

Analisis Kelayakan Penerima Bantuan Langsung Tunai di Desa Belendung dengan Pendekatan Metode Naive Bayes

1st Ahmad KamaludinUniversitas Buana Perjuangan Karawang
Karawang, Indonesia

if17.ahmadkamaludin@mhs.ubpkarawang.ac.id

2nd Deden WahiddinUniversitas Buana Perjuangan Karawang
Karawang, Indonesia

deden.wahiddin@ubpkarawang.ac.id

3rd Kiki Ahmad BaihaqiUniversitas Buana Perjuangan Karawang
Karawang, Indonesia

kikiahmad@ubpkarawang.ac.id

Abstract— Bantuan Langsung Tunai (BLT) adalah program pemerintah yang membantu masyarakat miskin untuk meningkatkan kesejahteraan. Di Desa Belendung, pemilihan penerima BLT masih dilakukan secara manual sehingga kurang tepat sasaran. Untuk menyelesaikan masalah ini, digunakan metode Naive Bayes dalam menentukan kelayakan penerima. Metode ini adalah algoritma pengklasifikasian statistik yang memperkirakan probabilitas keanggotaan dalam suatu kelas. Penelitian ini menerapkan kriteria seperti pekerjaan, penghasilan, tanggungan, status rumah, dan jumlah kendaraan. Data yang dianalisis menggunakan Python melibatkan 324 dataset penerima BLT tahun 2021, di mana 259 data untuk pelatihan dan 65 data untuk pengujian. Hasil uji menunjukkan akurasi 93%, presisi 100%, dan recall 96%, menandakan metode ini efektif.

Kata Kunci – Naive Bayes, klasifikasi, bantuan langsung tunai, Python

I. PENDAHULUAN

Kemiskinan tetap menjadi isu yang cukup serius dan masih dirasakan oleh sejumlah negara di seluruh dunia. Dalam rangka mengatasi masalah ini, pemerintah telah mengimplementasikan kebijakan tertentu, salah satunya adalah program bantuan langsung tunai (BLT). BLT dapat diartikan sebagai pemberian sejumlah uang tunai kepada masyarakat yang kurang mampu, setelah pemerintah melakukan kenaikan harga bahan bakar minyak (BBM) dengan mengurangi subsidi yang ada. Selisih dari pengurangan subsidi tersebut kemudian disalurkan kepada masyarakat miskin [1].

Bantuan langsung tunai yang kerap disingkat BLT adalah sebuah inisiatif pemerintah yang menyediakan dana bagi masyarakat yang hidup dalam kemiskinan. Program BLT bertujuan untuk mencapai distribusi kesejahteraan yang lebih merata di antara masyarakat. Untuk mendukung hal ini, diperlukan suatu sistem keputusan yang dapat membantu dalam pemilihan individu, sehingga pelaksanaan keputusan dapat dilakukan dengan cepat, tepat, dan akurat.

Program BLT telah dilaksanakan di Desa Belendung. Namun kenyataannya, tujuan pemberian bantuan kepada masyarakat kurang mampu tidak tercapai karena proses seleksi yang dilakukan secara manual memberikan hasil yang kurang optimal, sehingga menimbulkan banyak potensi kesalahan dalam pendataan dan seleksi.

Sebuah penelitian oleh Titis Diah Pangestuti (2020) tentang sistem pendukung keputusan rekrutmen pegawai baru menggunakan Naive Bayes Classifier mencatat akurasi 60% dengan 20 data uji. Penelitian lain oleh Febri Liantoni (2015) mengenai klasifikasi daun dengan metode K-Nearest Neighbor mencapai akurasi 86,67% berdasarkan 13 dari 15 data pengujian. Sebuah studi oleh Tohayah dan kolega (2021) menggunakan metode Topsis dan algoritma C4.5 untuk penerima bantuan langsung tunai dengan akurasi 76%, presisi 82,61%, dan recall 90,48%. Naive Bayes memiliki keunggulan dalam memerlukan sedikit data pelatihan untuk klasifikasi yang akurat.

Dalam penelitian ini, klasifikasi akan diterapkan menggunakan metode Naive Bayes. Data yang akan digunakan diambil dari penerima BLT Desa Belendung pada tahun 2021, dengan total 324 dataset. Sebanyak 259 dataset akan digunakan sebagai data pelatihan. Hasil dari pengujian dengan 259 data pelatihan dan 65 data pengujian akan dihitung menggunakan bahasa pemrograman Python.

