

Implementation of Data Mining for Classifying Electricity Subsidy Recipients in Stabat District Using the Naive Bayes Algorithm

by Laila Nurhidayah

Implementation of Data Mining for Classifying Electricity Subsidy Recipients in Stabat District Using the Naive Bayes Algorithm

Laila Nurhidayah^{1*)}, Ali Ikhwan²⁾

¹⁾²⁾Program Studi Sistem Informasi, Universitas Islam Negeri Sumatera Utara

^{*)}Correspondence author : lailanurhidayah191@gmail.com, Medan, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i2.3852>

Abstract

The electricity subsidy assistance program is a government initiative designed to enable low-income households to access electricity at reduced rates. Although the national data system is well-integrated, the economic circumstances of the population are dynamic. Such changes often lead to discrepancies between centrally recorded data and the actual situation on the ground, creating a risk that subsidies may not reach the intended recipients. Without an objective preliminary screening system, data submissions risk rejection and the accumulation of backlogged paperwork. This study aims to implement a web-based application serving as a local early screening tool to classify eligibility by automatically, rapidly, and measurably ranking the priority of prospective subsidy recipients; it employs a quantitative approach utilizing data mining techniques and the Naive Bayes algorithm. The Waterfall method was applied for software development. The results demonstrate that the developed system is capable of classifying the eligibility status of electricity subsidy recipients, specifically for 900 VA household customers. This research is expected to assist relevant agencies in identifying the household electricity customers entitled to receive the assistance.

Keywords: Classification, Data Mining, Naïve Bayes, Electricity Subsidy

Abstrak

Program bantuan subsidi listrik merupakan upaya yang dilakukan pemerintah kepada masyarakat kurang mampu agar dapat menikmati listrik dengan tarif listrik yang lebih rendah. Meskipun sistem pendataan nasional sudah terintegrasi dengan baik, kondisi ekonomi masyarakat di lapangan bersifat dinamis. Perubahan ini sering menyebabkan ketidaksesuaian antara data yang tercatat di pusat dengan kondisi aktual masyarakat, sehingga penyaluran subsidi berisiko tidak tepat sasaran. Tanpa adanya sistem penyaring awal yang objektif, pengusulan data pengusulan data berisiko mengalami penolakan dan menyebabkan terjadinya penumpukan berkas. Penelitian ini bertujuan untuk mengimplementasikan aplikasi berbasis web sebagai *early screening tool* lokal untuk mengklasifikasikan tingkat kualifikasi kelayakan dengan memeringkatkan prioritas calon penerima subsidi listrik secara otomatis, cepat, dan terukur menggunakan pendekatan kuantitatif dengan teknik data mining dan algoritma Naive Bayes. Untuk pengembangan perangkat lunak, penelitian ini menerapkan metode Waterfall. Hasil penelitian menunjukkan bahwa sistem yang dibangun mampu mengklasifikasikan status kelayakan penerima subsidi listrik terkhusus pelanggan rumah tangga 900 VA. Penelitian ini diharapkan dapat membantu instansi terkait dalam menentukan pelanggan listrik rumah tangga yang berhak menerima bantuan.

Kata Kunci: Klasifikasi, Data Mining, Naïve Bayes, Subsidi Listrik

PENDAHULUAN

Listrik merupakan salah satu kebutuhan yang mendasar bagi masyarakat, baik dalam aktivitas rumah tangga maupun industri (Almira & Ikhwan, 2021). Untuk menjaga keberlangsungan ekonomi dan meringankan beban masyarakat kurang mampu, PLN (Perusahaan Listrik Negara) memberikan subsidi listrik sesuai dengan Peraturan Menteri Energi dan Sumber Daya Mineral (ESDM) Republik Indonesia Tahun 2022 mengenai Pelaksanaan Usaha Ketenagalistrikan (Utiahman et al., 2025). Subsidi diberikan dalam bentuk tarif listrik yang lebih rendah dan bergantung pada jumlah pemakaian listrik setiap bulannya.

Namun dalam pelaksanaannya, pemberian subsidi listrik masih menghadapi permasalahan. Masih terdapat masyarakat yang sebenarnya mampu tetapi mendapatkan bantuan. Sebaliknya, masyarakat yang kurang mampu dan berhak mendapatkan subsidi, tetapi belum menerima bantuan. Hal ini terjadi karena data yang digunakan belum diperbarui secara rutin atau masih menggunakan data lama yang tidak lagi menggambarkan kondisi ekonomi masyarakat saat ini. Selain itu, dalam proses penyaluran di tingkat lokal, petugas menghadapi keterbatasan periode pengusulan, minimnya personel lapangan, dan ketatnya verifikasi tingkat desil menuju sistem Data Tunggal Sosial Ekonomi Nasional (DTSEN). Tanpa adanya sistem penyaring awal yang objektif, pengusulan data beresiko mengalami penolakan dan menyebabkan terjadinya penumpukan berkas.

Untuk mengatasi permasalahan tersebut, teknologi data mining dapat digunakan untuk meningkatkan ketepatan dan efektivitas proses klasifikasi yang memenuhi syarat untuk mendapatkan subsidi listrik. Penelitian ini menerapkan teknik data mining menggunakan algoritma Naïve Bayes untuk mengklasifikasikan data pelanggan ke dalam kategori layak dan tidak layak. Data yang digunakan berjumlah 1000 data pelanggan listrik 900 VA yang diperoleh melalui PT. PLN ULP Stabat serta Kelurahan dan Kantor Desa di Kecamatan Stabat. Data tersebut merupakan data terbaru yang disesuaikan dengan kondisi sosial ekonomi pelanggan saat ini yang terdiri dari 10 atribut sebagai dasar dalam proses klasifikasi.

Data mining merupakan teknik untuk memperoleh informasi penting dari sejumlah data besar yang melibatkan penggunaan pengetahuan di bidang statistik, matematika, dan pengenalan pola (Damuri et al., 2021). Data mining, sebagai teknik analisis tingkat lanjut mampu mengklasifikasikan penerima yang berhak secara akurat dan efisien (Pahlevi et al., 2023). Perancangan aplikasi klasifikasi penerima subsidi listrik berbasis web ini

memungkinkan petugas atau operator lokal untuk melakukan verifikasi awal secara fleksibel dan *real-time*. Sistem ini juga berfungsi sebagai *smart filter* lokal untuk memeringkat dan merekomendasikan pendaftar yang paling berhak sesuai dengan prioritas secara otomatis, cepat, dan akurat.

