Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

Klasifikasi Kelayakan Penerima Program Indonesia Pintar (PIP) Menggunakan Teknik Data Mining Naive Bayes

Ahmad Syah Lubis*, Lili Tanti, Ratih Puspasari

Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Potensi Utama, Medan, Indonesia
Email: 1,*ahmadsyahlubis62@gmail.com,2lili@potensi-utama.ac.id, 3ratih@potensi-utama.ac.id
Email Penulis Korespondensi: ahmadsyahlubis62@gmail.com*
Submitted: **20/10/2025**; Accepted: **27/11/2025**; Published: **31/12/2025**

Abstrak–Program Indonesia Pintar (PIP) merupakan inisiatif bantuan pendidikan dari pemerintah yang ditujukan bagi siswa dari keluarga kurang mampu. Namun, dalam pelaksanaannya, proses seleksi penerima PIP sering kali dilakukan secara manual dan subjektif, sehingga rentan terhadap kesalahan dan ketidaktepatan sasaran. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi kelayakan penerima PIP dengan menggunakan teknik data mining melalui algoritma Naive Bayes. Pengujian dilakukan dengan data historis siswa di SMA Laksamana Martadinata. Hasil evaluasi menunjukkan bahwa metode Naive Bayes menghasilkan performa yang memuaskan, dengan akurasi sebesar 95% pada data pengujian dan 90% pada data baru. Sistem ini diharapkan dapat mendukung pihak sekolah dalam proses seleksi penerima PIP secara lebih objektif, efisien, dan akurat.

Kata Kunci: Program Indonesia Pintar; klasifikasi kelayakan; Naive Bayes; Sistem pendukung keputusan; Data mining

Abstract–The Indonesia Smart Program (PIP) is a government-funded educational assistance initiative aimed at underprivileged students. However, the selection process for recipients is often conducted manually and subjectively, which increases the risk of inaccuracy and misallocation. This research aims to develop a classification system for determining PIP eligibility using data mining techniques with the Naive Bayes algorithm. Testing was conducted using historical student data from SMA Laksamana Martadinata. The evaluation results show that the Naive Bayes method performs well, achieving 95% accuracy on testing data and 90% accuracy on new data. This system is expected to assist schools in conducting the PIP recipient selection process more objectively, efficiently, and accurately.

Keywords: Program Indonesia Pintar; eligibility classification; Naive Bayes; decision support system; data mining

1. PENDAHULUAN

Data mining melibatkan analisis sejumlah besar kumpulan data observasi untuk menemukan hubungan yang tidak terduga serta merangkum data dengan cara baru yang bermanfaat dan mudah dipahami oleh pengguna. Penggunaan sistem informasi terdistribusi mendorong pembentukan koleksi data besar di berbagai bidang [1]. Pendidikan merupakan elemen krusial dalam suatu negara karena memiliki peran signifikan dalam transformasi sosial masyarakat, baik di negara maju maupun berkembang. Pendidikan yang merata dan memadai di semua lapisan masyarakat akan berdampak pada kemajuan negara. Di Indonesia, pendidikan menjadi prioritas utama dalam pembangunan nasional, sebagaimana tercantum dalam Amandemen UUD 1945 yang menyatakan bahwa negara berkewajiban mencerdaskan kehidupan bangsa dan meningkatkan kesejahteraan umum. Menanggapi hal tersebut, pemerintah meluncurkan program Indonesia Pintar. PIP merupakan salah satu bentuk bantuan biaya pendidikan yang mencakup biaya akademik dan biaya hidup harian sesuai dengan jenjang pendidikan [2].

Metode *Naive Bayes* dapat digunakan sebagai teknik untuk sistem penentuan penerima bantuan PIP melalui perhitungan probabilitas antar kriteria yang ditentukan. [3]. Berdasarkan penelitian oleh [4] Kasus yang dianalisis adalah klasifikasi penerima Program Indonesia Pintar, dengan variabel independen: pekerjaan, penghasilan, serta variabel dependen: PIP (layak/tidak), jumlah dataset yang digunakan adalah 129, label kelas Program Indonesia Pintar (layak/tidak layak), algoritma yang diterapkan adalah algoritma *Naive Bayes* , teknik evaluasi AUC sebesar 89%.