II. DATA DAN METODE

A. Algoritma Naïve Bayes

Pengklasifikasi Naïve Bayes adalah model probabilistik yang mudah dipahami yang menggunakan frekuensi kumpulan data dan kombinasi nilai untuk menghasilkan sekumpulan probabilitas. Bergantung pada nilai variabel kelas, metode ini menggunakan Teorema Bayes untuk mengasumsikan bahwa semua kualitas bersifat independen atau tidak saling bergantung. Naïve Bayes adalah metode yang mengelompokkan permasalahan ke dalam berbagai kelas berdasarkan karakteristik kesamaan dan perbedaan dengan menggunakan statistik untuk memprediksi probabilitas suatu kelas [5]. Persamaan Teorema Bayes adalah:

$$P(F|X) = \frac{P(X|F)}{P(X)} \cdot P(F)$$

Ket :

X : Sampel data yang memiliki *class* (label) yang tidak diketahui.

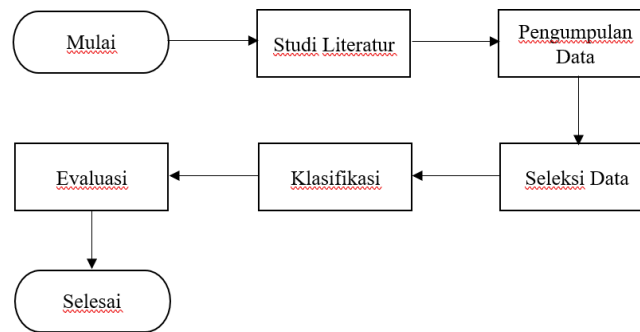
F : Hipotesis bahwa X adalah data *class* (label).

$P(F)$: Probabilitas Hipotesis F .

$P(X|F)$: Probabilitas berdasarkan kondisi pada hipotesis.

$P(X)$: Peluang dari data sampel yang diamati (probabilitas F).

B. Prosedur Penelitian



Gambar 1 Prosedur Penelitian

Prosedur penelitian memiliki tahapan seperti pada Gambar 2. Tahapan pertama adalah studi literatur, di mana peneliti mengumpulkan informasi dari bermacam sumber seperti internet, jurnal, perpustakaan, dan berbagai referensi lainnya yang terkait dengan topik penelitian. Pada tahap selanjutnya dilakukan pengumpulan data warga penerima BLT Desa Belendung tahun 2021 di Kantor Desa Belendung. Tahap berikutnya adalah seleksi data, yaitu proses memilih atribut yang akan dipakai dan menghapus atribut yang tidak dibutuhkan untuk penelitian. Seleksi variabel dilakukan untuk menghilangkan variabel yang tidak diperlukan dalam proses klasifikasi. Pada tahap klasifikasi, peneliti menggunakan bahasa pemrograman Python melalui platform Google Colab dan menerapkan algoritma klasifikasi Naïve Bayes. Data yang telah melalui proses klasifikasi kemudian dievaluasi menggunakan confusion matrix untuk menghitung nilai akurasi, presisi, dan recall.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data diperoleh dari aparatur Desa Belendung, Kecamatan Klari, Kabupaten Karawang pada tahun 2021 dalam format Microsoft Excel. Data tersebut mencakup beberapa variabel, yaitu nama, jenis kelamin, alamat, pekerjaan, penghasilan, jumlah tanggungan, status kepemilikan rumah, jumlah kendaraan, serta keterangan. Dataset yang tersedia berjumlah 324, sebanyak 259 data digunakan untuk tujuan pelatihan (training), sementara 65 data lainnya digunakan untuk pengujian (testing).

Tabel 1 Data Penerima Bantuan

NO	NAMA	JK	ALAMAT	PEKERJAAN	PENG	TANG	STAT RM	KEN	KET
1	Asep	L	Krajan	Petani	1.500.000	1	Milik Sendiri	1	Dapat
2	Wawan	L	Krajan	Petani	2.000.000	3	Sewa	1	Dapat
3	Didin	L	Krajan	Pedagang	6.000.000	3	Milik Orang lain	2	Tidak Dapat
4	Cecep	L	Krajan	Buruh Lepas	2.000.000	2	Milik Sendiri	1	Dapat
5	Dadang	L	Krajan	Petani	1.300.000	2	Milik Orang tua	1	Dapat
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
259	Carkem	P	Pundong	Karyawan	5.000.000	4	Sewa	2	Tidak Dapat

B. Seleksi Data

Proses penyaringan dilakukan pada tahap ini untuk menghapus atribut yang memiliki data yang tidak lengkap atau mengandung kesalahan penulisan. Setelah itu, proses dilanjutkan dengan pemilihan atribut untuk memusatkan perhatian pada data yang relevan. Atribut yang akan digunakan meliputi:

Tabel 2 Seleksi Variabel

Variabel	Digunakan	Keterangan
Nama	X	Tidak dipakai
Jenis Kelamin	√	Dipakai
Alamat	X	Tidak dipakai
Pekerjaan	√	Dipakai
Penghasilan	√	Dipakai
Tanggungan	√	Dipakai
Status Kepemilikan Rumah	√	Dipakai
Jumlah Kendaraan	√	Dipakai
Keterangan	√	Dipakai

Atribut yang telah dipilih kemudian akan dihitung secara otomatis menggunakan bahasa pemrograman Python.

