
Klasifikasi Penggunaan Daya Listrik Rumah Tangga Menggunakan Metode *Naïve Bayes*

Fitriani¹⁾, Sriani²⁾, Suhardi³⁾

^{1,2,3)}Prodi Ilmu Komputer, Fakultas Sains dan Teknologi
Universitas Islam Negeri Sumatera Utara

*Corresponding Email: fitriani lubis9912@gmail.com

Abstrak

Listrik merupakan kebutuhan primer dalam kehidupan manusia yang hampir tidak dapat dipisahkan dari berbagai aktivitas sehari-hari. PLN sebagai penyedia listrik nasional memberikan subsidi kepada pelanggan rumah tangga kurang mampu dengan daya 450 VA dan 900 VA. Namun, penyaluran subsidi ini dinilai belum tepat sasaran karena keterbatasan data penduduk yang digunakan. Penelitian ini mengusulkan metode klasifikasi ulang penerima subsidi menggunakan algoritma *Naïve Bayes*. Dengan menggunakan 720 data latih dan sejumlah variabel penentu kelayakan, diperoleh akurasi model sebesar 92%. Hasil ini menunjukkan bahwa metode *Naïve Bayes* efektif untuk mendukung pengambilan keputusan dalam penyaluran subsidi listrik secara lebih tepat dan objektif.

Kata Kunci: *Naïve Bayes*, Subsidi Listrik, Klasifikasi, Machine Learning, Python, Data Mining.

Abstract

Electricity is a primary need in human life and is almost inseparable from daily activities. PLN, as the national electricity provider, offers subsidies to low-income households with power capacities of 450 VA and 900 VA. However, the distribution of these subsidies is considered inaccurate due to the limitations of the population data used. This study proposes a reclassification method for subsidy recipients using the *Naïve Bayes* algorithm. By utilizing 720 training data and several eligibility-determining variables, the model achieved an accuracy of 92%. These results indicate that the *Naïve Bayes* method is effective in supporting decision-making for a more accurate and objective electricity subsidy distribution.

Keywords: *Naïve Bayes*, Electricity Subsidy, Classification, Machine Learning, Python, Data Mining.

PENDAHULUAN

Listrik merupakan komoditas vital dalam kehidupan manusia modern. Berbagai aspek kehidupan, mulai dari kegiatan domestik rumah tangga hingga kegiatan produktif dalam industri dan bisnis, sangat tergantung pada pasokan listrik yang andal. Di Indonesia, PT Perusahaan Listrik Negara (PLN) menjadi satu-satunya penyedia utama listrik yang menjangkau hampir seluruh wilayah. Untuk

menjamin keadilan dalam akses terhadap energi listrik, pemerintah menetapkan kebijakan subsidi listrik bagi kelompok masyarakat berpenghasilan rendah (Jollyta et al., 2020).

Subsidi listrik biasanya diberikan kepada pelanggan dengan daya 450 VA dan 900 VA. Namun, dalam praktiknya, kebijakan ini belum sepenuhnya berjalan sesuai harapan. Banyak pelanggan tidak mampu yang justru tidak menerima subsidi, sementara pelanggan yang tergolong mampu masih menikmatinya. Penyebab utama dari permasalahan ini adalah ketidakakuratan dalam proses identifikasi dan klasifikasi pelanggan layak subsidi (Rahayu et al., 2021).

Klasifikasi adalah proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. Algoritma klasifikasi yang banyak digunakan secara luas, yaitu decision/classification trees, bayesian classifiers/naive bayes classifiers, neural networks, algoritma genetika, rough sets, k-nearest neighbor, metode rule based, memory based reasoning, dan support vector machines (SVM) (Retnoningsih & Pramudita, 2020). Klasifikasi data terdiri dari 2 langkah proses. Pertama adalah learning (fase training), dimana algoritma klasifikasi dibuat untuk menganalisa data training lalu direpresentasikan dalam bentuk rule klasifikasi. Proses kedua adalah klasifikasi, dimana data testing digunakan untuk memperkirakan akurasi dari rule klasifikasi (Huda et al., 2020)