Penelitian sebelumnya menunjukkan bahwa algoritma Naïve Bayes dapat ⁵digunakan dalam proses klasifikasi penerima bantuan. Penelitian oleh Amanda Sari dkk. Dengan menggunakan Naïve Bayes untuk menentukan kelayakan penerima Program Indonesia Pintar atau PIP. Hasil menunjukkan, berdasarkan data 207 siswa algoritma tersebut memperoleh tingkat akurasi sebesar 89%, sehingga dapat membantu proses seleksi penerima bantuan berdasarkan kriteria yang ditentukan (Sari et al., 2024). Selanjutnya, penelitian Zulfa dan Atifah menerapkan algoritma ⁵Naïve Bayes dalam ⁸mengklasifikasikan kelayakan penerima Bantuan Langsung Tunai Dana Desa (BLT-DD). Hasil pengujian menunjukkan bahwa metode Bernoulli Naïve Bayes memperoleh akurasi tertinggi sebesar 91%, sedangkan Gaussian Naïve Bayes memperoleh akurasi sebesar 90%. Hasil dari kedua penelitian tersebut menunjukkan bahwa algoritma Naïve Bayes mampu memberikan hasil yang baik dalam mengklasifikasikan kelayakan penerima bantuan.

Berdasarkan penelitian yang telah dilakukan sebelumnya, terdapat ruang untuk pengembangan lebih lanjut mengenai klasifikasi subsidi listrik menggunakan metode klasifikasi Naïve Bayes. Oleh karena itu, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan metode klasifikasi yang lebih tepat sasaran dan efisien untuk memprediksi penerimaan bantuan listrik di Kecamatan Stabat.

METODE PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif dengan memanfaatkan teknik data mining dan algoritma Naive Bayes untuk mengklasifikasikan kelayakan pelanggan listrik berdasarkan atribut sosial dan ekonomi yang telah ditentukan. Pendekatan kuantitatif digunakan karena penelitian ini mengolah data dan melakukan perhitungan probabilitas untuk menghasilkan keputusan yang objektif (Subhaktiyasa, 2024).

Tahapan data mining dilakukan ⁹menggunakan pendekatan Knowledge Discovery in Database (KDD) yang terdiri atas beberapa tahapan, yaitu sebagai berikut.

1. Data *selection*, pada tahap ini dipilih 1000 data pelanggan listrik 900 VA yang sesuai dengan kebutuhan penelitian. Dari data tersebut digunakan 10 atribut sosial dan ekonomi sebagai dasar penentuan kelayakan calon penerima subsidi listrik.
2. Data *preprocessing*, dilakukan untuk memastikan data dalam kondisi yang lengkap, konsisten, dan siap digunakan, yang meliputi pemeriksaan data kosong, data ganda, kesalahan penulisan, dan ketidaksesuaian format.
3. Data *transformation*, dilakukan dengan menyesuaikan format atribut agar siap diproses oleh algoritma Naïve Bayes tanpa mengubah informasi penting di dalamnya.
4. Data mining, tahap ini menerapkan algoritma Naïve Bayes untuk mengklasifikasikan kelayakan pelanggan berdasarkan 10 atribut yang digunakan. Data dibagi menjadi data latih dan data uji untuk membangun serta menguji model.
5. Evaluation, tahapan dilakukan untuk mengukur kinerja model dengan menggunakan *confusion matrix* dengan parameter *accuracy*, *precision*, *recall*, dan *F-1 score*. Pelanggan yang dinyatakan layak kemudian diurutkan berdasarkan probabilitas layak tertinggi sebagai prioritas verifikasi.

Metode Algoritma Naïve Bayes

Naive Bayes merupakan salah satu metode klasifikasi berbasis probabilitas dan statistik (Suganda et al., 2022). Klasifikasi merupakan suatu proses dalam data mining yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu (Susana, 2022). Berikut merupakan perhitungan klasifikasi menggunakan rumus teorema Bayes:

$$P(C|X_1, X_2, \dots, X_n) = P(C)P(X_1|C)P(X_2|C) \dots P(X_n|C)$$

Keterangan :

$P(C|X_1, X_2, \dots, X_n)$: Probabilitas data termasuk dalam kelas tertentu.

$P(C)$: Probabilitas awal kelas layak atau tidak layak.

$P(X_i|C)$: Probabilitas atribut ke-i terhadap kelas.

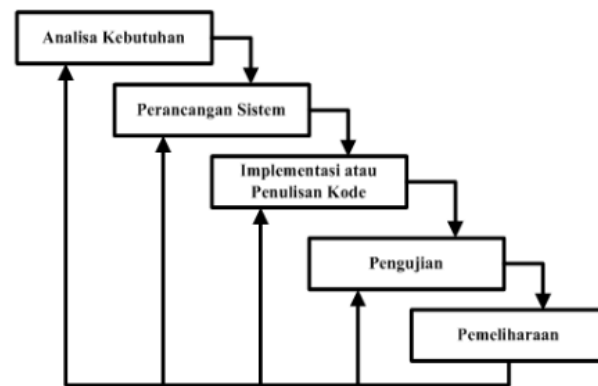
X_i : Atribut ke-i yang digunakan dalam klasifikasi.

n : Jumlah atribut yang digunakan.

Metode Pengembangan Sistem

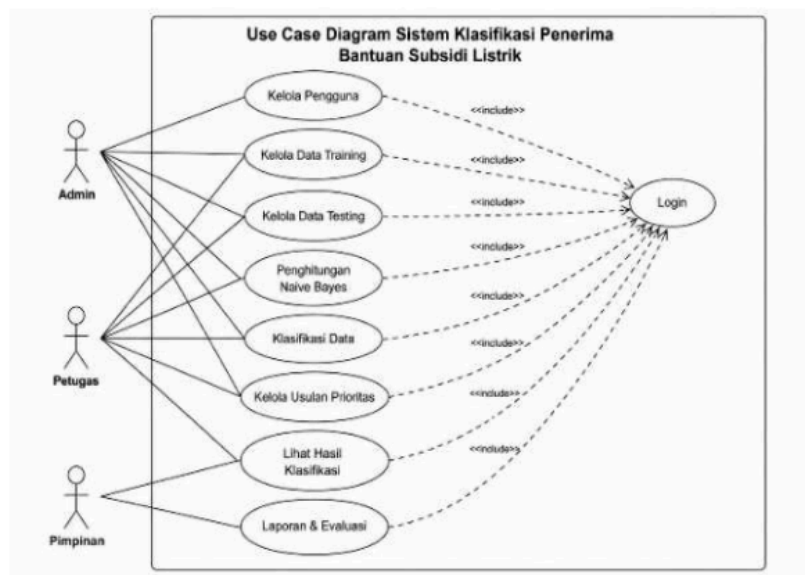
Pengembangan sistem dilakukan dengan menggunakan metode *Waterfall*. Metode *waterfall* merupakan metode dalam pengembangan perangkat lunak klasik yang memiliki pendekatan sederhana dan terstruktur, yang mana setiap tahapannya dilakukan secara berurutan

(Nurwijayanti & Syahrullah Syahrullah, 2023). Berikut merupakan tahapan dalam metode *Waterfall*.