Sedangkan penelitian yang dilakukan penulis dengan judul "Klasifikasi Kelayakan Penerima Program Indonesia Pintar (PIP) Menggunakan Teknik Data Mining". Kasus yang dianalisis adalah klasifikasi kelayakan penerima Program Indonesia Pintar (PIP), dengan variabel independen: alat transportasi, pendidikan ayah, pekerjaan ayah, penghasilan ayah, pendidikan ibu, pekerjaan ibu, penghasilan ibu, serta variabel dependen: PIP (layak/tidak), jumlah dataset yang digunakan adalah 1151, label kelas Program Indonesia Pintar (layak/tidak layak), algoritma yang diterapkan adalah algoritma *Naive Bayes* , teknik evaluasi *Confusion matrix* dengan hasil akurasi 95%, presisi 98%, recall 92%, dan F1-Score 95%. Penelitian ini juga menyajikan hasil evaluasi untuk data baru yang terdiri dari 30 sampel dengan akurasi 90%, presisi 89%, recall 94%, dan F1-Score 91%.

Program Indonesia Pintar (PIP) merupakan salah satu program bantuan pendidikan dari pemerintah Indonesia yang bertujuan untuk memberikan dukungan biaya kepada peserta didik yang berasal dari keluarga kurang mampu. Namun, dalam praktiknya, proses penentuan kelayakan penerima PIP di banyak sekolah, termasuk di SMA Laksamana Martadinata, masih menghadapi berbagai kendala. Proses seleksi sering dilakukan secara manual dan subjektif, sehingga menimbulkan potensi ketidaktepatan dalam penentuan siswa yang benar-benar layak menerima bantuan. Selain itu, jumlah siswa yang mengajukan bantuan semakin meningkat setiap tahun, sehingga sekolah membutuhkan metode yang lebih sistematis, objektif, dan efisien untuk melakukan klasifikasi kelayakan PIP. Pada era digital, pemanfaatan teknik data mining telah banyak digunakan dalam proses

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

pengambilan keputusan berbasis data. Salah satu algoritma yang banyak digunakan adalah Naïve Bayes, karena algoritma ini memiliki keunggulan dalam hal kecepatan, efisiensi, serta kemampuan menangani data berdimensi banyak dan bersifat kategorikal. Selain itu, Naïve Bayes memiliki performa yang stabil meskipun jumlah data besar dan distribusi antar kelas tidak seimbang. Oleh karena itu, metode ini dinilai cocok untuk proses klasifikasi kelayakan PIP yang melibatkan berbagai atribut seperti kondisi ekonomi, pekerjaan orang tua, prestasi akademik, dan status sosial siswa.

Berdasarkan penelitian oleh [5], kasus yang dianalisis adalah klasifikasi penentuan penerima Program Indonesia Pintar, dengan variabel independen: pekerjaan orang tua, penghasilan orang tua, pemegang KKS (Kartu Keluarga Sejahtera), pemegang SKTM (Surat Keterangan Tidak Mampu), serta variabel dependen: PIP (ya/tidak), jumlah dataset yang digunakan adalah 50, label kelas penerima Program Indonesia Pintar (ya/tidak), algoritma yang diterapkan adalah algoritma C4.5, teknik evaluasi presisi sebesar 100%.

Berdasarkan penelitian oleh [6], kasus yang dianalisis adalah klasifikasi kelayakan penerima bantuan pangan pokok, dengan variabel independen: PKH, jumlah tanggungan, kepala keluarga, jumlah penghasilan, serta variabel dependen: klasifikasi kelayakan penerima bantuan pangan pokok (layak/tidak layak), jumlah dataset yang digunakan adalah 135, label kelas klasifikasi kelayakan penerima bantuan pangan pokok (layak/tidak layak), algoritma yang diterapkan adalah algoritma *Naive Bayes* , teknik evaluasi presisi sebesar 88%.

Berdasarkan penelitian oleh [7], kasus yang dianalisis adalah klasifikasi penerima Program Indonesia Pintar, dengan variabel independen: jumlah pendapatan, jumlah tanggungan, pekerjaan, serta variabel dependen: PIP (ya/tidak), jumlah dataset yang digunakan adalah 400, label kelas Program Indonesia Pintar (layak/tidak layak), algoritma yang diterapkan adalah algoritma C4.5 dan *Naive Bayes* , teknik evaluasi AUC sebesar 100%.

Berdasarkan penelitian oleh [8], kasus yang dianalisis adalah klasifikasi penerima Program Indonesia Pintar, dengan variabel independen: jenis tempat tinggal, alat transportasi, penerima KPS, pekerjaan ayah, penghasilan ayah, pekerjaan ibu, penghasilan ibu, status penerima KIP, serta variabel dependen: PIP (layak/tidak), jumlah dataset yang digunakan adalah 130, label kelas Program Indonesia Pintar (layak/tidak layak), algoritma yang diterapkan adalah algoritma C4.5, teknik evaluasi presisi sebesar 85%.

