersebut menjadi format yang dapat digunakan, dan mengklasifikasikannya ke dalam satu kategori atau lebih yang telah ditentukan sebelumnya [2][11]. Klasifikasi merupakan metode yang digunakan untuk melakukan prediksi pada suatu nilai yang didalamnya terdapat atribut dan biasanya memperoleh seperangkat rules yang disebut aturan[12]. Proses klasifikasi bertujuan menemukan model atau fungsi yang akurat untuk menjelaskan dan membedakan antara kelas atau konsep yang menghasilkan prediksi kelas pada objek dengan label yang tidak diketahui [2]. Salah satu metode yang digunakan dalam klasifikasi adalah metode *Naïve Bayes* dengan mengelompokkan data statistik pada probabilitas sederhana untuk melakukan prediksi kejadian di masa depan berdasarkan hasil kejadian sebelumnya, dan menjumlahkan kombinasi dan frekuensi nilai dalam *dataset*[11].

Beberapa penelitian sebelumnya menerapkan algoritma *Naïve Bayes* untuk klasifikasi penerima program bantuan dari Pemerintah. Penelitian Huriah dan Nuris klasifikasi penentuan Bantuan Sosial UMKM menghasilkan *Accuracy* 95,43%, *Precision* 97,87%, dan *Recall* 93,88% dengan menggunakan beberapa atribut yaitu NIK, Nama, Alamat, Bidang Usaha, Pekerjaan dan Hasil[13]. Penelitian lain dilakukan oleh Muhammad Adi Mukti *et.al*[14] klasifikasi menggunakan algoritma *Naïve Bayes* dengan metode CRISP-DM yang dilakukan dengan dua cara, yaitu perhitungan manual dan perhitungan pada *tools Rapid Miner*. Perhitungan manual dengan jumlah data *training* sebanyak 282 dan data *testing* sebanyak 2 yang dilakukan dengan *confusion matrik* menghasilkan *Accuracy*, *Precision*, dan *Recall* masing-masing sebesar 100%. Pada pengujian menggunakan *Rapid Miner* dengan jumlah data *training* sebanyak 197 dan data *testing* 85 menghasilkan *performance* nilai *Accuracy* sebesar 98,82%, *Precision* sebesar 100%, dan *Recall* sebesar 96,97%. Penelitian yang dilakukan Anwar Pauji *et.al* klasifikasi penerima BLT menggunakan *K-Nearest Neighbor* menghasilkan nilai *Accuracy* sebesar 68,82% dengan prediksi Layak dan Ternyata True Layak sebesar 104, prediksi Layak dan Ternyata True Tidak Layak sebesar 16, prediksi Tidak Layak dan Ternyata True Tidak Layak sebesar 13, dengan nilai *Class Recall* layak sebesar 86,67%, dan *Class Recall* sebesar 26%[4]. Sedangkan penelitian yang dilakukan Musdalifah *et.al*[15] klasifikasi menggunakan Algortima C45 menghasilkan nilai akurasi sebesar 72,20% dengan *recall* kelas “menerima” sebesar 99,47% dan kelas *recall* “Tidak Menerima” sebesar 6,64%.

Berdasarkan temuan dari penelitian sebelumnya, algoritma *Naïve Bayes* menunjukkan akurasi yang lebih tinggi dalam mengklasifikasikan penerima program bantuan pemerintah, khususnya penerima Bantuan Langsung Tunai (BLT). Penelitian ini bertujuan untuk menerapkan algoritma *Naïve Bayes* dalam mengklasifikasikan penerima Bantuan Langsung Tunai (BLT) di desa Pesanggrahan, sehingga meningkatkan proses seleksi penerima BLT berdasarkan kriteria yang telah ditetapkan yaitu keluarga miskin, keluarga miskin yang tidak terkait bantuan lain, dan lansia [16].