Gambar 1. Metode *Waterfall*

- Analisis kebutuhan, tahap analisis kebutuhan dilakukan identifikasi kebutuhan sistem berdasarkan permasalahan yang terjadi pada penyaluran program subsidi listrik di Kecamatan Stabat, khususnya mengenai ketidaktepatan sasaran pada pelanggan dengan daya 900 VA. Proses analisis dilakukan dengan mengumpulkan kebutuhan pengguna dan menyesuaikannya dengan tujuan sistem, yaitu membantu proses klasifikasi penerima subsidi listrik secara objektif.
- Perancangan sistem dilakukan dengan menggunakan perangkat pemodelan seperti UML (*Unified Modelling Language*) sehingga alur kerja dan hubungan antar komponen dapat dipahami dengan lebih mudah.

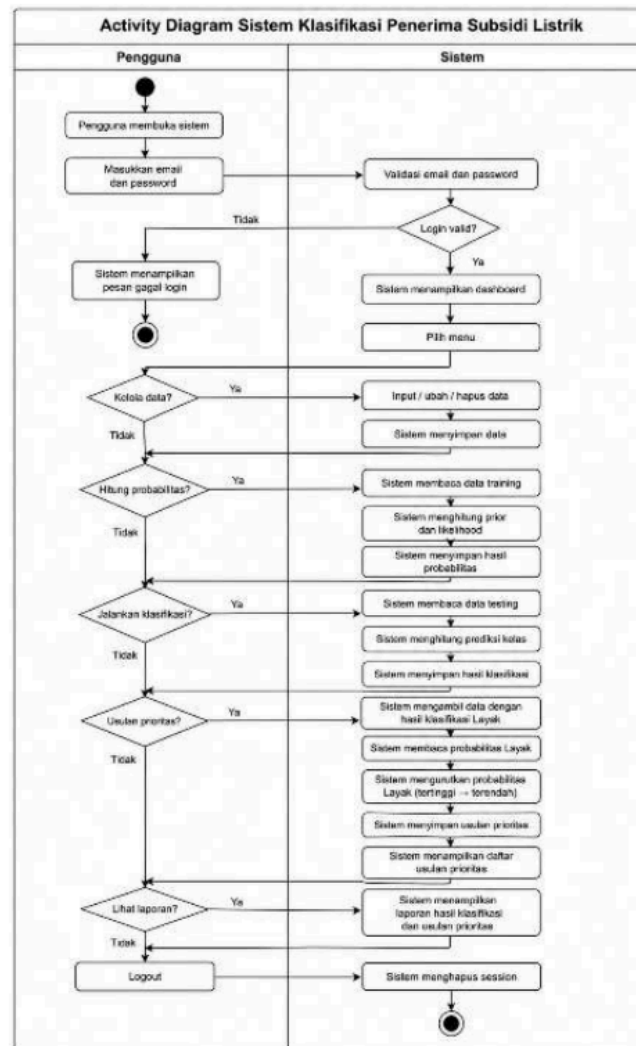


7

Gambar 2. Use Case Diagram

4

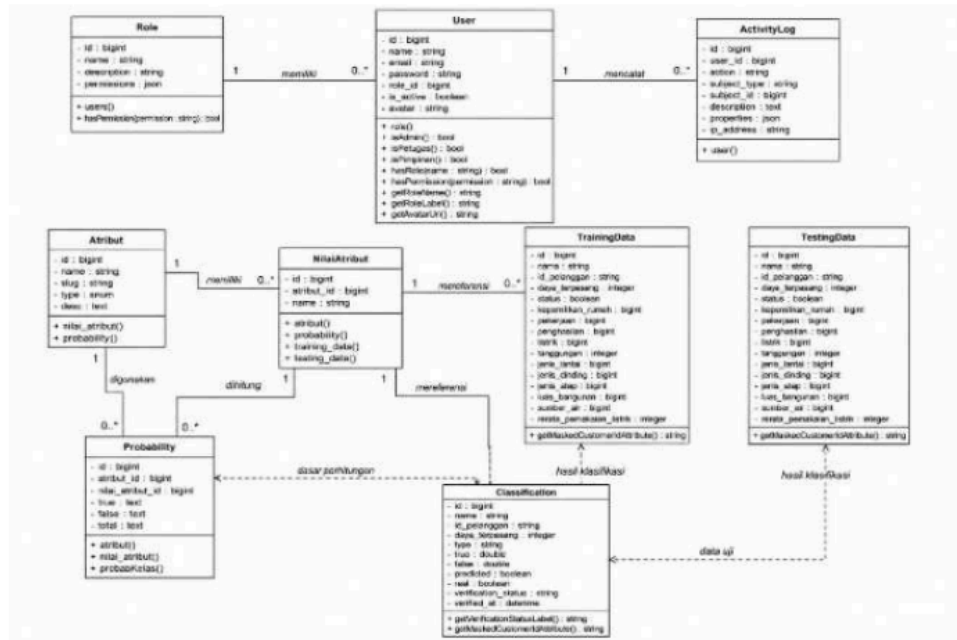
Use case diagram pada gambar 2 diatas, digunakan untuk memberikan gambaran umum mengenai seluruh elemen dan fungsi utama yang tersedia pada sistem (Pratama et al., 2022). Pada sistem yang dibuat, use case diagram berfungsi untuk menunjukkan hak akses dan aktivitas yang dapat dilakukan oleh setiap pengguna yaitu admin, petugas, dan pimpinan.



7
Gambar 3. Activity Diagram

Activity diagram pada gambar 3 diatas, digunakan untuk memodelkan alur aktivitas yang dilakukan oleh pengguna dari proses awal hingga akhir (KOMHARUDIN & Aura, 2025). Proses dimulai dari login yang divalidasi oleh sistem. Kemudian admin masuk ke halaman utama dan memilih menu klasifikasi. Pada tahap ini admin dapat memilih metode input, yaitu mengunggah file untuk memproses banyak data sekaligus atau mengisi form secara manual untuk data calon penerima. Setelah data dikirimkan, sistem

memprosesnya menggunakan algoritma Naive Bayes, mengambil data latih dari basis data, menghitung probabilitas, dan menentukan apakah pelanggan termasuk dalam kategori Layak atau Tidak Layak. Setelah proses klasifikasi selesai, admin dapat menjalankan usulan prioritas. Admin juga dapat mengakses menu riwayat untuk melihat daftar dan detail hasil perhitungan, sebelum keluar dan sistem menutup sesi untuk menyelesaikan proses.