Berdasarkan penelitian oleh [9], kasus yang dianalisis adalah klasifikasi penerima Program Indonesia Pintar, dengan variabel independen: pekerjaan, penghasilan, serta variabel dependen: PIP (layak/tidak), jumlah dataset yang digunakan adalah 129, label kelas Program Indonesia Pintar (layak/tidak layak), algoritma yang diterapkan adalah algoritma *Naive Bayes* , teknik evaluasi AUC sebesar 89%. Berdasarkan penelitian terdahulu, terlihat bahwa masih terdapat beberapa kekurangan yang belum terjawab dan menjadi dasar perlunya penelitian ini dilakukan. Penelitian oleh [10] menggunakan metode C4.5 tetapi hanya melibatkan dua kriteria sehingga hasil klasifikasi kurang komprehensif. Penelitian oleh [11] menerapkan K-NN, namun dataset yang digunakan relatif kecil sehingga akurasinya belum optimal. Penelitian oleh [12] menggunakan algoritma Naïve Bayes namun konteksnya pada ketepatan pemberian beasiswa reguler, bukan PIP, sehingga variabelnya tidak relevan dengan kondisi SMA Laksamana Martadinata. Penelitian [13] meneliti PIP tetapi belum menggunakan algoritma pembanding sehingga tidak diketahui keunggulan model. Sementara penelitian oleh [14] menggunakan metode SVM, namun memerlukan parameter tuning yang cukup kompleks sehingga kurang efisien untuk diterapkan di sekolah dengan keterbatasan sumber daya teknis.

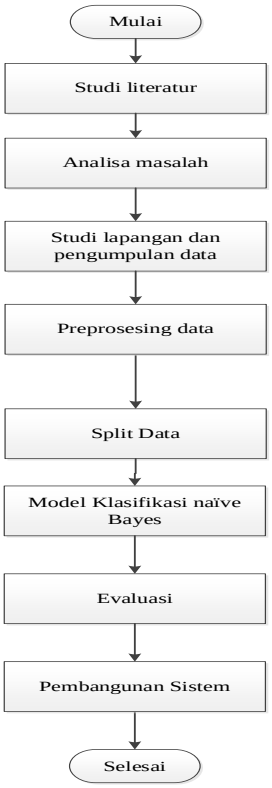
Celah penelitian (research gap) yang muncul adalah belum adanya penelitian yang secara spesifik menerapkan algoritma Naïve Bayes pada kasus klasifikasi kelayakan PIP di SMA Laksamana Martadinata dengan variabel yang lebih lengkap, dataset yang lebih representatif, dan pengujian akurasi yang lebih rinci. Penelitian ini hadir untuk mengisi gap tersebut dengan menggunakan model klasifikasi yang efisien, mudah diimplementasikan, dan sesuai dengan karakteristik data sekolah. Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan metode Naïve Bayes dalam melakukan klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) di SMA Laksamana Martadinata serta menganalisis tingkat akurasinya. Diharapkan penelitian ini dapat membantu sekolah memperoleh sistem pendukung keputusan yang lebih objektif, cepat, dan akurat sehingga proses penyaluran bantuan PIP menjadi lebih tepat sasaran dan transparan.

2. METODOLOGI PENELITIAN

2.1 Rancangan Penelitian

Metode prosedur perancangan adalah pengerjaan dari suatu sistem dilakukan secara berurutan atau secara linear. Jadi jika langkah satu belum dikerjakan maka tidak akan bisa melakukan pengerjaan langkah 2, 3 dan seterusnya. Secara otomatis tahapan ke-3 akan bisa dilakukan jika tahap ke-1 dan ke-2 sudah dilakukan.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160



Gambar 1. Tahapan Penelitian

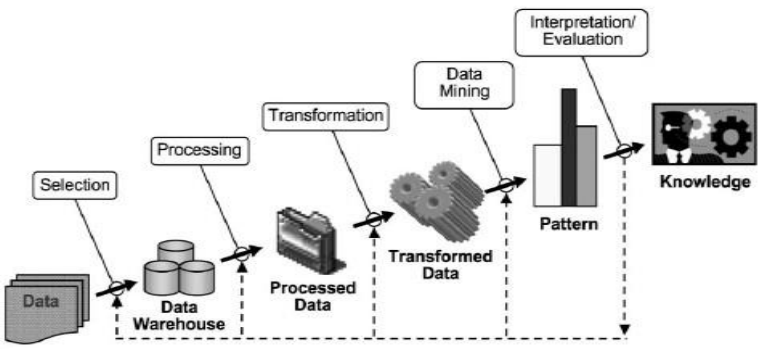
Keterangan gambar 1 :

a. Studi Literatur
Digunakan untuk mempelajari dan memperdalam pemahaman peneliti tentang penentuan penerima Program Indonesia Pintar (PIP) dan metode *Naive Bayes* secara spesifik, serta membaca jurnal atau referensi terkait penelitian.