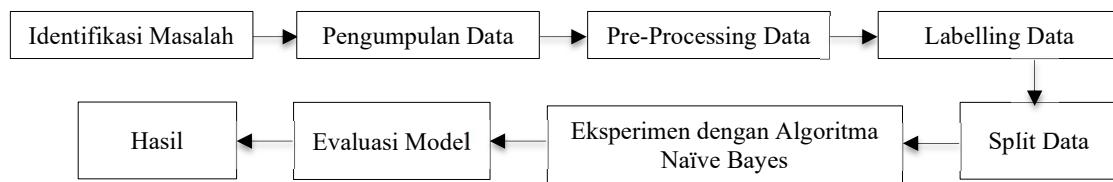
2. METODE

2.1. Jenis Penelitian

Penelitian ini menggunakan *mixed method*, yaitu menggabungkan metode kualitatif dan kuantitatif untuk memahami fenomena yang terjadi secara mendalam, sehingga menghasilkan pemahaman yang menyeluruh dan mendalam tentang permasalahan yang terjadi. Penelitian kualitatif merupakan studi yang menggambarkan peristiwa secara menyeluruh dengan metode pengumpulan data melalui wawancara (*interview*), studi dokumentasi, dan studi lapangan (*observasi*). Sedangkan penelitian kuantitatif merupakan penelitian yang mengacu pada data *numeric* untuk menjawab rumusan masalah serta memahami masalah dan fenomena yang muncul dalam objek penelitian [9].

2.2. Tahapan Penelitian

Untuk menyelesaikan permasalahan yang akan dibahas, dibutuhkan beberapa tahapan-tahapan penelitian yang ditunjukkan pada gambar 1 berikut.



Gambar 1. Tahapan Penelitian

Berdasarkan tahapan penelitian yang diuraikan pada gambar di atas, maka penjelasan masing-masing tahapan tersebut antara lain :

1. Identifikasi Masalah

Identifikasi masalah menjadi tahapan penting untuk memahami permasalahan yang terjadi pada objek penelitian, sehingga penulis dan pembaca memahami tujuan dari penelitian yang dilakukan [17].

2. Pengumpulan data

Tahap pengumpulan data merupakan pengumpulan informasi yang relevan dengan tujuan penelitian yang mencakup fenomena dan masalah yang muncul dalam objek penelitian. Pada penelitian ini, penulisan melakukan pengumpulan data melalui *interview* kepada pihak terkait, dan melakukan studi dokumentasi dengan cara mempelajari dan menganalisis dokumen-dokumen yang berkaitan dengan data penerima Bantuan Langsung Tunai (BLT), sehingga menghasilkan data yang relevan dan akurat [17]. Penelitian ini menggunakan *sample* data sebanyak 153 data penduduk penerima Bantuan Langsung Tunai terdiri dari 6 atribut, yaitu Nama, Jenis Kelamin (JK), Usia, Pekerjaan, Terkait Bantuan Lain, dan Dusun. Data penduduk penerima Bantuan Langsung Tunai ditunjukkan pada tabel 1 berikut.

Tabel 1. Data Penduduk Penerima Bantuan Langsung Tunai

No	Nama	Jenis kelamin	Usia	Pekerjaan	Terkait Bantuan Lain	Dusun
1	Dada	P	Lansia	Buruh Tani	Tidak	Krajan
2	Fatimah	P	Lansia	Buruh Tani	Tidak	Krajan
3	Hawe	P	Lansia	Buruh Tani	Tidak	Krajan
4	Pur	L	Lansia		Tidak	Gedang
5	Adeem	L	Lansia		Tidak	Krajan
6	Sunardi	L	Dewasa	Pedagang	Tidak	Krajan
7	Jan	P	Dewasa		Tidak	Acem
8	Dahrani	P	Lansia	Petani	Ya	Gedang
9	Sahma	P	Lansia	Petani	Ya	Gedang
...
153	Sa'ati	P	Lansia	Ibu Rumah Tangga	Iya	Krajan

3. *Pre-processing* data

Pre-processing data atau disebut dengan *cleaning data* merupakan proses yang dilakukan untuk menghapus data yang tidak lengkap, menghapus data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pengetikan data, membersihkan atribut kosong dan atribut yang tidak digunakan saat pemodelan [18][19].

4. *Labelling* data

Pada penelitian ini, tahapan *labelling data* dilakukan setelah *pre-processing* data dan dilakukan secara manual menggunakan aplikasi *Microsoft office Excel* dengan menentukan klasifikasi status layak dan tidak layak [20].

5. *Split* Data

Split data merupakan pembagian *dataset*, untuk memberi perintah pada algoritma agar dapat mengklasifikasi data ke dalam tiga kategori, yaitu positif, negatif, dan netral. Hasil split data ini digunakan untuk mengevaluasi model yang diperoleh dari pengolahan data, sehingga menghasilkan nilai akurasi yang optimal [21].