Gambar 4. Class Diagram

Class Diagram merupakan diagram UML yang digunakan untuk menggambarkan struktur statis suatu sistem yang terdiri dari kelas, atribut, metode, dan hubungan antar kelas. Diagram ini berfungsi sebagai representasi rancangan basis data dan objek yang digunakan dalam pengembangan perangkat lunak. Gambar 4 diatas Merupakan Class Diagram pada sistem klasifikasi penerima bantuan subsidi listrik.

c. Implementasi

Pada tahap implementasi, desain sistem yang telah dibuat kemudian diterapkan ke dalam bentuk program atau aplikasi. Sistem yang dibangun berbasis website dengan menggunakan bahasa pemrograman PHP dan framework Laravel (Fuad & Asia, 2024).

d. Pengujian

Pada tahap pengujian bertujuan untuk memastikan semua komponen sistem dapat bekerja sama dengan baik. Selama tahap pengujian sistem, penulis menggunakan pendekatan Black Box Testing (Abdullah et al., 2022).

e. Pemeliharaan

Tahap ini meliputi proses pemeliharaan dan pengembangan lanjutan. Sistem dapat disesuaikan atau ditingkatkan sesuai dengan kebutuhan pengguna yang mungkin akan berubah seiring waktu. Dengan adanya evaluasi dan pemeliharaan yang berkelanjutan, sistem akan mampu mendukung kebutuhan pengguna dalam jangka panjang dan tetap

HASIL DAN PEMBAHASAN

Perhitungan Algoritma Naïve Bayes

Naïve Bayes Classifier adalah metode klasifikasi dengan berdasarkan nilai probabilitas. Algoritma ini menghitung probabilitas tertinggi dari berbagai kelas untuk menentukan klasifikasi akhir. Data yang peneliti gunakan dalam penelitian ini berjumlah 1000 data yang terdiri dari data testing sebanyak 800 data dan data uji sebanyak 200 data. Adapun untuk sampel perhitungan manual peneliti mengambil 10 data sebagai data latih dan 2 data sebagai uji secara acak.

Tabel 1.
Sampel Beberapa Dataset untuk Data Latih

Nama	Pekerjaan	Penghasilan	Jumlah Tanggungan	Kepemilikan Rumah	Jenis Lantai	Jenis Dinding	Jenis Atap	Luas Bangunan	Sumber Air	Pemakaian Listrik (kWh)	Status
Juraida	Belum/Tidak Bekerja	≤1.000.000	2	Milik Sendiri	Semen	Tembok	Seng	>36 - 60 m2	Sumur	>50- 100	Layak
Sri Diana	Karyawan Swasta	>1.000.000 - 1.800.000	3	Milik Sendiri	Keramik	Kayu/Papan	Seng	>60 - 90 m2	Sumur	>100- 150	Layak
Sunardi	Buruh Harian Lepas	>1.000.000 - 1.800.000	1	Sewa	Keramik	Tembok	Seng	>60 - 90 m2	Sumur	>100- 150	Tidak Layak
Samsul	Petani	>1.000.000 - 1.800.000	2	Milik Sendiri	Keramik	Kayu/Papan	Seng	>60 - 90 m2	Sumur	>50- 100	Layak
Eko Purnama	Petani	≤1.000.000	3	Milik Sendiri	Keramik	Tembok	Asbes	>36 - 60 m2	PDAM	>50- 100	Layak
Fitriadi	Petani	>3.200.000	1	Milik Sendiri	Keramik	Kayu/Papan	Seng	>36 - 60 m2	PDAM	>150	Tidak Layak
Rubiadi	Buruh Harian Lepas	>1.000.000 - 1.800.000	2	Milik Sendiri	Semen	Tembok	Asbes	>36 - 60 m2	Sumur	≤50	Layak
Siti Khadijah	Pedagang	>1.800.000 - 2.500.000	2	Milik Sendiri	Keramik	Tembok	Seng	>60 - 90 m2	Sumur	>100- 150	Tidak Layak
Julianto	Buruh Harian Lepas	>1.000.000 - 1.800.000	3	Sewa	Keramik	Tembok	Seng	>36 - 60 m2	PDAM	>100- 150	Layak
Hariani	Wiraswasta	>2.500.000 - 3.200.000	2	Milik Sendiri	Keramik	Tembok	Seng	>60 - 90 m2	Sumur	>50- 100	Tidak Layak

Tabel 2.
Sampel Beberapa Dataset untuk Data Uji

	Mualip	Rahmat Faisal
Pekerjaan	Karyawan Swasta	Buruh Harian Lepas
Penghasilan	>3.200.000	>1.000.000 - 1.800.000
Jumlah Tanggungan	2	3
Kepemilikan Rumah	Milik Sendiri	Milik Sendiri
Jenis Lantai	Keramik	Semen
Jenis Dinding	Tembok	Kayu/Papan
Jenis Atap	Seng	Seng
Luas Bangunan	>36 - 60 m ²	>60 - 90 m ²
Sumber Air	Sumur	Sumur
Pemakaian Listrik	>100-150	>50-100

1. Menghitung Probabilitas Prior

Tahapan pertama dalam melakukan klasifikasi dengan Naïve Bayes adalah dengan menghitung probabilitas awal dari setiap kelas dengan rumus: $P(C_i) = n_i / N$. dengan keterangan $P(C_i)$ adalah probabilitas prior kelas ke-i, n_i merupakan jumlah data atau sampel yang masuk, dan N adalah total dari seluruh data training.

- $P(\text{Layak}) = 6/10 = 0,60$
- $P(\text{Tidak Layak}) = 4/10 = 0,40$

2. Menghitung Probabilitas Fitur (Likelihood)

Setelah mengetahui nilai probabilitas prior, langkah selanjutnya dalam algoritma naïve bayes adalah menghitung nilai likelihood atau probabilitas setiap kriteria pada masing-masing kelas.

Tabel 3.
Perhitungan Probabilitas Fitur

Kriteria	Kategori	P(XilLayak)	P(XilTidak)
Pekerjaan	Belum/Tidak Bekerja	1/6=0,166	0/4=0
	Buruh Harian Lepas	1/6=0,166	1/4=0,25
	Petani	2/6=0,333	1/4=0,25
	Pedagang	1/6=0,166	1/4=0,25
	Wiraswasta	1/6=0,166	0/4=0
	Karyawan Swasta	0/6=0	1/4=0,25
Penghasilan	≤1.000.000	2/6=0,333	0/4=0
	>1.000.000 - 1.800.000	2/6=0,333	1/4=0,25
	>1.800.000 - 2.500.000	0/6=0	1/4=0,25
	>2.500.000 - 3.200.000	0/6=0	1/4=0,25
	>3.200.000	0/6=0	1/4=0,25
Jumlah Tanggungan	1	0/6=0	2/4=0,50
	2	3/6=0,50	2/4=0,50
	3	3/6=0,50	0/4=0

Kepemilikan Rumah	Sewa	1/6=0,166	1/4=0,25
	Milik Sendiri	4/6=0,666	4/4=1
Jenis Lantai	Semen	2/6=0,333	0/4=0
	Keramik	5/6=0,833	3/4=0,75
Jenis Dinding	Kayu/Papan	2/6=0,333	1/4=0,25
	Tembok	4/6=0,666	3/4=0,75
Jenis Atap	Seng	4/6=0,666	4/4=1
	Asbes	0/6=0	0/4=0
Luas Bangunan	>36 - 60 m2	4/6=0,666	1/4=0,25
	>60 - 90 m2	2/6=0,333	3/4=0,75
Sumber Air	Sumur	4/6=0,666	3/4=0,75
	PDAM	2/6=0,333	1/4=0,25
Pemakaian Listrik	≤50	1/6=0,166	0/4=0
	>50-100	3/6=0,50	1/4=0,25
	>100-150	2/6=0,333	2/4=0,50
	>150	0/6=0	1/4=0,25

3. Menghitung Probabilitas Fitur (Likelihood)

Berikut adalah proses perhitungan skor posterior untuk data uji.