Klasifikasi yang dilakukan selama ini masih bersifat manual atau berdasarkan data statis yang tidak diperbarui secara periodik. Dalam era digital dan data besar (big data), pendekatan tradisional semacam ini menjadi kurang relevan. Oleh karena itu, diperlukan metode yang mampu menangani data dalam jumlah besar dan menghasilkan klasifikasi yang cepat, efisien, dan akurat. Salah satu metode yang dapat digunakan adalah algoritma *Naïve Bayes* (Siregar et al., 2025).

Dari beberapa penelitian sebelumnya, *Naïve Bayes* telah diteliti oleh beberapa ahli, Pada penelitian yang dilakukan oleh (Purnamasari et al., 2018) yang berjudul "Performansi Klasifikasi Dosen Berprestasi Menggunakan Metode Naive Bayes Classifier" berdasarkan analisa hasil evaluasi data mining untuk perhitungan

performasi menggunakan klasifikasi dengan algoritma Naive Bayes menghasilkan nilai akurasi sebesar 91,96%. Dan hasil penelitian yang dilakukan oleh (Putri et al., 2021) yang berjudul “Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode Naive Bayes dan KNN ” menghasilkan klasifikasi penerima manfaat dari 6.491 penerima KKS dengan metode KNN menghasilkan nilai akurasi 66,18% dengan sebaran di 5 kelurahan sedangkan algoritma Naive Bayes terklasifikasi dengan true di masing-masing kelurahan dengan sebaran rute Argasanya 1.196 KKS clas precision 100% dan nilai akurasi algoritma Naive Bayes sebesar 99,88%. Hal ini menunjukkan nilai akurasi yang tinggi untuk penelitian menggunakan metode Naive Bayes.

Naïve Bayes merupakan salah satu teknik klasifikasi yang bersifat probabilistik dan banyak digunakan dalam bidang data mining dan machine learning. Metode ini bekerja berdasarkan prinsip Teorema Bayes dan mampu menangani data berukuran besar maupun kecil dengan performa yang baik. Keunggulannya terletak pada kecepatan komputasi dan kemampuannya menangani data dengan atribut kategorikal maupun numerik (Sriani et al., 2024).

Penelitian ini bertujuan untuk menerapkan metode *Naïve Bayes* dalam proses klasifikasi pelanggan listrik rumah tangga yang layak mendapatkan subsidi. Penelitian dilakukan di wilayah kerja PT. PLN Kotanopan, Kabupaten Mandailing Natal. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam peningkatan kualitas pelayanan publik, khususnya dalam distribusi subsidi yang lebih adil dan tepat sasaran.

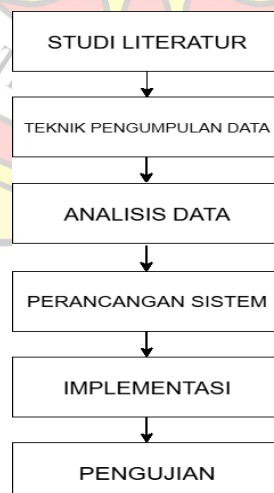
METODE PENELITIAN

Jenis penelitian ini adalah penelitian kuantitatif dengan pendekatan eksperimen komputasi, di mana model klasifikasi dibangun dan diuji menggunakan dataset pelanggan listrik rumah tangga. Fokus utamanya adalah pengujian performa algoritma *Naïve Bayes* dalam memprediksi kelayakan penerima subsidi listrik berdasarkan variabel-variabel yang telah ditentukan (Kawani, 2019).

Pendekatan ini dipilih karena mampu menguji keterandalan metode klasifikasi dalam kondisi data nyata yang mencerminkan karakteristik sosial ekonomi masyarakat (Adyaksa, 2022). Selain itu, pendekatan komputasi memungkinkan analisis dilakukan secara sistematis dan replikatif, sehingga hasilnya dapat dijadikan dasar bagi pengambilan keputusan yang lebih tepat dan terukur oleh pihak penyedia layanan listrik. Dengan metode ini, diharapkan sistem klasifikasi yang dibangun dapat memberikan kontribusi signifikan terhadap efisiensi distribusi subsidi dan peningkatan akurasi penentuan sasaran kebijakan.