6. Pemodelan dengan Algoritma *Naïve Bayes*

Setelah pembagian data *training* dan data *testing* selasai, tahap selanjutnya adalah pemodelan untuk menentukan hasil klasifikasi menggunakan algoritma *Naïve Bayes* [19]. Implementasi algoritma *Naïve Bayes* untuk memprediksi dan mengklasifikasi data ke dalam kategori berdasarkan probabilitas dan dinyatakan dalam persamaan (1) berikut [11] [22]:

$$P(H|X) = \frac{P(X|H)P(H)}{P(x)} \quad (1)$$

Dimana :

- X : Data dengan klasifikasi yang tidak diketahui
- H : Hipotesis data untuk kelas tertentu
- $P(H|X)$: Probabilitas hipotesis H , dengan syarat X , sebagai probabilitas posterior
- $P(X|H)$: Probabilitas X ditentukan berdasarkan kondisi dalam hipotesis
- $P(H)$: Probabilitas hipotesis H (prior probabilitas)
- $P(X)$: Probabilitas X

Pada penelitian ini, pemodelan *Naïve Bayes* dengan metode *Gaussian* yang diimplementasikan menggunakan *library scikit-learn* pada bahasa pemrograman *Python*.

7. Evaluasi model

Evaluasi model merupakan tahapan yang dilakukan setelah hasil klasifikasi diperoleh, tahap ini dilakukan untuk mengetahui hasil akurasi terbaik dengan tingkat akurasi yang optimal[10]. Proses evaluasi ini dilakukan menggunakan pengujian *confusion matrix* yang terdiri *true positive* (Layak dan memang benar layak), *false negative* (Layak dan sebenarnya tidak layak), *true negative* (Tidak layak dan memang benar tidak layak), dan *false negative* (tidak layak padahal sebenarnya layak) [9]. Untuk menghitung nilai meliputi *accuracy*, *precision*, *recall*, dan *F-1 Score* menggunakan persamaan berikut [22].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F - 1 score = 2 * \frac{Precision * Recall}{Precision + recall} \quad (5)$$

8. Hasil

Tahap ini merupakan tahap akhir dari penelitian dan penarikan kesimpulan dari proses dan informasi penelitian.

3. HASIL DAN PEMBAHASAN

3.1. Pre-processing Data

Pre-processing data dilakukan untuk pembersihan data untuk menghasilkan mutu dan konsistensi *dataset* yang akan digunakan pada proses analisis. Tahap *pre-processing* ini terdiri dari penghapusan data duplikat dan penghapusan baris kosong.

1. Menghapus data duplikat

Menghapus data duplikat merupakan langkah awal tahap *pre-processing* data, yaitu dengan data yang tercatat lebih dari satu. Proses ini dilakukan karena data ganda berpotensi menimbulkan bias, sehingga harus dihapus agar hasil analisis menjadi lebih valid. Proses menghapus data duplikat dilakukan dengan menggunakan Bahasa pemrograman *Python* dengan membaca data *Excel* dari penyimpanan *google drive* ke dalam *environment Google Colab* yang terdiri dari 153 data. Berikut ini merupakan data awal sebelum menghapus data duplikat ditunjukkan pada gambar 2.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 153 entries, 0 to 152
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   NO              153 non-null    int64  
 1   NAMA             153 non-null    object  
 2   JENIS KELAMIN   153 non-null    object  
 3   USIA             153 non-null    object  
 4   PEKERJAAN        150 non-null    object  
 5   TIDAK TERKAIT BANTUAN LAIN 153 non-null    object  
 6   DUSUN            153 non-null    object  
dtypes: int64(1), object(6)
memory usage: 8.5+ KB
```

Gambar 2. Data Duplikat Sebelum Dihapus

Pada proses ini, penghapusan data duplikat dilakukan pada variabel NAMA, yang terdiri dari tiga data dengan mempertahankan data NAMA dengan kemunculan pertama, sehingga jumlah data terbaru setelah penghapusan data duplikat menjadi 150 data ditunjukkan pada gambar 3 berikut.