Data dengan nilai “Mualip”, “Karyawan Swasta”, “>3.200.000”, “2”, “Milik Sendiri”, “Keramik”, “Tembok”, “Seng”, “>36-60m2”, “Sumur”, “>100-150”

- $P(\text{Layak}) = 0,60 \times [0 \times 0 \times 0,50 \times 0,666 \times 0,833 \times 0,666 \times 0,666 \times 0,666 \times 0,666 \times 0,166] = 0$
- $P(\text{Tidak Layak}) = 0,40 \times [0,25 \times 0,50 \times 1 \times 0,75 \times 0,75 \times 0,75 \times 1 \times 0,25 \times 0,75 \times 0,50] = 0,00026367$

Berdasarkan perhitungan di atas, karena nilai $P(\text{Tidak Layak}) > P(\text{Layak})$, maka data uji atas nama Mualip diklasifikasikan sebagai Tidak Layak.

Data dengan nilai “Rahmat Faisal”, “Buruh Harian Lepas”, “>1.000.000 - 1.800.000”, “3”, “Milik Sendiri”, “Semen”, “Kayu/Papan”, “Seng”, “>60 - 90 m2”, “Sumur”, “>50-100”

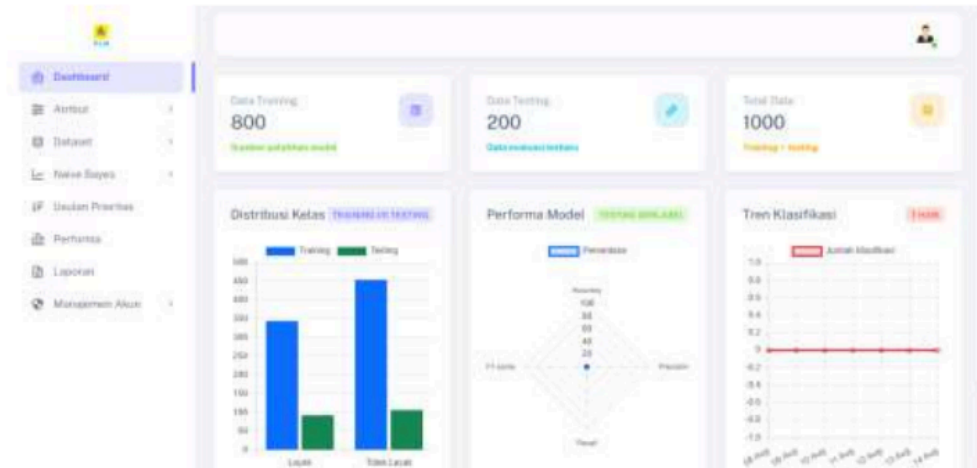
- $P(\text{Layak}) = 0,60 \times [0,166 \times 0,333 \times 0,50 \times 0,666 \times 0,333 \times 0,333 \times 0,666 \times 0,333 \times 0,666 \times 0,50] = 0,00012598$
- $P(\text{Tidak Layak}) = 0,40 \times [0,25 \times 0,25 \times 0 \times 1 \times 0 \times 0,25 \times 1 \times 0,75 \times 0,75 \times 0,25] = 0$

Implementasi Sistem

Implementasi sistem merupakan tahap merealisasikan rancangan aplikasi ke dalam bentuk sistem yang dapat digunakan. Sistem yang dibangun memiliki beberapa menu yang tersedia, yaitu menu login, menu dashboard, menu atribut dan nilai atribut, menu dataset, menu probabilitas, menu hasil klasifikasi, menu performa, menu usulan prioritas, menu manajemen akun, dan menu laporan.

1. Halaman Dashboard

Halaman ini menampilkan antarmuka utama sistem yang menyajikan ringkasan visualisasi data mengenai klasifikasi penerimaan subsidi listrik. Adapun menu dashboard dapat dilihat pada gambar 5.



Gambar 5. Tampilan Menu Dashboard

2. Halaman Atribut

Halaman atribut menunjukkan manajemen atribut (fitur/kriteria) yang digunakan dalam proses klasifikasi. Terdapat total 10 atribut (variabel input), seperti Pekerjaan, Penghasilan, Jumlah Tanggungan, Kepemilikan Rumah, Jenis Lantai, Jenis Dinding, Jenis Atap, Luas Bangunan, Sumber Air, dan Pemakaian Listrik. Halaman ini mendefinisikan tipe data dari masing-masing atribut, baik berupa data Kategorikal (teks/pilihan) maupun Numerik (angka).

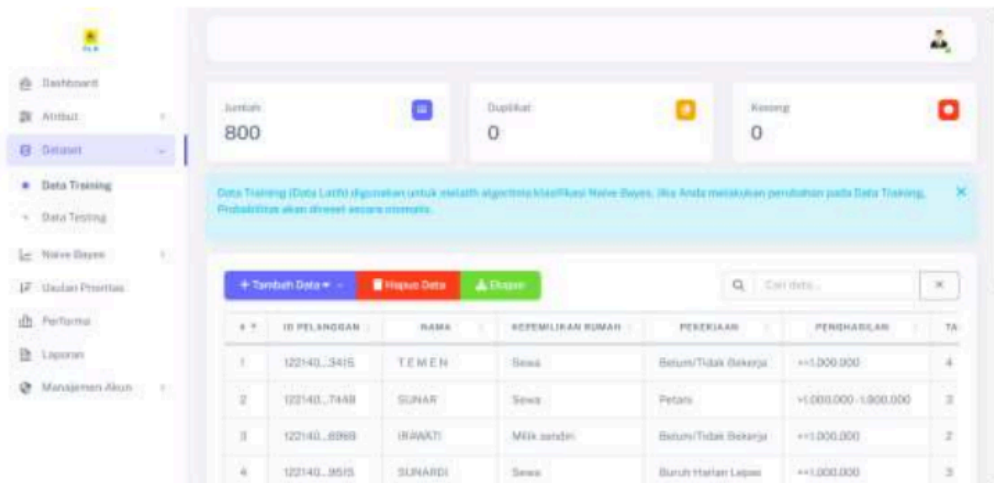
The Atribut page displays a table of attributes with the following columns: No, Nama, Tipe Atribut, Keterangan, and Aksi.