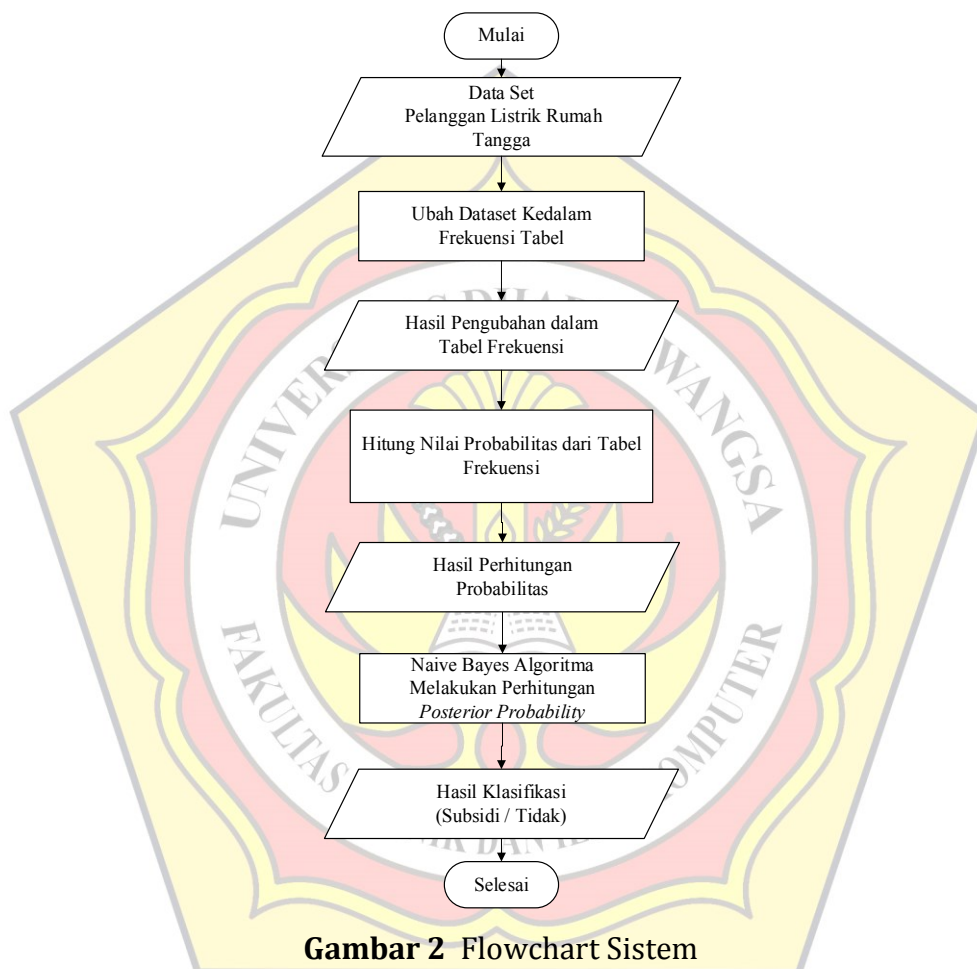
Data yang digunakan dalam penelitian terdiri atas dua jenis, yaitu data latih sebanyak 720 entri pelanggan rumah tangga yang telah diklasifikasikan berdasarkan kelayakan menerima subsidi, serta 500 data responden yang diperoleh melalui kuesioner daring (Google Form). Kuesioner tersebut berisi isian terkait status penerimaan subsidi dan sejumlah variabel yang berpengaruh, seperti daya listrik terpasang, status kepemilikan rumah, luas bangunan, jenis dinding, lantai, dan atap, jumlah penghuni rumah, pendapatan bulanan, jumlah peralatan listrik, serta rata-rata tagihan bulanan.