b. Analisis Masalah
Analisis masalah digunakan untuk memecahkan masalah dan menganalisis data dalam kajian sebelum perancangan atau perhitungan.

c. Pengumpulan Data
Tahap ini dilakukan dengan mengambil data siswa penerima dan tidak penerima PIP pada SMA Laksamana Martadinata. Data meliputi variabel seperti pendapatan orang tua, jumlah tanggungan, status rumah, pekerjaan orang tua, dan nilai rapor. Setelah seluruh data dikumpulkan, langkah berikutnya adalah memastikan bahwa data berada dalam kondisi siap digunakan melalui proses preprocessing.

d. Preprocessing
Preprocessing bertujuan untuk membersihkan data, mengintegrasikan data, dan mentransformasi data agar data siswa yang digunakan dalam proses klasifikasi lebih tepat dan akurat. Tahap preprocessing ini melibatkan beberapa langkah, seperti mengisi data yang hilang atau kosong, mengurangi noise data yang mengandung kesalahan atau outlier, serta menangani data yang tidak konsisten. Setelah preprocessing selesai, data yang sudah bersih kemudian digunakan dalam proses pemodelan.



Gambar 2. Tahapan KDD

Tahapan data mining dibagi menjadi bagian-bagian yaitu :

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

a. *Selection* (Pemilihan)
Pada tahap ini, data yang relevan dipilih dari berbagai sumber (misalnya: file excel, database, sistem informasi, dsb). Tidak semua data akan digunakan, hanya data yang sesuai dengan tujuan analisis. Tujuan yaitu mengambil subset data yang paling sesuai dan relevan untuk dianalisis.
b. *Processing* (Pemrosesan)
Pada tahap ini, data mentah diproses dan dibersihkan. Ini mencakup penghapusan data yang tidak lengkap, duplikat, atau tidak konsisten. Tujuan adalah menyiapkan data agar dapat digunakan untuk analisis lebih lanjut.
c. *Transformation* (Transformasi)
Data yang telah diproses kemudian ditransformasikan ke dalam format atau struktur yang sesuai dengan metode data mining yang akan digunakan. Bisa berupa normalisasi, pengkodean ulang, atau pembuatan variabel baru. Tujuan Adalah menyesuaikan data agar lebih optimal untuk dianalisis.
d. Data *Mining*
Ini adalah inti dari proses KDD, yaitu penerapan teknik analisis seperti klasifikasi, clustering, asosiasi, dsb. untuk menemukan pola tersembunyi dalam data. Tujuan Adalah menggali informasi dan pola dari data yang sudah ditransformasikan.
e. *Interpretation*/*Evaluation* (Interpretasi/Evaluasi)
Tahap ini mengevaluasi apakah pola yang ditemukan bermanfaat, valid, dan dapat diandalkan untuk mendukung pengambilan keputusan. Tujuan adalah menilai apakah informasi yang diperoleh dari data mining memang benar-benar penting dan relevan [15]
f. Model Klasifikasi Metode *Naive Bayes*
Pada tahap ini dilakukan perhitungan peluang prior, likelihood, dan posterior berdasarkan rumus Teorema Bayes. Model dibangun dengan menggunakan data latih untuk menghasilkan pola klasifikasi penerima PIP. Model yang telah terbentuk kemudian diuji untuk mengetahui tingkat akurasi, presisi, recall, dan error rate melalui tahap evaluasi.
g. Evaluasi
Tahap ini dilakukan dengan menghitung performa model menggunakan *confusion matrix*. Evaluasi bertujuan untuk mengukur akurasi klasifikasi dan memastikan model layak digunakan untuk menentukan kelayakan PIP. Setelah model dievaluasi, hasilnya diinterpretasikan untuk menghasilkan keputusan kelayakan PIP pada penelitian ini.
h. Pembangunan Sistem
Tahap ini dilakukan untuk desain sistem menggunakan UML, yang mencakup diagram use case, class diagram, sequence diagram, dan activity diagram, kemudian desain database dan antarmuka.

2.2 Data Mining

Data mining merupakan istilah yang digunakan untuk menemukan pengetahuan tersembunyi dalam database. Data mining adalah proses semi-otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi serta mengidentifikasi informasi dan pengetahuan potensial yang bermanfaat dari database besar[16]. Tujuan data mining adalah mengidentifikasi informasi yang berguna dan pengetahuan untuk mendukung pengambilan keputusan yang lebih baik. Dalam konteks bisnis, data mining dapat membantu perusahaan memahami perilaku pelanggan, meningkatkan efisiensi operasi, meningkatkan kualitas produk, dan mengoptimalkan strategi pemasaran. Data mining sering disebut sebagai knowledge discovery in database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, penggunaan data historis untuk menemukan keteraturan, pola, atau hubungan dalam set data berukuran besar [17].