```
<class 'pandas.core.frame.DataFrame'>
Index: 150 entries, 0 to 151
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   NO              150 non-null    int64  
 1   NAMA             150 non-null    object  
 2   JENIS KELAMIN   150 non-null    object  
 3   USIA             150 non-null    object  
 4   PEKERJAAN        147 non-null    object  
 5   TIDAK TERKAIT BANTUAN LAIN 150 non-null    object  
 6   DUSUN            150 non-null    object  
dtypes: int64(1), object(6)
memory usage: 9.4+ KB
```

Gambar 3. Data Duplikat Setelah Dihapus

2. Menghapus Baris Kosong

Penghapusan baris yang berisi nilai *null* merupakan salah satu prosedur yang dilakukan. Baris yang tidak memiliki data lengkap berpotensi menyebabkan kesalahan dalam proses analisis maupun pelatihan model, oleh karena itu baris yang mengandung nilai kosong pada satu kolom diidentifikasi dan dihapus dari *dataset*. Langkah awal dilakukan dengan menghapus spasi pada setiap entri bertipe data string agar tidak terbaca sebagai nilai kosong. Selanjutnya, simbol tertentu diganti dengan *NaN*, sehingga lebih mudah dikenali sebagai data hilang. Proses penghapusan baris ini

dilakukan dengan menggunakan *Python* menggunakan fungsi *dropna*. Sebelum dilakukan penghapusan baris kosong jumlah data terdiri dari 150 data ditunjukkan pada gambar 4 berikut.

```
<class 'pandas.core.frame.DataFrame'>
Index: 150 entries, 0 to 151
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   NO              150 non-null     int64  
 1   NAMA             150 non-null     object  
 2   JENIS KELAMIN   150 non-null     object  
 3   USIA             150 non-null     object  
 4   PEKERJAAN        147 non-null     object  
 5   TIDAK TERKAIT BANTUAN LAIN 150 non-null     object  
 6   DUSUN            150 non-null     object  
dtypes: int64(1), object(6)
memory usage: 9.4+ KB
```

Gambar 4. Data Sebelum Penghapusan Baris Kosong

Setelah dilakukan penghapusan baris yang kosong data menjadi 147 data dengan menghapus 3 data pada baris kosong. Hasil penghapusan baris kosong ditunjukkan pada gambar 5 berikut.