Tabel 3 Data dan Atribut setelah tahap seleksi

NO	JK	PEK	PENG	TANG	STAT	KEN	KET
1	L	Petani	1.500.000	1	Milik Sendiri	1	Dapat
2	L	Petani	2.000.000	3	Sewa	1	Dapat
3	L	Pedagang	6.000.000	3	Milik Orang lain	2	Tidak Dapat
4	L	Buruh Lepas	2.000.000	2	Milik Sendiri	1	Dapat
5	L	Petani	1.300.000	2	Milik Orang tua	1	Dapat
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
259	P	Karyawan	5.000.000	4	Sewa	2	Tidak Dapat

C. Implementasi

1. Pengujian Sistem

Peneliti menguji perhitungan Naïve Bayes dengan bantuan Google Colab. Langkah awal yang dilakukan yaitu import library NumPy, Matplotlib, Pandas, dan Scikit-learn. Kemudian memasukkan dataset yang berformat CSV dan melakukan print dataset untuk melihat apakah data sudah terbaca oleh sistem. Setelah dataset muncul dan tidak terjadi error, lanjut import LabelEncoder untuk mengubah data string menjadi angka agar proses bisa terbaca oleh sistem, lalu masukkan perintah variabel encoder enc. Selanjutnya, transformasikan data string menjadi angka. Dataset yang sudah ditransformasi akan terlihat seperti pada Gambar 2.

```
#print dataset
dataset
```

	JENIS KEL	PEKERJAAN	PENDAPATAN	TANGGUNGAN	STATUS RUMAH	JUMLAH KENDARAAN	KETERANGAN
0	0	3	1500000	1	1	1	0
1	0	3	2000000	3	0	1	0
2	0	2	6000000	3	1	2	1
3	0	0	2000000	2	1	1	0
4	0	3	1300000	2	0	1	0
...	...	...	...	...	...	...	...
254	0	0	2000000	3	0	1	0
255	1	3	2000000	3	1	1	0
256	1	1	6000000	4	1	2	1
257	0	0	1800000	2	0	1	0
258	0	3	2000000	3	0	1	0

259 rows × 7 columns

Gambar 2 Dataset Setelah ditransformasi

Jika data sudah ditransformasi, selanjutnya membuat variabel X dan y. Tahap selanjutnya yaitu memasukkan data test dengan cara membuat X_train, X_test, y_train, dan y_test dengan test size atau jumlah data test sebanyak 65 dan random state 0. Hasilnya bisa dilihat di Gambar 3.

```
#masukkan xtrain,xtest dengan menggunakan 33 data test
from sklearn.model_selection import train_test_split
#(membagi xtrain, xtest & ytrain, ytest)
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=65, random_state=0)

#print hasil
print(x_test)
```

[	0	0	2200000	3	1	1]
[	1	2	7000000	5	0	3]
[	1	3	1300000	2	1	1]
[	1	1	6000000	4	1	2]
[	0	2	6000000	3	1	2]
[	0	1	3000000	3	1	1]
[	0	3	1800000	3	0	1]
[	1	0	1800000	2	0	1]
[	0	3	1500000	1	1	1]
[	1	3	1300000	2	0	1]
[	0	2	7000000	5	0	3]
[	0	3	1800000	3	0	1]

Gambar 3 Import train data split dan hasil x test

Masukkan feature scaling dengan import StandardScaler agar perhitungan data sesuai, dan mentransform X_train dan X_test dengan menggunakan fit transform. Lalu, masukkan library algoritma Naïve Bayes untuk dapat menghitung confusion matrix.