2.3 *Naive Bayes*

Naive Bayes Classifier merupakan metode klasifikasi yang berbasis pada teorema Bayes. Metode pengklasifikasian ini menggunakan probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya, sehingga dikenal sebagai Teorema Bayes [18]. Keunggulan *Naive Bayes* telah diamati bahwa klasifikasi *Naive Bayes* mudah diimplementasikan dan cepat. Ini akan konvergen lebih cepat daripada model diskriminatif seperti regresi logistik. Membutuhkan lebih sedikit data pelatihan.
Langkah Penyelesaian *Naive Bayes*
Metode *Naive Bayes* terdiri dari tiga langkah utama:
1. Perhitungan peluang prioritas
2. Perhitungan peluang kondisional
3. Dan pemilihan kategori [19]

Prediksi Bayes didasarkan pada teorema Bayes dengan formula umum sebagai berikut :

$$P(H|E) = P(E|H) \times P(H) / P(E) \quad \text{.....(1)}$$

Keterangan:
P(H|E): Probabilitas akhir bersyarat suatu hipotesis H terjadi jika bukti E terjadi.
P(E|H): Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H.
P(H): Probabilitas awal hipotesis H terjadi tanpa memandang bukti apapun.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

P(E): Probabilitas awal bukti E terjadi tanpa memandang hipotesis atau bukti yang lain .
Kelebihan metode ini adalah:
1. Memperbaiki data yang berbentuk angka dan data yang terpisah-pisah.
2. Algoritma ini efisien dalam hal komputasi, sehingga dapat mengolah data dalam jumlah besar dengan cepat.
3. *Naive Bayes* bekerja dengan baik untuk data kategorikal.
4. Algoritma *Naive Bayes* cukup handal dalam menangani data yang hilang.
5. Cepat serta menghemat waktu selama estimasi.
6. Kuat terhadap karakteristik yang tidak relevan.
Kekurangan Metode *Naive Bayes* :
1. *Naive Bayes* dapat menjadi sensitif terhadap fitur atau atribut yang tidak relevan.
2. Memperkirakan bahwa karakteristiknya bebas.
3. Ini tidak berlaku; jika probabilitas bersyarat adalah 0, probabilitas prediksi juga akan menjadi 0 [19].

2.4. Evaluasi Model *Confusion matrix*

Tahap ini bertujuan untuk mengevaluasi kemampuan atau performa algoritma dalam hal akurasi model, yang memerlukan perhitungan matriks kebingungan (*Confusion matrix*).. Matriks ini mencakup empat kombinasi berbeda berdasarkan nilai prediksi dan nilai aktual. Tabel evaluasi model dapat dirujuk pada tabel berikut [20]:

Tabel 1. *confusion matrix*

Confusion matrix		**Nilai aktual**	
		Positif	**Negatif**
Nilai Prediksi	**Positif**	*True Positives*	*False Positives*
	Negatif	*False Negatives*	*True Negatives*

Confusion matrix dimaksudkan untuk memvisualisasikan hasil prediksi dan kondisi aktual dari data yang dihasilkan oleh algoritma pembelajaran mesin. Hal ini dicapai melalui penghitungan metrik seperti akurasi, presisi, recall, dan F1-score. Rumus untuk keempat metrik tersebut disajikan dalam tabel berikut:

Tabel 2 Rumus *Confusion matrix*

No	Pengukuran	Rumus
1	Akurasi	$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$
2	Presisi	$\text{Precision} = \frac{TP}{TP+FP}$
3	Recall	$\text{Recall} = \frac{TP}{TP+FN}$
4	F1-Score	$\text{F1-Score} = \frac{2x Recall\ x\ Precision}{Recall+Precision}$

3. HASIL DAN PEMBAHASAN

3.1 Analisis Kebutuhan

Data yang digunakan dalam penelitian ini berasal dari SMA Laksamana Martadinata sebagai bagian dari proses pendataan dan verifikasi calon penerima Program Indonesia Pintar (PIP). Program ini ditujukan untuk membantu peserta didik yang berasal dari keluarga kurang mampu agar tetap dapat mengakses pendidikan tanpa hambatan ekonomi. Metode Pengumpulan data yang dilakukan pada penelitian ini berupa wawancara dengan kepala sekolah dari SMAS Laksamana Martadina, kemudian melakukan pengumpulan data berdasarkan data dari SMAS Laksamana Martadina. Data yang digunakan pada penelitian ini sebanyak 1151 dengan memiliki Variabel Independen seperti Alat Transportasi, Pendidikan Ayah, Pekerjaan Ayah, Penghasilan Ayah, Pendidikan Ibu, Pekerjaan Ibu, Penghasilan Ibu, alasan layak PIP (PKH, Yatim, Miskin/Rentan miskin). Data siswa/i pada SMAS Laksamana Martadina dapat dilihat pada Tabel 3