No	NAMA	Tipe Atribut	KETERANGAN	Aksi
1	Kepemilikan Rumah	Kategorikal	Status kepemilikan rumah (sudah/belum memiliki).	[Edit] [Hapus]
2	Pekerjaan	Kategorikal	Jenis pekerjaan (gaji tetap/gaji harian).	[Edit] [Hapus]
3	Penghasilan	Kategorikal	Kelompok penghasilan (rendah/tinggi).	[Edit] [Hapus]
4	Tanggungan	Numerik	Jumlah tanggungan (jumlah tanggungan).	[Edit] [Hapus]
5	Jenis Lantai	Kategorikal	Jenis lantai (lantai beton/lantai kayu).	[Edit] [Hapus]
6	Jenis Dinding	Kategorikal	Jenis dinding (dinding bata/dinding plester).	[Edit] [Hapus]
7	Jenis Atap	Kategorikal	Jenis atap (atap beton/atap genteng).	[Edit] [Hapus]

Gambar 6. Halaman Atribut

3. Halaman Data Latih

Halaman data latih berisikan sekumpulan data historis pelanggan subsidi listrik. Halaman ini berfungsi sebagai tempat untuk melihat, mengelola, dan memproses data yang digunakan agar sistem dapat mempelajari pola tertentu untuk kebutuhan klasifikasi.

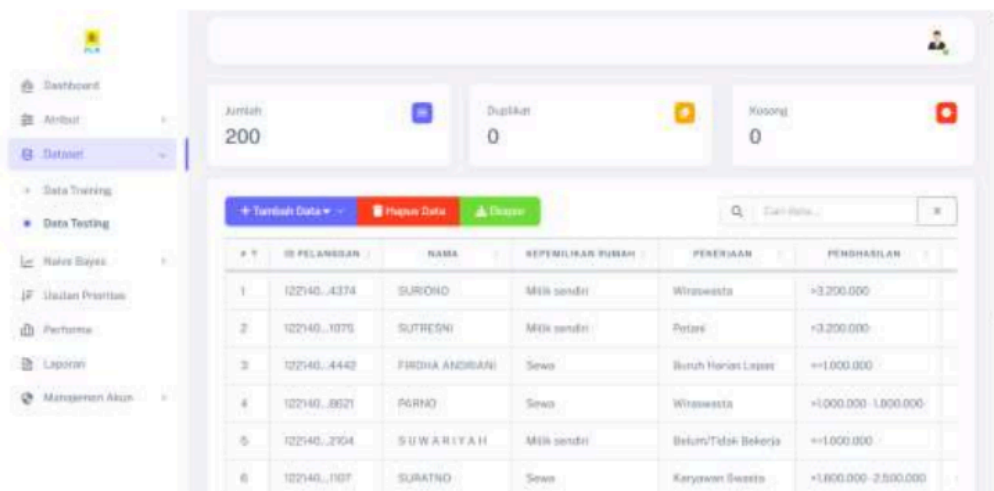


#	ID PELANGGAN	NAMA	KEPEMILIKAN RUMAH	PEKERJAAN	PENDHASILAN	TA
1	122140...3415	T E M E N	Sewa	Belum/Tidak Bekerja	<=1.000.000	4
2	122140...7448	SUNAR	Sewa	Petani	>1.000.000 -1.900.000	3
3	122140...8969	IRAKATI	Milik sendiri	Belum/Tidak Bekerja	<=1.000.000	2
4	122140...9515	SUNARDI	Sewa	Buruh Harian Lepas	<=1.000.000	3

Gambar 7. Halaman Data Latih

4. Halaman Data Uji

Merupakan halaman yang berisi data baru yang belum pernah dipelajari oleh sistem selama proses pelatihan. Halaman ini berfungsi untuk menguji data baru menggunakan model Naive Bayes yang sudah terbentuk untuk melihat seberapa akurat sistem dalam memprediksi kelas kelayakan yang sebenarnya.

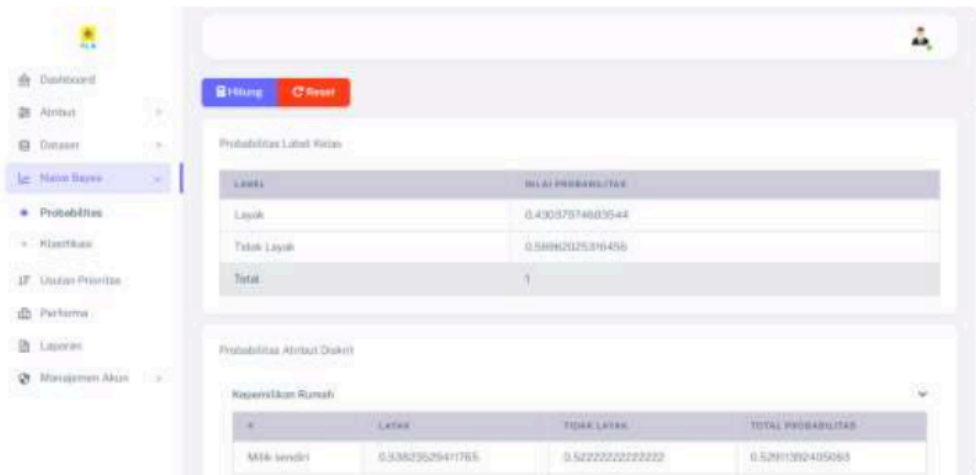


#	ID PELANGGAN	NAMA	KEPEMILIKAN RUMAH	PEKERJAAN	PENDHASILAN
1	122140...4374	SURONO	Milik sendiri	Wiraswasta	>3.200.000
2	122140...1075	SUTRESNI	Milik sendiri	Petani	<=3.200.000
3	122140...4442	FIRDIHA ANDRIANI	Sewa	Buruh Harian Lepas	<=1.000.000
4	122140...8621	GORNO	Sewa	Wiraswasta	>1.000.000 -1.900.000
5	122140...2754	S U W A R I Y A H	Milik sendiri	Belum/Tidak Bekerja	<=1.000.000
6	122140...1107	SURATNO	Sewa	Karyawan Swasta	>1.800.000 -2.500.000

Gambar 8. Halaman Data Uji

5. Halaman Probabilitas

Pada halaman ini berisi hasil perhitungan peluang nilai dari setiap kriteria terhadap kelas yang ada. Halaman Probabilitas bisa dilihat pada gambar di bawah ini.