Input y prediksi classifier untuk X_test, lalu hitung confusion matrix dengan import confusion matrix dan masukkan y_test dan y_pred. Jika sudah, print hasil dari confusion matrix yang menunjukkan:

- "Class Layak" dan "Prediksi Layak" = 43
- "Class Tidak Layak" dan "Prediksi Layak" = 0
- "Class Layak" dan "Prediksi Tidak Layak" = 4
- "Class Tidak Layak" dan "Prediksi Tidak Layak" = 18

Hasil hitung dapat dilihat pada Gambar 4.

```
#prediksi hasil xtest
y_pred = classifier.predict(x_test)

#menghitung confusion matrix
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print("confusion matrix")
print(cm)
```

```
confusion matrix
[[43  0]
 [ 4 18]]
```

Gambar 4 Prediksi xtest classifier dan Confusion matrix

Langkah terakhir adalah menampilkan hasil akurasi dengan menggunakan laporan klasifikasi (classification report) berdasarkan y_test dan y_prediksi, serta menghitung skor akurasi (accuracy score).

```
#menghitung akurasi skor
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

#menampilkan hasil akurasi
akurasi = classification_report(y_test, y_pred)
print("Tingkat Akurasi Algoritma Naive Bayes")
print("Akurasi :", akurasi)
akurasi = accuracy_score(y_test, y_pred)
print("Tingkat Akurasi : %d persen" %(akurasi*100))
```

Tingkat Akurasi Algoritma Naive Bayes	precision	recall	f1-score	support
Akurasi :				
0	0.91	1.00	0.96	43
1	1.00	0.82	0.90	22
accuracy			0.94	65
macro avg	0.96	0.91	0.93	65
weighted avg	0.94	0.94	0.94	65

Tingkat Akurasi : 93 persen

Gambar 5 Hasil akurasi, presisi, dan recall

Hasil pengujian sistem menunjukkan bahwa dari 65 data uji, diperoleh akurasi sebesar 93% (0,93), presisi mencapai 100% (1,00), dan nilai recall sebesar 96% (0,96).

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

1. Proses klasifikasi dilakukan dengan membagi data menjadi 259 data pelatihan dan 65 data pengujian. Tahapannya meliputi pengujian atribut untuk menentukan nilai rata-rata (mean), perhitungan standar deviasi, penentuan probabilitas kelas, penghitungan distribusi Gaussian, dan diakhiri dengan analisis confusion matrix untuk mengevaluasi tingkat akurasi, presisi, dan recall.
2. Penerapan algoritma Naïve Bayes dalam proses klasifikasi dievaluasi berdasarkan nilai akurasi, presisi, dan recall, dengan pembagian data sebanyak 259 untuk pelatihan (training) dan 65 untuk pengujian (testing). Hasil pengujian menggunakan bahasa pemrograman Python menunjukkan akurasi sebesar 93%, presisi mencapai 100%, dan recall sebesar 96%.

B. Saran

1. Diharapkan penggunaan algoritma klasifikasi lain dapat dilakukan untuk membandingkan hasil perhitungan antara algoritma Naïve Bayes dan algoritma lainnya. Hal ini juga dapat menjadi acuan bagi penelitian berikutnya yang membahas permasalahan serupa.
2. Agar memperoleh hasil yang lebih akurat, disarankan untuk menambahkan sampel data yang lebih beragam.

V. PENGAKUAN

Naskah ilmiah ini adalah bagian dari penelitian tugas akhir yang dilakukan oleh Ahmad Kamaludin dengan judul Penerapan Metode Naïve Bayes untuk Menguji Kelayakan Data Penerima Bantuan Langsung Tunai pada Desa Belendung. Penelitian ini dibimbing oleh Deden Wahiddin, M.Kom., dan Kiki Ahmad Baihaqi, M.Kom.

VI. DAFTAR PUSTAKA

- [1] I. Arfanda, W. Ramdhan, R. A. Yusda, dan H. Artikel, "Naive Bayes Dalam Menentukan Penerima Bantuan Langsung Tunai," vol. 1, no. 1, 2021, doi: 10.47709/briliance.vxix.xxxx.
- [2] Titis Diah Pangestuti, "RANCANG BANGUN SISTEM PENDUKUNG KEPUTUSAN PENERIMAAN KARYAWAN BARU MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," 2020.
- [3] F. Liantoni dan H. Nugroho, "KLASIFIKASI DAUN HERBAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN KNEAREST NEIGHBOR".
- [4] N. Ariesanto Ramdhan dan H. Bhakti, "METODE TOPSIS DAN ALGORITMA C4.5 DALAM MENENTUKAN PENERIMA BANTUAN LANGSUNG TUNAI," *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, vol. 3, no. 2, hlm. 188–196, 2021.
- [5] S. Lestari, M. Badrul, P. Studi Sistem Informasi, dan S. Nusa Mandiri Jakarta, "IMPLEMENTASI KLASIFIKASI NAIVE BAYES UNTUK PREDIKSI KELAYAKAN PEMBERIAN PINJAMAN PADA KOPERASI ANUGERAH BINTANG CEMERLANG," vol. 7, no. 1, 2020.