Tabel 3. Fitur Yang digunakan

	Nama Fitur	**Keterangan**
1	Nomor	Nomor urutan untuk data yang akan diseleksi dalam penerima PIP
2	Nama	Nama siswa untuk data yang akan diseleksi dalam penerima PIP
3	Jenis Kelamin	Jenis kelasmin untuk data yang akan diseleksi dalam penerima PIP
4	NISN	NISN untuk data yang akan diseleksi dalam penerima PIP
5	Tempat Lahir	Tempat lahir untuk data yang akan diseleksi dalam penerima PIP
6	Tanggal Lahir	Tanggal lahir untuk data yang akan diseleksi dalam penerima PIP
7	NIK	NIK untuk data yang akan diseleksi dalam penerima PIP
8	Agama	Agama siswa dalam data yang akan diseleksi dalam penerima PIP

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

	Nama Fitur	Keterangan
9	Alamat	Alamat siswa untuk data yang akan diseleksi dalam penerima PIP
10	Alat Transportasi	Alat transportasi yang di gunakan siswa
11	Nama ayah	Nama ayah dari siswa dalam penentuan kelayakan pemberian beasiswa PIP
12	Jenjang pendidikan ayah	Jenjang pendidikan ayah dari siswa dalam penentuan kelayakan pemberian beasiswa PIP seperti S2, S1, D1, D2, D3, D4, SMA/ Sederajat, SMP/ Sederajat, SD/ Sederajat, Putus SD, Tidak sekolah
13	Pekerjaan ayah	Pekerjaan ayah dari siswa dalam penentuan kelayakan pemberian beasiswa PIP seperti Buruh, Karyawan Swasta, lainnya, nelayan, pedagang besar, pedang kecil, petani, Peternak, Wiraswasta, PNS/ TNI/ POLRI, sudah meninggal, tidak bekerja, tidak dapat diterapkan, wirausaha
14	Penghasilan ayah	Penghasilan ayah dari siswa dalam penentuan kelayakan pemberian beasiswa PIP
15	Nama ibu	Nama ibu dari siswa dalam penentuan kelayakan pemberian beasiswa PIP
16	Jenjang pendidikan ibu	Jenjang ibu ayah dari siswa dalam penentuan kelayakan pemberian beasiswa PIP seperti S2, S1, D1, D2, D3, D4, SMA/ Sederajat, SMP/ Sederajat, SD/ Sederajat, Putus SD, Tidak sekolah
17	Pekerjaan ibu	Pekerjaan ibu dari siswa dalam penentuan kelayakan pemberian beasiswa PIP seperti Buruh, Karyawan Swasta, lainnya, nelayan, pedagang besar, pedang kecil, petani, Peternak, Wiraswasta, PNS/ TNI/ POLRI, sudah meninggal, tidak bekerja, tidak dapat diterapkan, wirausaha
18	Penghasilan ibu	Penghasilan ibu dari siswa dalam penentuan kelayakan pemberian beasiswa PIP
19	Rombel saat ini	Status rombel siswa dalam penentuan kelayakan penerima PIP
20	Status layak PIP	Status layak PIP yang terdiri dari layak dan tidak layak
21	Alasan Layak PIP	Alasan layak PIP seperti Tidak ada, pemegang PKH/KPS/KKS, Siswa miskin/ Rentan miskin, Sudah mampu, Yatim Piatu/Panti Asuhan/Panti Sosial

3.2.Hasil Penelitian

1. Preprocesing Data

Setelah dilakukan proses *Preprocessing* data yang diperoleh adalah 1066 siswa/i. Maka proses Pre-processing dapat dilihat pada Tabel 4 berikut :

a. *Missing Value*

Missing Value dalam preprocessing mengacu pada data yang tidak tersedia atau absen pada fitur atau variabel dalam dataset. Ini merupakan nilai yang tidak terisi, yang dapat terjadi akibat berbagai faktor seperti kesalahan pengumpulan data, responden yang tidak memberikan jawaban, atau data yang tidak dapat diukur. Teknik umum untuk menangani missing value meliputi:

a) Penghapusan Data (Deletion): Menghilangkan baris atau kolom yang memiliki missing value jika jumlahnya minimal.

b) Imputasi: Mengganti data yang salah menjadi lebih akurat dan formal.