```
<class 'pandas.core.frame.DataFrame'>
Index: 147 entries, 0 to 151
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   NO              147 non-null     int64  
 1   NAMA             147 non-null     object  
 2   JENIS KELAMIN   147 non-null     object  
 3   USIA             147 non-null     object  
 4   PEKERJAAN        147 non-null     object  
 5   TIDAK TERKAIT BANTUAN LAIN 147 non-null     object  
 6   DUSUN            147 non-null     object  
dtypes: int64(1), object(6)
memory usage: 9.2+ KB
```

Gambar 5. Data Setelah Penghapusan Baris Kosong

3.2. Labelling Data

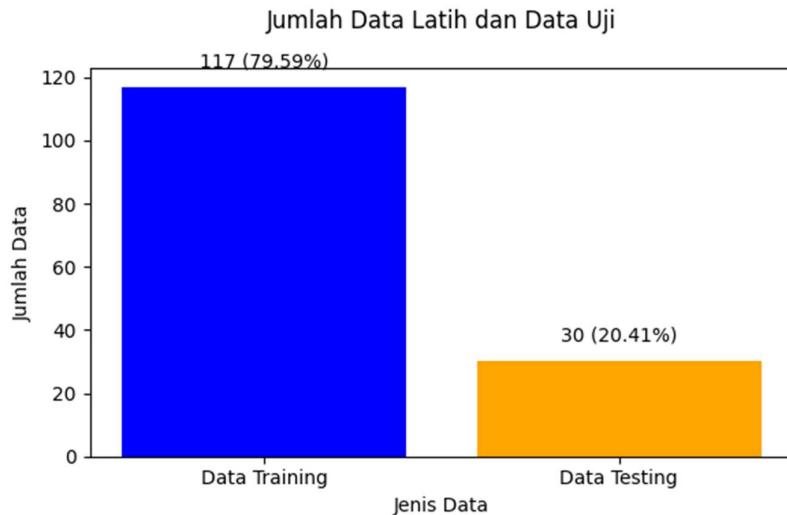
Proses *labelling* dengan mengamati setiap data satu per satu, lalu menetapkan label “Layak” atau “Tidak Layak” sesuai dengan kriteria yang ditentukan. Tabel 2 berikut merupakan hasil *labelling* data.

Tabel 2. Hasil Labelling Data

No	Nama	Jenis Kelamin	Usia	Pekerjaan	Terkait Bantuan Lain	Dusun	Status
1	Dada	P	Lansia	Buruh Tani	Tidak	Krajan	Layak
2	Fatimah	P	Lansia	Buruh Tani	Tidak	Krajan	Layak
3	Hawe	P	Lansia	Buruh Tani	Tidak	Krajan	Layak
4	Arwe	P	Lansia	Buruh Tani	Tidak	Krajan	Layak
5	Mabruk	L	Lansia	Buruh Tani	Tidak	Krajan	Layak
6	Asiah H	P	Lansia	Buruh Tani	Tidak	Krajan	Layak
7	Diyati	P	Dewasa	Petani	Tidak	Gedang	Tidak Layak
8	Pandi	L	Dewasa	Petani	Tidak	Acem	Tidak Layak
9	Asan Mustafa	L	Dewasa	Buruh Pabrik	Tidak	Gedang	Tidak Layak
...
147	Nisa	P	Lansia	Ibu Rumah Tangga	Ya	Banuaju	Tidak Layak

3.3. Split Data

Pada tahap *split data*, dataset dibagi menjadi dua bagian, yaitu data *training* sebanyak 117 data (79,59%) dan data *testing* sebanyak 30 data (20,41%). Pada gambar 6 berikut merupakan hasil dari pembagian data.



Gambar 6. Jumlah Data Latih dan Data Uji

3.4. Pemodelan Data dengan Algoritma *Naïve Bayes*

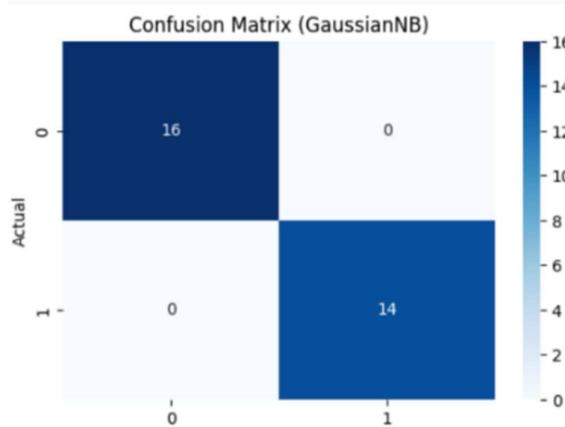
Langkah awal pada metode *Naïve Bayes* adalah mengubah data kategorikal menjadi *numeric* berupa token frekunesi kata. Tujuannya adalah untuk menganalisis pola kata yang paling umum dalam setiap kategori klasifikasi. Untuk membagi data *training* dan data *testing* menggunakan *train_test_split* dengan rasio perbandingan 80:20. Fungsi *fit()* untuk model *GaussianNB* digunakan untuk melakukan proses pelatihan yang sesuai untuk data yang terdistribusi secara normal menggunakan metode TF-IDF (*Term Frequency - Inverse Document Frequency*). Setelah pelatihan, model digunakan untuk prediksi data *testing* dan hasil prediksi untuk memudahkan interpretasi hasil klasifikasi, dan menentukan seberapa baik model dapat mengklasifikasi data. Hasil Prediksi untuk klasifikasi penerima Bantuan Lansung Tunai menggunakan algoritma *Naïve Bayes* menggunakan metode *GaussianNB* ditunjukkan pada gambar 7 berikut.

...	Predicted				
	USIA	PEKERJAAN	TIDAK TERKAIT BANTUAN LAIN	Actual	Predicted
125	DEWASA	IBU RUMAH TANGGA		TIDAK	tidak layak
51	LANSIA	BURUH TANI		TIDAK	layak
138	LANSIA	IBU RUMAH TANGGA		IYA	tidak layak
19	LANSIA	BURUH TANI		TIDAK	layak
104	LANSIA	TIDAK KERJA		IYA	tidak layak
12	LANSIA	BURUH TANI		TIDAK	layak
76	LANSIA	BURUH TANI		TIDAK	layak
31	LANSIA	PETANI		IYA	tidak layak
81	LANSIA	PETANI		TIDAK	tidak layak
9	LANSIA	BURUH TANI		TIDAK	layak
26	LANSIA	BURUH TANI		TIDAK	layak
96	DEWASA	IBU RUMAH TANGGA		TIDAK	tidak layak
143	DEWASA	IBU RUMAH TANGGA		TIDAK	tidak layak

Gambar 7. Prediksi Status Kelayakan Penerima Bantuan Langsung Tunai

3.5. Evaluasi Model

Setelah hasil klasifikasi algoritma *Naïve Bayes* diperoleh, langkah selanjutnya adalah melakukan evaluasi untuk menentukan performa dan akurasi pada model klasifikasi dengan menentukan perbandingan hasil prediksi menggunakan *confusion matrik*, sebagaimana pada gambar 8 di bawah ini.