Adapun tahapan penelitian yang akan dilakukan dalam penelitian ini adalah sebagai berikut :



Gambar 1 Kerangka Penelitian

Tahapan selanjutnya adalah tahapan desain layout dan coding atau pembangunan aplikasi, hal-hal yang termasuk kedalam tahapan perancangan ini adalah proses pengetikan layout tampilan aplikasi, pengetikan program, penerapan metode dan logika program. Tahapan ini adalah gambaran tahapan seluruh rangkaian system dan proses penelitian yang digambarkan dalam bentuk flowchart dibawah ini.



Gambar 2 Flowchart Sistem

HASIL DAN PEMBAHASAN

Data yang dipergunakan dalam penelitian adalah data pelanggan yang menerima subsidi dari pemerintah terhadap pembayaran tagihan rekening listrik. Data ini diambil dari kuisioner yang dibuat dengan menggunakan google form dan disebarakan melalui media whatsapp, target responden dari kuesioner ini adalah pelanggan listrik yang berdomisili di Kabupaten Mandailing Natal dengan jumlah target populasi sebanyak 500 populasi (responden), dan pada dataset akan

menggunakan 720 data yang merupakan kondisi yang harus terpenuhi apakah pelanggan berhak mendapatkan subsidi listrik atau tidak.

Representasi dataset pada penelitian ini berjumlah total 720 data, yang berisikan kondisi atau syarat benar layak menerima subsidi atau tidak layak menerima subsidi. Dataset yang dipergunakan untuk menghasilkan data latih, dari 720 data yang dipergunakan disini akan ditampilkan sebanyak 10 data yang akan dipergunakan untuk melakukan proses kalkulasi klasifikasi Naive Bayes.

Tabel 1. Dataset

Daya Listrik	Subsidi Listrik	Tempat Tinggal	Luas Bangunan	Lantai Terluas	Dinding Terluas	Atap Terluas
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Tembok	Seng
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Tembok	Asbes
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Tembok	Genteng Tanah Liat
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Setengah Tembok	Seng
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Setengah Tembok	Asbes
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Setengah Tembok	Genteng Tanah Liat
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Kayu	Seng
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Kayu	Asbes
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Kayu	Genteng Tanah Liat
450 VA	Tidak	Bebas Sewa / Rumah Dinas	$\leq 36 \text{ m}^2$	Tanah	Bambu	Seng

Selain dataset, penelitian ini juga akan menggunakan data responden yang akan digunakan terhadap hasil pengujian data latih yang dihasilkan. Tabel 2 berikut ini adalah merupakan tabel dari data responden yang telah berhasil dikumpulkan, total data responden yang berhasil dikumpulkan adalah berjumlah 500 data responden, pada tabel 2 akan ditampilkan sebanyak 10 data.

Tabel 2. Tabel Data Responden

ID Pelanggan	Daya Listrik Terpasang	Memperoleh Subsidi Listrik	Kepemilikan Tempat Tinggal	Luas Bangunan	Lantai Terluas	Dinding Terluas	Atap Terluas
121410 085604	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Semen	Setengah Tembok	Seng
124110 013194	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Tanah	Kayu	Seng
124110 017915	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Tanah	Setengah Tembok	Seng
124110 027617	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Keramik	Tembok	Seng
124110 210051	450 VA	Ya	Milik Sendiri	$36 \text{ m}^2 < \text{Luas} \leq 70 \text{ m}^2$	Keramik	Tembok	Seng
124110 028543	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Semen	Tembok	Seng
124110 019666	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Keramik	Tembok	Seng
124110 030571	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Semen	Tembok	Seng
124110 119328	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Semen	Tembok	Seng
124110 231292	450 VA	Ya	Milik Sendiri	$\leq 36 \text{ m}^2$	Semen	Tembok	Seng