Tabel 4. Statistik Missing Value

No	Nama Fitur	Missing Value
1	Nomor	0
2	Nama	0
3	Jenis Kelamin	0
4	NISN	0
5	Tempat Lahir	0
6	Tanggal Lahir	0

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

No	Nama Fitur	Missing Value
7	NIK	13
8	Agama	0
9	Alamat	0
10	Alat Transportasi	19
11	Nama ayah	16
12	Jenjang pendidikan ayah	17
13	Pekerjaan ayah	0
14	Penghasilan ayah	0
15	Nama ibu	0
16	Jenjang pendidikan ibu	20
17	Pekerjaan ibu	0
18	Penghasilan ibu	0
19	Rombel saat ini	0
20	Status layak PIP	0
21	Alasan Layak PIP	0

b. *Outlier*
Dari data set yang diperoleh ada 19 data yang outlier pada fitur nama ayah dan penghasilan ayah.
c. Seleksi Fitur
Ada beberapa fitur yang diseleksi diantaranya adalah :

Tabel 5. Seleksi Fitur

No	Nama Fitur	Keterangan
1	Nomor	Tidak digunakan
2	Nama	Tidak digunakan
3	Jenis Kelamin	Tidak digunakan
4	NISN	Tidak digunakan
5	Tempat Lahir	Tidak digunakan
6	Tanggal Lahir	Tidak digunakan
7	NIK	Tidak digunakan
8	Agama	Tidak digunakan
9	Alamat	Tidak digunakan
10	Alat Transportasi	Digunakan
11	Nama ayah	Tidak digunakan
12	Jenjang pendidikan ayah	Digunakan
13	Pekerjaan ayah	Digunakan
14	Penghasilan ayah	Digunakan

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

15	Nama ibu	Tidak digunakan
16	Jenjang pendidikan ibu	Digunakan
17	Pekerjaan ibu	Digunakan
18	Penghasilan ibu	Digunakan
19	Rombel saat ini	Tidak digunakan
20	Status layak PIP	Digunakan
21	Alasan Layak PIP	Digunakan

2. Split Data

Proses pembagian *data set* dilakukan yaitu 20% untuk testing dan 80% untuk data training. Maka hasil penentuan split data sebagai berikut :
Data Testing = 20% * 1066 = 213
Data Training = 80% * 1066 = 853
3. Model Klasifikasi *Naive Bayes*
Data set sebanyak 853 data, 448 sebagai penerima dan 405 tidak menerima.
P (Tidak) = 405/853 = 0.4747
P (Ya) = 448/853 = 0.5252
Menghitung probabilitas atribut

Tabel 6. Variabel dan Nilai Probabilitas Prior (Training)

Kode Atribut	Atribut/Variabel	Total	Data		P(X)Ci	
			Ya	Tidak	Ya	Tidak
		853	448	405	0.5252	0,4748
Alat Transportasi	Andong/ bendi/ sado/ dokar/ delman/ becak	3	1	2	0,0025	0,0045
	Angkutan umum/ bus/ pete-pete	218	97	121	0,2395	0,2701
	Jalan Kaki	466	268	198	0,6617	0,4420
	kendaraan Pribadi	34	13	21	0,0321	0,0469
	Lainnya	15	6	9	0,0148	0,0201
	Mobil Pribadi	1	1	0	0,0025	0,0000
	Mobil/Bus Antar Jemput	7	5	2	0,0123	0,0045
	Ojek	7	5	2	0,0123	0,0045
	Sepeda	18	12	6	0,0296	0,0134
	Sepeda Motor	84	40	44	0,0988	0,0982
Pendidikan Ayah	S2	0	0	0	0,0000	0,0000
	S1	21	7	14	0,0173	0,0313
	D1	4	1	3	0,0025	0,0067
	D2	1	1	0	0,0025	0,0000
	D3	10	5	5	0,0123	0,0112
	D4	1	0	1	0,0000	0,0022
	SMA/ Sederajat	587	294	293	0,7259	0,6540
	SMP/ Sederajat	177	108	69	0,2667	0,1540
	SD/ Sederajat	44	27	17	0,0667	0,0379
	Putus SD	5	3	2	0,0074	0,0045
	Tidak Sekolah	3	2	1	0,0049	0,0022
Pekerjaan Ayah	Buruh	60	40	20	0,0988	0,0446
	Karyawan Swasta	246	127	119	0,3136	0,2656