**Gambar 8. Confusion Matrix (Gaussian NB)**

Evaluasi *precision*, *recall*, *F-1 score* untuk menilai akurasi prediksi model terhadap setiap kelas menggunakan *classification report* dengan menguji tingkat akurasi, ketepatan prediksi model dan jumlah seluruh prediksi dibandingkan total seluruh data ditunjukkan pada gambar 9.

Classification Report (GaussianNB):				
	precision	recall	f1-score	support
layak	1.00	1.00	1.00	16
tidak layak	1.00	1.00	1.00	14
accuracy			1.00	30
macro avg	1.00	1.00	1.00	30
weighted avg	1.00	1.00	1.00	30

=====

Accuracy (GaussianNB): 1.0000

=====

Gambar 9. Classification Report (Gaussian NB)

Selanjutnya menghitung *confusion matrix* untuk memperoleh hasil prediksi yang sangat baik, dimana seluruh data uji berhasil diklasifikasikan dengan benar. Terdapat 16 data dengan kategori kayak berhasil diprediksi dengan tepat, dan 14 data dalam kategori tidak layak dengan diprediksi dengan akurat, tanpa ada kesalahan klasifikasi. Dengan demikian, nilai *true positive* (TP) dan *true negative* (TN) masing-masing adalah 16 dan 14, sedangkan nilai *false positive* dan *false negative* adalah, nilai hasil perhitungan akurasi sebesar 1.00 atau 100% dengan persamaan (2) berikut.

$$\text{Accuracy} = \frac{16 + 14}{16 + 14 + 0 + 0} = \frac{30}{30} = 1 \times 100\% = 1.00$$

Berdasarkan hasil perhitungan di atas, dapat disimpulkan bahwa algoritma *Naïve Bayes* dengan metode *gausianNB* mampu mengklasifikasi data dengan hasil yang sangat tinggi dan sesuai dengan karakteristik yang sudah ditentukan.

Setelah melakukan perhitungan akurasi, selanjutnya adalah melakukan perhitungan *classification report* pada *precision*, *recall*, dan *F-1 score* pada setiap kelas. *Precision* dihitung berdasarkan data kategori layak dan tidak terdapat satupun yang salah diklasifikan sebagai kategori layak padahal sebenarnya tidak layak, sehingga *precision*-nya adalah 16 dibagi 16 menghasilkan nilai 1.00. kelas *precision* tidak layak terdapat 14 data yang seluruhnya berhasil diprediksi dengan benar tanpa kesalahan prediksi positif, sehingga menghasilkan nilai *precision* 1.00 dengan persamaan (3) berikut.

$$\text{Precision Layak} = \frac{16}{16 + 0} = 1.00$$

$$Precision \text{ Tidak Layak} = \frac{14}{14 + 0} = 1.00$$

Tahap selanjutnya setelah menghitung *precision* adalah melakukan perhitungan *recall* untuk kelas layak yang terdiri dari 16 data kategori layak, dimana seluruh data berhasil diprediksi dengan tepat tanpa ada kesalahan klasifikasi menjadi tidak layak atau FN 0. Hal ini menghasilkan nilai *recall* sebesar 16 dibagi 16+0, menghasilkan nilai 1.00. Sedangkan untuk kelas tidak layak, terdapat 16 data yang seluruhnya berhasil diprediksi dengan benar tanpa ada kesalahan klasifikasi sebagai layak, sehingga perhitungan *recall*-nya adalah 14 dibagi 14+0 menghasilkan nilai 1.00 dengan persamaan (4) berikut.

$$Recall \text{ Layak} = \frac{16}{16 + 0} = 1.00$$

$$Recall \text{ Tidak Layak} = \frac{14}{14 + 0} = 1.00$$

Setelah melakukan perhitungan *recall*, tahap selanjutnya adalah melakukan perhitungan *f-1 score* dengan mengkombinasikan nilai *precision* dan nilai *recall*, sehingga metrik ini mampu memberikan gambaran yang jelas dan lebih menyeluruh mengenai hasil kinerja model, tidak seperti hasil akurasi yang hanya menunjukkan persentase prediksi benar secara umum. Hasil perhitungan *f-1 score* lebih menekankan keseimbangan antara ketepatan prediksi positif dan kemampuan model dalam mengenali seluruh data positif yang ada.