Tahap selanjutnya dalam proses ini adalah melakukan kalkulasi dengan menggunakan metode Naïve Bayes Classifier untuk menentukan membangun data latih kriteria kelayakan pelanggan yang menerima subsidi listrik. Selanjutnya dilakukan perhitungan probabilitas kelas pada masing-masing kelas, baik pada kelas dengan target Ya dan target Tidak. Perhitungan yang akan dilakukan terkait dengan hasil dari tabel data latih, selanjutnya dilakukan perhitungan sebagai berikut :

1. Menghitung total data latih, dalam pelatihan data yang dipergunakan berjumlah 720 data.
2. Menghitung total jumlah variabel Daya Terpasang, pada dengan label target “Ya” dan target “Tidak” pada masing-masing label yaitu 450 VA dan 900 VA. Dari tabel

data latih diperoleh total nilai daya listrik terpasang pada masing label sebagai berikut :

Label 450 VA = 432 data

Label 900 VA = 288 data

3. Menghitung probabilitas pada masing-masing label dengan membagi total data masing-masing label dengan seluruh total data dengan persamaan :

$$= \frac{\sum \text{Total Data Ya (450 VA)}}{\sum \text{Total Data}}$$

$$= \frac{432}{720} = 0.58333$$

$$= 0.6 \text{ (pembulatan)}$$

$$= \frac{\sum \text{Total Data Ya (900 VA)}}{\sum \text{Total Data}}$$

$$= \frac{288}{720} = 0.4$$

Sehingga dari perhitungan tersebut diperoleh nilai probabilitas untuk label 450 VA = 0.6 dan label 900 VA = 0.4

4. Melakukan perhitungan yang sama terhadap seluruh nilai variabel nilai yang ada. Untuk nilai perhitungan keseluruhan probabilitas kelas dapat dilihat pada table berikut ini :

Tabel 3. Tabel Probabilitas

Atribut / Variable		Jumlah	YA	TIDAK	P (X)C	
		Data			YA	TIDAK
Total		720	432	288	0.6	0.4
X1	450 VA	504	288	216	0.571429	0.428571
	900 VA	216	144	72	0.666667	0.333333
X2	Bebas Sewa / Rumah Dinas	144	0	144	0	1
	Sewa / Kontrak	288	288	0	1	0
	Milik Sendiri	288	216	72	0.75	0.25
X3	<=36 m2	180	144	36	0.8	0.2
	<=36 m2 Luas <=70 m2	180	144	36	0.8	0.2

	70 m2	360	216	144	0.6	0.4
X4	Tanah	240	168	72	0.7	0.3
	Semen	240	168	72	0.7	0.3
	Keramik	240	168	72	0.7	0.3
X5	Tembok	180	126	54	0.7	0.3
	Setengah Tembok	180	126	54	0.7	0.3
	Kayu	180	126	54	0.7	0.3
	Bambu	180	126	54	0.7	0.3
X6	Seng	240	168	72	0.7	0.3
	Asbes	240	168	72	0.7	0.3
	Genteng Tanah Liat	240	168	72	0.7	0.3

Setelah tahapan kalkulasi probabilitas priority, selanjutnya dilakukan perhitungan terhadap data responden yang tertera pada tabel 4.3 ,terhadap probabilitas ya dan tidak. Berikut akan dilakukan perhitungan dengan menggunakan salah satu data responden sebagai berikut :

1. Classifier

Pada tahapan ini akan dijelaskan teknik-teknik yang akan digunakan dalam pengklasifikasian data, pertama hitung setiap probabilitas kategori dan probabilitas setiap kelas dari data latih.

$$P(C|X) = \frac{P(X|C)}{P(X)}$$

Keterangan :

$P(C|X)$ = Probabilitas hipotesis C berdasarkan kondisi X (Posteriori Probability)

$P(X|C)$ = Probabilitas X berdasarkan hipotesis C

$P(X)$ = Probabilitas dari X

X = Data dengan class yang belum diketahui.

C = Hipotesis data X yang merupakan class spesifik

a. Melakukan perhitungan Probabilitas Prior

Pada perhitungan ini akan dipergunakan data sampel sebanyak 144 data, yang diperoleh dari data responden.