Probabilitas Label Kelas

Label	Nilai Probabilitas
Layak	0.43027574603544
Tidak Layak	0.56972425396456
Total	1

Probabilitas Atribut Diskrit

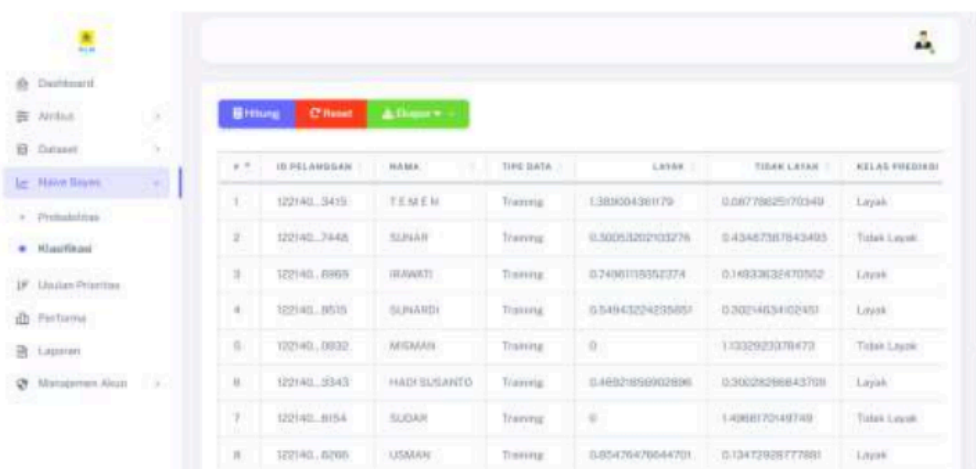
Kepemilikan Rumah:

	LAYAK	TIDAK LAYAK	TOTAL PROBABILITAS
Milik sendiri	0.53823529411765	0.52222222222222	0.5291392405068

Gambar 9. Halaman Probabilitas

6. Halaman Hasil Klasifikasi

Halaman ini menampilkan proses pengambilan keputusan oleh sistem. Pada halaman ini, algoritma Naive Bayes membandingkan nilai probabilitas akhir dari dua kelas, yaitu 'Layak' dan 'Tidak Layak'. Keputusan Kelas Prediksi diambil berdasarkan nilai bobot probabilitas yang paling tinggi.

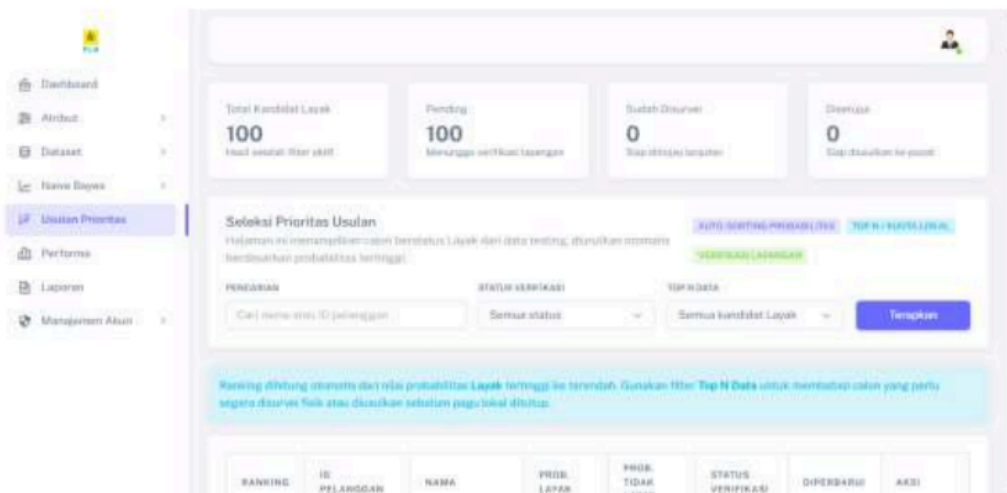


No	ID PELAMBSAN	NAMA	TIPE DATA	LAYAK	TIDAK LAYAK	KELAS PREDIKSI
1	122140_3415	F E S E N	Training	1.383004361179	0.08779625170349	Layak
2	122140_7445	SUNAR	Training	0.50053202103276	0.43457367843493	Tidak Layak
3	122140_8969	IRAWATI	Training	0.7490115351274	0.14933632470502	Layak
4	122140_8515	SUNARDI	Training	0.54843254235851	0.30214034102481	Layak
5	122140_0932	MIDHAN	Training	0	1.332923378473	Tidak Layak
6	122140_5343	HADI SUSANTO	Training	0.4692185902896	0.30028266843708	Layak
7	122140_8154	SUDAR	Training	0	1.408817049749	Tidak Layak
8	122140_6266	USMAN	Training	0.85476476644701	0.12472926777681	Layak

Gambar 10. Halaman Hasil Klasifikasi

7. Halaman Usulan Prioritas

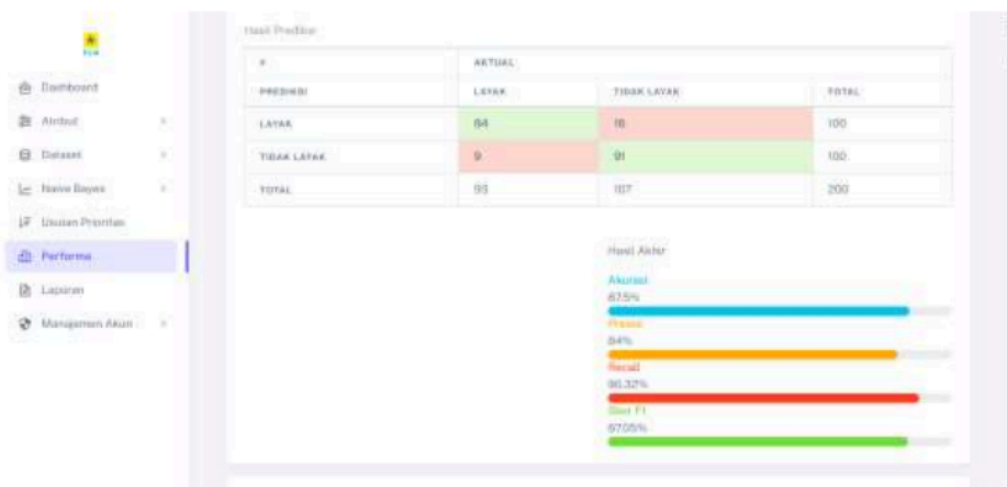
Halaman ini digunakan untuk menampilkan daftar calon penerima subsidi listrik yang telah melalui proses klasifikasi dan memperoleh status layak. Pada halaman ini, data diurutkan secara otomatis berdasarkan nilai probabilitas tertinggi, sehingga petugas dapat dengan lebih mudah menentukan calon penerima yang paling diprioritaskan untuk ditinjaulanjuti.