Berdasarkan hasil pengujian pada kategori layak, nilai *precision* tercatat 1.00 dan *recall* juga menunjukkan nilai 1.00. hal yang sama juga terlihat pada kategori Tidak Layak, di mana kedua metrik tersebut sama-sama memperoleh nilai sempurna 1.00. hasil pengujian tersebut memperlihatkan bahwa algoritma *Naïve Bayes* mampu menganalisis dan mengklasifikasi data dengan sangat baik, tanpa menghasilkan kesalahan prediksi baik kelas positif maupun kelas *negative* dengan persamaan (5) berikut.

$$f1 - score \text{ Layak} = 2X \frac{1.00 \times 1.00}{1.00 + 1.00} = 2X \frac{1.00}{2.00} = 1.00$$

$$f1 - score \text{ Tidak Layak} = 2X \frac{1.00 \times 1.00}{1.00 + 1.00} = 2X \frac{1.00}{2.00} = 1.00$$

Perhitungan *f-1 score* merupakan penggabungan hasil *precision* dan *recall*, dan hasil nilai yang peroleh adalah 1.00. nilai maksimal ini menjadi bukti bahwa model menghasilkan nilai yang optimal, sebab mampu menjaga keseimbangan antara ketepatan prediksi dan kelengkapan deteksi.

Berdasarkan hasil analisis tersebut, metode *Naïve Bayes* yang diterapkan pada penelitian ini bekerja sangat efektif dalam menyelesaikan permasalahan klasifikasi pada data yang diuji.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian, implementasi algoritma *Naïve Bayes* untuk memudahkan proses klasifikasi penerima Bantuan Langsung Tunai (BLT) secara objektif. Hasil pengujian menunjukkan performa model yang sangat baik, dengan menggunakan *confusion matrix* menunjukkan bahwa data uji dapat diprediksi tanpa adanya kesalahan klasifikasi dengan nilai *accuracy*, *precision*, *recall*, *f1-score* masing-masing sebesar 100%. Penggunaan Bahasa pemrograman *Python* dikombinasikan dengan *Metode Naïve Bayes Gaussian* memungkinkan pengembangan model yang mampu mengklasifikasikan data dengan akurat sesuai kriteria yang ditetapkan oleh pemerintah desa, yaitu keluarga miskin, tidak terkait bantuan lain, dan lansia. Hal tersebut menjadi solusi untuk menyelesaikan permasalahan dan mengurangi kesalahan yang dilakukan secara manual, sehingga proses seleksi penerima Bantuan Langsung Tunai (BLT) tepat sasaran berdasarkan kriteria dan pengolahan data secara *real time* dan dapat dimanfaatkan secara berkelanjutan oleh pemerintah desa.

Klasifikasi Penerima Bantuan Langsung Tunai menggunakan Algoritma *Naïve Bayes* menghasilkan kurasi yang sangat baik, oleh karena itu pada penelitian selanjutnya disarankan penerapan algoritma klasifikasi seperti *Support Vector Machine* (SVM), *Convulucional Neural*

Network (CNN), atau algoritma optimasi seperti *particle swam optimization* (PSO) dan *Genetic Algorhthm* (GA), sebagai perbandingan hasil akurasi pada masing-masing algoritma sesuai dengan karakteristik data kebutuhan desa.

5. DAFTAR PUSTAKA

- [1] A. A. H. Tasyifa Nafsiah Muthmainnah, Apriade Voutama, “Menggunakan Algoritma K-Means,” vol. 8, no. 3, pp. 3916–3919, 2024.
- [2] A. Surahman and U. Hayati, “Implementasi Algoritma *Naïve Bayes* Untuk Prediksi Penerima Bantuan Sosial,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 347–352, 2023, doi: 10.36040/jati.v7i1.6302.
- [3] A. . & Aldiansyah and D. Santoso, “Implementasi Metode SAW pada Sistem Pendukung Keputusan Pemilihan Supplier.,” *J. Tek. Inform.*, vol. 4, no. 2, pp. 87–95, 2016.
- [4] Anwar Pauji, S. Aisyah, A. Surip, R. Saputra, and I. Ali, “Implementasi Algoritma K-Nearest Neighbor Dalam Menentukan Penerima Bantuan Langsung Tunai,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 1, pp. 21–27, 2022, doi: 10.32485/kopertip.v4i1.114.