1. Hitung jumlah target setiap kelas.

Atribut Target	Jumlah Data	Probabilitas
Ya	108	0.82
Tidak	25	0.08

2. Hitung *Confusion Matrix*

		Kelas	
		Layak (Ya)	Tidak Layak (Tidak)
Prediksi	True	108	0
	False	11	25

- 1) Accuracy

$$\frac{TP + TN}{TP + TN + FN + FP} \times 100\% = \frac{108 + 25}{108 + 11 + 0 + 25} = 92\%$$

- 2) Recall

$$\frac{TP}{TP + FN} \times 100\% = \frac{108}{108 + 11} \times 100\% = 90.9\%$$

- 3) Precision

$$\frac{TP}{TP + FP} \times 100\% = \frac{108}{108 + 0} \times 100\% = 100\%$$

Implementasi

Pada pengujian yang akan dilakukan terhadap metode Naive Bayes data akan dikalkulasikan dengan menggunakan bahasa pemrograman python menggunakan IDE Jupyter Notebook. Sebelum masuk kedalam tahap pengujian,

data latih yang berjumlah 720 data akan diubah terlebih dahulu kedalam bentuk data integral.

```
#Import Library Yang dibutuhkan
#import panda untuk panggil file excel
import pandas as pd
import numpy as np
```

Gambar 3. Library yang di butuhkan

Gambar diatas adalah sebuah kode untuk membaca dan mengelola data (biasanya dari file Excel/CSV) dan melakukan perhitungan matematis atau statistik yang diperlukan dalam pemrosesan data, seperti yang dilakukan dalam proyek klasifikasi dengan Naïve Bayes

Daya_Listrik_Terpasang	Memperoleh_Subsidi_Listrik	Kepemilikan_Tempat_Tinggal	Luas_Bangunan	Jenis_Lantai_Terluas	Jenis_Dinding_Terluas
0	0	1	0	0	0
1	0	1	0	0	0
2	0	1	0	0	0
3	0	1	0	0	1
4	0	1	0	0	1

Gambar 4. Tampilan data 5 teratas

Daya_Listrik_Terpasang	Kepemilikan_Tempat_Tinggal	Luas_Bangunan	Jenis_Lantai_Terluas	Jenis_Dinding_Terluas
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	1
4	0	0	0	1
...
715	1	2	2	2
716	1	2	2	2
717	1	2	2	4
718	1	2	2	4
719	1	2	2	4

720 rows × 5 columns

Gambar 5. Output Data variable x

Setelah proses pemilihan variabel x selesai, selanjutnya adalah menampilkan data dari kolom “Memperoleh Subsidi Listrik”. Perlu diingat bahwa syntax python bersifat case sensitive, sehingga penulisan huruf besar dan kecil akan sangat berpengaruh dalam proses perhitungan.

Fungsi naive bayes gaussian ini disediakan oleh library Sklearn, sehingga perlu dilakukan pemanggilan library sklearn sebelum memanggil fungsi Naive Bayes Gaussian, penulisan syntax ini dapat dilihat pada gambar 4.6 berikut ini:

```
#Panggil Naive bayes dari Sklearn
from sklearn.naive_bayes import GaussianNB

#Panggil fungsi Naive Bayes
nbc = GaussianNB()

##Data training Untuk Klasifikasi Naive Bayes
data_training=nbc.fit(x,y)
```

Gambar 6. Tampilan fungsi Gaussian dari sklearn.