Gambar 11. Halaman Usulan Prioritas

8. Halaman Performa

Halaman performa menunjukkan evaluasi berupa ringkasan keberhasilan suatu model dalam mengklasifikasikan data secara tepat. Adapun halaman performa dapat dilihat pada Gambar 12.



Gambar 12. Halaman Performa

KESIMPULAN DAN REKOMENDASI

Penelitian ini menghasilkan sistem klasifikasi penerima subsidi listrik pelanggan 900 VA di Kecamatan Stabat dengan menerapkan teknik Data Mining menggunakan algoritma Naïve Bayes. Data yang digunakan sebanyak 1000 data yang terbagi menjadi 800 data latih dan 200 data uji, dengan 10 atribut sosial ekonomi yang digunakan untuk mengklasifikasikan penerima subsidi listrik ke dalam kategori Layak dan Tidak Layak. Hasil pengujian menunjukkan bahwa algoritma Naïve Bayes memperoleh nilai akurasi sebesar 87,5%, sehingga model memiliki kemampuan yang baik dalam melakukan klasifikasi. Selain itu, hasil klasifikasi Layak disajikan dalam bentuk usulan prioritas berdasarkan nilai probabilitas Layak dari yang tertinggi hingga terendah.

REFERENSI

1. Abdullah, R. W., Hartanti, D., Permatasari, H., Septyanto, A. W., & Abi Bagaskara, Y. (2022). Penerapan Data Mining untuk Memprediksi Jumlah Produk Terlaris Menggunakan Algoritma Naive Bayes Studi Kasus (Toko Prapti). *Jurnal Ilmiah Informatika Global*, 13(1).
2. Almira, A., & Ikhwan, A. (2021). Implementasi Data Mining Menggunakan Algoritma Fp-Growth pada Analisis Pola Pencurian Daya Listrik. 6(2), 442–448.
3. Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 219. <https://doi.org/10.30865/jurikom.v8i6.3655>
4. Fuad, M. N., & Asia, S. N. (2024). Penerapan Metode Naïve Bayes Untuk Penentuan Kelayakan Pembuatan Sertifikat Tanah Berbasis Web. *Jurnal Teknik AMATA*, 5(1).
5. KOMHARUDIN, K., & Aura, S. U. (2025). Perancangan Sistem Informasi Perpustakaan Berbasis Diagram UML (Unified Modeling Language). *SEMNASTERA (Seminar Nasional Teknologi Dan Riset Terapan)*, 7, 93–98.
6. Nurwijayanti, K., & Syahrullah Syahrullah. (2023). Klasifikasi Diagnosa Penyakit Diabetes Dengan Metode Naïve Bayes Berbasis Web. *Jurnal Kecerdasan Buatan Dan Teknologi Informasi*, 2(3), 115–121.
7. Pahlevi, R., Negara, E. S., Sutabri, T., & Herdiansyah, M. I. (2023). Penerapan Metode Naive Bayes untuk Menentukan Klasifikasi Kelayakan Penerimaan Bantuan Rehabilitasi dan Pembangunan Sekolah pada Dinas Pendidikan dan Kebudayaan Kabupaten

Banyuasin Magister Teknik Informatika , Universitas Bina Darma Provinsi Sumatera Sel. 9(2), 1176–1188.

8. Pratama, F. D., Zufria, I., & Triase, T. (2022). Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 7(1), 77–84.
<https://doi.org/10.36341/rabit.v7i1.2217>
9. Putra, M. S., Solikin, I., Duit, V. S., & Choiriyah, M. (2025). Implementation Of Forward Chaining In Expert System For Computer Troubleshooting. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 11(2), 257–264.
<https://doi.org/10.33330/jurtekxi.v11i2.3517>
10. Sari, A., Kania, I., & Oktasari, N. (2024). Optimasi Seleksi Penerima Bantuan PIP di SD Negeri 017107 Kisaran Naga dengan Metode Naïve Bayes. 1(2), 51–58.
11. Subhaktiyasa, P. G. (2024). Menentukan populasi dan sampel: Pendekatan metodologi penelitian kuantitatif dan kualitatif. *Jurnal Ilmiah Profesi Pendidikan*, 9(4), 2721–2731.
12. Suganda, G., Asfi, M., Subagio, R. T., & Kusuma, R. P. (2022). BEASISWA KARTU INDONESIA PINTAR (KIP) KULIAH MENGGUNAKAN NAÏVE. 9(2), 193–199.
<https://doi.org/10.30656/jsii.v9i2.4376>
13. Susana, H. (2022). Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet. *Jurnal Riset Sistem Informasi Dan Teknologi Informasi (JURSISTEKNI)*, 4(1), 1–8.
14. Utiahman, S. A., Dalai, H., & Sabudi, A. S. (2025). *Jurnal Teknologi dan Manajemen Informatika Optimasi K-Nearest Neighbor Dengan Particle Swarm Optimization Pada Klasifikasi Pelanggan Listrik Rumah Tangga Bersubsidi*. 11(1), 1–13.
15. Yusmandari Putera, D., Lestari Rahayu, S., & Artikel, I. (2023). Penerapan Data Mining Dalam Memprediksi Tingkat Perceraian Pasangan Muda Di Kota Medan Menggunakan Metode NAIVE BAYES. *Jurnal Rekayasa Sistem*, 1(1), 18.

Implementation of Data Mining for Classifying Electricity Subsidy Recipients in Stabat District Using the Naive Bayes Algorithm

ORIGINALITY REPORT

21%
SIMILARITY INDEX

17%
INTERNET SOURCES

12%
PUBLICATIONS

14%
STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Perguruan Tinggi Pelita Bangsa Student Paper	8%
2	journal.thamrin.ac.id Internet Source	5%
3	Submitted to Universitas Brawijaya Student Paper	2%
4	jurnal.unismuhpalu.ac.id Internet Source	2%
5	Ahmad Rizki Saputra, Rina - Wati, Tri - Susilowati. "KLASIFIKASI STATUS KELAYAKAN PENERIMA BLT-DD DI PEKON SUKA BANJAR II UJUNG REMBUN MENGGUNAKAN NAÏVE BAYES", Technologia : Jurnal Ilmiah, 2026 Publication	1%
6	Submitted to Universitas Malikussaleh Student Paper	1%
7	ejurnal.provisi.ac.id Internet Source	1%
8	Husein Nur Ansyari, Sriani Sriani, Armansyah Armansyah. "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Diabetes di Puskesmas Tegal Sari", Indonesian Journal of Innovation Multidisipliner Research, 2026 Publication	1%
9	Submitted to Universitas Sanata Dharma Student Paper	1%
10	nanopdf.com Internet Source	1%
11	samgriffin.info Internet Source	1%