- [5] N. A. Kilo, M. R. Katili, and I. K. Hasan, “Research in the Mathematical and Natural Sciences Perbandingan Metode K-Means dan K-Medoids Dengan Validitas Davies-Bouldin Indeks , Dunn Indeks dan Indeks Connectivity Pada Pengelompokan Masyarakat Penerima Bantuan Langsung Tunai,” vol. 4, no. 1, pp. 8–15, 2025, doi: 10.55657/rmns.v4i1.190.
- [6] J. S. Informasi, T. Informasi, W. Dari, and A. Y. Sari, “Implementasi Data Mining Dengan *Naïve Bayes* Untuk Prediksi Penerima Bantuan Langsung Tunai (Blt) Warga Desa Xyz,” vol. 2, no. 2, pp. 29–40, 2023.
- [7] BPK, “Instruksi Presiden Republik Indonesia Nomor 4 Tahun 2022 Tentang Percepatan Penghapusan Kemiskinan Ekstrem,” *Badan Pemeriksaan Keuangan*, no. 146187, pp. 1–15, 2022, [Online]. Available: <https://peraturan.bpk.go.id/Details/211477/inpres-no-4-tahun-2022>
- [8] U. Habibah and M. Rosyda, “Sistem Pendukung Keputusan Penerima Bantuan Langsung Tunai Dana Desa di Pekandangan Menggunakan Metode AHP-TOPSIS,” *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 404, 2022, doi: 10.30865/mib.v6i1.3471.
- [9] K. Penerima and B. Di Desa, “Analisis Perbandingan Varian *Naïve Bayes* dalam Klasifikasi,” vol. 8, pp. 281–292, 2025.
- [10] D. Kurniadi, F. Nuraeni, and M. Firmansyah, “Klasifikasi Masyarakat Penerima Bantuan Langsung Tunai Dana Desa Menggunakan *Naïve Bayes* dan SMOTE,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 2, pp. 309–320, 2023, doi: 10.25126/jtiik.20236453.
- [11] N. Alfiah and P. S. Informasi, “Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan Menggunakan Metode Naive Bayes,” vol. XVI, pp. 32–40, 2021.
- [12] R. B. Syadewo and N. Riza, “KLASIFIKASI PENERIMAAN DANA BANTUAN PADA DUSUN JATI BENING,” vol. 7, no. 2, pp. 1220–1226, 2023.
- [13] D. A. Huriah, N. D. Nuris, B. Usaha, D. Mining, A. Naive, and B. Dalam, “Klasifikasi penerima bantuan sosial umkm menggunakan algoritma *Naïve Bayes*,” vol. 7, no. 1, pp. 360–365, 2023.
- [14] Muhamad Adhi Mukti, W. Wiyanto, and Ismasari Nawangsih, “Implementasi Algoritma Naive Bayes Untuk Prediksi Calon Penerima Blt,” *J. Inform. Teknol. dan Sains*, vol. 5, no. 3, pp. 383–392, 2023, doi: 10.51401/jinteks.v5i3.3106.
- [15] M. Daud, R. Juita, and C. D. Suhendra, “Penerapan Metode Algoritma C4 . 5 Untuk Klasifikasi Kelayakan Penerima Program Bantuan Pada Dinas Sosial Kabupaten Manokwari,” vol. 5, no. 1, pp. 271–278, 2025.
- [16] “Perkades BLT 2023.”

- [17] I. Arfanda, W. Ramdhan, and R. A. Yusda, “Naive Bayes Dalam Menentukan Penerima Bantuan Langsung Tunai Digital Transformation Technology (Digitech) | e-ISSN : 9999-9999,” no. June 2021, 2023, doi: 10.47709/digitech.v1i1.1091.
- [18] I. Hadianti *et al.*, “PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK MENENTUKAN,” vol. 7, no. 6, pp. 3616–3620, 2023.
- [19] H. Nurrifqi *et al.*, “Klasifikasi Penerima Bansos Menggunakan Algoritma Naive Bayes,” vol. 10, no. 1, pp. 683–695, 2023.
- [20] E. Poerwandono and F. J. Perwitosari, “Penerapan Data Mining Untuk Penilaian Kinerja Karya Di PT Riksa Dinar Djaya Menggunakan Metode Naive Bayes Classification,” vol. 5, no. 1, pp. 336–340, 2023.
- [21] N. S. Ramadan and D. Darwis, “PERBANDINGAN METODE *NAÏVE BAYES* DAN SVM UNTUK SENTIMEN ANALISIS MASYARAKAT TERHADAP SERANGAN RANSOMWARE PADA DATA KIP-K,” vol. 8, no. 1, pp. 12–23, 2025.
- [22] M. Rifat and M. Akhriza, “Klasifikasi Penerima Bantuan Sosial Program Keluarga Harapan (PKH) Kecamatan Gondanglegi Menggunakan Metode *Naïve Bayes* Pendahuluan Metode Penelitian,” vol. 24, pp. 289–296, 2025.