Pengujian

Evaluasi dilakukan untuk mengukur kinerja model secara menyeluruh melalui sejumlah metrik, yaitu akurasi, presisi, recall, dan f1-score. Proses ini bertujuan untuk mengetahui sejauh mana model *Naïve Bayes* mampu menghasilkan prediksi yang akurat dan konsisten terhadap data uji yang diberikan. Hasil evaluasi ini menjadi acuan penting dalam menilai efektivitas model serta memberikan dasar bagi pengembangan dan penyempurnaan sistem ke depannya (Syahrudin & Kurniawan, 2018). Berdasarkan implementasi metode *Naïve Bayes*, diperoleh hasil evaluasi sebagai berikut.

```
##Melakukan Klasifikasi Berdasarkan Data Inputan
y_predict = data_training.predict(Data_Test)
#print(y_klasifikasi)

c:\python37\lib\site-packages\sklearn\base.py:451:
with feature names
"X does not have valid feature names, but"
```

Gambar 7. Pengujian Input data

Proses input yang dilakukan pada gambar ini selanjutnya akan diproses dengan melakukan prediksi atau klasifikasi dengan menggunakan syntax pada gambar

berikutnya syntax ini berfungsi untuk melakukan kalkulasi terhadap data input. Terlihat puna tanda warning berwarna merah saat syntax di eksekusi. Tanda warning ini tampil dikarena jumlah masukan atau input yang diisikan hanya berjumlah 1 dan hanya menghasilkan matrix 1 dimensi, sedangkan data latih yang sudah dilakukan sebelumnya menggunakan matrix 2 dimensi.

```
##Cetak Hasil Klasifikasi Berdasarkan Data Inputan
## Jika Hasil == 0 "Klasifikasi Subsidi Tidak Layak (Tidak)"
## Jika Hasil == 1 "Klasifikasi Subsidi Layak (Ya)"
if y_predict == 1 :
    Hasil = "Tidak Layak Subsidi Listrik"
elif y_predict == 0 :
    Hasil = "Layak Subsidi Listrik"
else :
    Hasil = "Tidak Diketahui"

print ("Hasil Klasifikasi Subsidi Listrik =", Hasil)
```

Hasil Klasifikasi Subsidi Listrik = Layak Subsidi Listrik

Gambar 8. Output Hasil Pengujian Responden

Perhitungan akurasi ini bertujuan untuk mendapatkan informasi seberapa akurat proses pelatihan yang telah dilakukan terhadap data pengujian. Data yang dikalkulasi disini adalah data latih yang berjumlah 720 data. Data ini akan dibagi menjadi dua bagian sehingga hanya akan menghitung 360 data yang terdiri dari 180 data uji dan 180 data latih, proses pemilihan data ini akan dilakukan secara random dengan kondisi random state 0, yang artinya proses random akan dilakukan setiap kali data ditampilkan.

```
from sklearn.metrics import classification_report
#Cetak Klasifikasi Report
print (classification_report(y_test, nbc.predict(x_test)))
```

	precision	recall	f1-score	support
0	0.91	1.00	0.95	108
1	1.00	0.69	0.82	36
accuracy			0.92	144
macro avg	0.95	0.85	0.89	144
weighted avg	0.93	0.92	0.92	144

Gambar 8. Report Hasil Kalkulasi

Berdasarkan hasil evaluasi yang ditampilkan melalui classification report, model Naïve Bayes berhasil mencapai akurasi sebesar 92% terhadap data uji, yang menunjukkan performa klasifikasi yang cukup tinggi. Untuk kelas 0, model menunjukkan recall sempurna (1.00) dan f1-score sebesar 0.95, sedangkan untuk kelas 1, meskipun precision mencapai 1.00, recall-nya hanya 0.69 sehingga f1-score-nya menurun menjadi 0.82. Hal ini mengindikasikan bahwa model sangat baik dalam mengenali kelas mayoritas, tetapi masih memiliki keterbatasan dalam mendeteksi kelas minoritas. Oleh karena itu, meskipun performa keseluruhan cukup memuaskan, diperlukan penyempurnaan lebih lanjut agar klasifikasi pada kedua kelas menjadi lebih seimbang.

KESIMPULAN

Dari hasil penelitian yang telah dilakukan, dihasilkan kesimpulan sebagai berikut :

1. Untuk menghasilkan data latih yang baik harus menggunakan data yang relatif banyak, hal ini bertujuan untuk menghasilkan nilai akurasi pengujian yang baik.
2. Pada penelitian ini digunakan data latih dengan jumlah 720 data dengan kondisi variabel yang sesuai dengan aturan pemberian subsidi listrik.
3. Dalam proses pelatihan dilakukan splitting data dari total data yang dipergunakan, proses splitting data ini bertujuan untuk menghasilkan uji dan data uji secara random.
4. Dalam hasil kalkulasi terhadap data uji dan data latih pada penelitian diperoleh nilai akurasi sebesar 92%, nilai recall 92% dan precision 100% sehingga dapat dikatakan hasil penelitian ini memiliki nilai akurasi yang sangat baik dalam melakukan pengujian.

SARAN

Pada penelitian ini proses pengolahan data dilakukan secara manual tidak berjalan secara otomatis, sehingga perlu adanya sebuah aplikasi utuh yang memiliki form inputan terhadap data uji, data latih dan form pengujian, sehingga memudahkan siapapun yang tidak memiliki keahlian pemrograman untuk menggunakan aplikasi dari hasil penelitian ini

DAFTAR PUSTAKA

- Adyaksa, Iwa (2022). *Prototype Monitoring Dan Alarm Pendeteksi Banjir Dengan Internet of Things Menggunakan Metode Fuzzy Sugeno Berbasis Mikrokontroler*. 11(2), 154–166.
- Huda, A. S., Awangga, R. M., & Fathonah, N. S. (2020). *Prediksi Penerimaan Pegawai Baru Dengan Metode Naive Bayes*. Kreatif Industri Nusantara.

- Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). Konsep Data Mining Dan Penerapan (1st ed.). CV Budi Utama.
- Kawani, G. P. (2019). Implementasi Naive Bayes. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 1(2), 73–81. <https://doi.org/10.20895/inista.v1i2.73>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Purnamasari, I., Afnisari, K., & Sitasi, C. (2018). Performansi Klasifikasi Dosen Berprestasi Menggunakan Metode Naive Bayes Classifier. *Paradigma*, XX(2), 45–50. <https://doi.org/10.31294/p.v20i2.3788>
- Putri, H., Purnamasari, A. I., Dikananda, A. R., Nurdiawan, O., & Anwar, S. (2021). Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN. *Building of Informatics, Technology and Science (BITS)*, 3(3), 331–337. <https://doi.org/10.47065/bits.v3i3.1093>
- Rahayu, W. I., Prianto, C., & Novia, E. A. (2021). Perbandingan Algoritma K-Means dan Naive Bayes untuk Memprediksi Prioritas Pembayaran Tagihan Rumah Sakit Berdasarkan Tingkat Kepentingan pada PT. Pertamina (Persero). *Jurnal Teknik Informatika*, 13(2), 1–8.
- Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*, 7(2), 156. <https://doi.org/10.51211/biict.v7i2.1422>
- Rosaly, R., & Prasetyo, A. (2019). Pengertian Flowchart Beserta Fungsi dan Simbol-simbol Flowchart yang Paling Umum Digunakan. <https://www.nesabamedia.com>, 2, 2.
- Siregar, R. G., Lubis, A. H., & Ikhsan, M. (2025). Klasifikasi Pengaruh Negatif Game online Bagi Remaja Menggunakan Algoritma Naïve Bayes. *Device : Journal of Information System, Computer Science and Information Technology*, 6(1), 232–244. <https://doi.org/10.46576/device.v6i1.6655>
- Sriani, Suhardi, & Saufa Yardha, L. (2024). Analisis Sentimen Pengguna Aplikasi Mobile Jkn Menggunakan Algoritma Naïve Bayes. *Journal of Science and Social Research*, 4307(2), 555–563. <http://jurnal.goretanpena.com/index.php/JSSR>
- Syahrudin, A. N., & Kurniawan, T. (2018). Input dan Output pada Bahasa Pemrograman Python. *Jurnal Dasar Pemrograman Python STMIK*, June 2018, 1–7. <https://www.researchgate.net/publication/338385483>