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

Kode Atribut	Atribut/Variabel	Total	Data		P(X)Ci	
			Ya	Tidak	Ya	Tidak
	Lainnya	28	19	9	0,0469	0,0201
	Nelayan	2	2	0	0,0049	0,0000
	Pedagang Besar	1	0	1	0,0000	0,0022
	Pedang Kecil	23	10	13	0,0247	0,0290
	Petani	15	8	7	0,0198	0,0156
	Peternak	1	1	0	0,0025	0,0000
	Wiraswasta	413	211	202	0,5210	0,4509
	PNS/ TNI/ POLRI	27	3	24	0,0074	0,0536
	Sudah Meninggal	19	14	5	0,0346	0,0112
	Tidak Bekerja	5	4	1	0,0099	0,0022
	Tidak Dapat Diterapkan	2	2	0	0,0049	0,0000
	Wirausaha	11	7	4	0,0173	0,0089
Penghasilan Ayah	Kurang dari Rp 500,000	17	16	1	0,0395	0,0022
	Rp 1.000.000 -Rp 1.999.999	471	267	204	0,6593	0,4554
	Rp 2.000.000 - Rp 4.999.999	186	50	136	0,1235	0,3036
	Rp 5.000.000 - Rp 20.000.000	1	1	0	0,0025	0,0000
	Rp 500.000 - Rp 999.999	154	96	58	0,2370	0,1295
	Tidak Berpenghasilan	24	18	6	0,0444	0,0134
Pendidikan Ibu	S2	1	0	1	0,0000	0,0022
	S1	24	7	17	0,0173	0,0379
	D1	5	2	3	0,0049	0,0067
	D2	0	0	0	0,0000	0,0000
	D3	8	3	5	0,0074	0,0112
	D4	1	1	0	0,0025	0,0000
	SMA/ Sederajat	559	276	283	0,6815	0,6317
	SMP/ Sederajat	194	112	82	0,2765	0,1830
	SD/ Sederajat	50	37	13	0,0914	0,0290
	Putus SD	3	3	0	0,0074	0,0000
	Tidak Sekolah	8	7	1	0,0173	0,0022
Pekerjaan Ibu	Buruh	14	8	6	0,0198	0,0134
	Karyawan Swasta	34	11	23	0,0272	0,0513
	Lainnya	43	19	24	0,0469	0,0536
	Nelayan	1	1	0	0,0025	0,0000
	Pedagang Besar	0	0	0	0,0000	0,0000
	Pedang Kecil	27	15	12	0,0370	0,0268
	Petani	13	10	3	0,0247	0,0067
	Peternak	0	0	0	0,0000	0,0000
	Wiraswasta	62	32	30	0,0790	0,0670
	PNS/ TNI/ POLRI	5	1	4	0,0025	0,0089
	Sudah Meninggal	1	1	0	0,0025	0,0000
	Tidak Bekerja	650	348	302	0,8593	0,6741

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

Kode Atribut	Atribut/Variabel	Total	Data		P(X)Ci	
			Ya	Tidak	Ya	Tidak
Penghasilan Ibu	Tidak Dapat Diterapkan	1	1	0	0,0025	0,0000
	Wirausaha	2	1	1	0,0025	0,0022
	Kurang dari Rp 500,000	28	15	13	0,0370	0,0290
	Rp 1.000.000 -Rp 1.999.999	75	38	37	0,0938	0,0826
	Rp 2.000.000 - Rp 4.999.999	23	3	20	0,0074	0,0446
	Rp 5.000.000 - Rp 20.000.000	0	0	0	0,0000	0,0000
	Rp 500.000 - Rp 999.999	70	40	30	0,0988	0,0670
	Tidak Berpenghasilan	657	352	305	0,8691	0,6808
Alasan Layak PIP	Tanpa Keterangan	492	91	401	0,2247	0,8951
	Pemegang PKH/KPS/KKS	41	40	1	0,0988	0,0022
	Siswa miskin/ Rentan miskin	312	310	2	0,7654	0,0045
	Sudah mampu	1	0	1	0,0000	0,0022
	Yatim Piatu/Panti Asuhan/Panti Sosial	7	7	0	0,0173	0,0000

Table di atas menunjukan nilai subset dari masing-masing atribut pada kelas Ya dan kelas Tidak. Nilai tersebut akan digunakan untuk perhitungan pada probabilitas kelas yang akan terpilih.

3.3. Pembahasan

Metode *Naive Bayes* dipilih karena memiliki keunggulan berupa kecepatan perhitungan, kemampuan menangani dataset berdimensi banyak, serta sering digunakan untuk permasalahan klasifikasi dengan distribusi atribut kategorikal. Prinsip kerja *Naive Bayes* adalah menghitung probabilitas setiap kelas berdasarkan frekuensi kemunculan setiap atribut dalam dataset latih. Menghitung nilai probabilitas pada setiap kelas data *testing*

Tabel 7. Hasil Evaluasi

No.	NIS	Nama	Data Aktual	Model Evaluasi
1	1300000001	Aisyah Fitri	Tidak	Tidak
2	1300000002	Akiko Muethia	Ya	Ya
3	1300000003	Anggi Alhbiyansa	Tidak	Tidak
4	1300000004	Annisa Rahayu	Tidak	Tidak
5	1300000005	Arrahma Kirana Putri	Tidak	Tidak
6	1300000006	Aulia Aisyah Safitri	Tidak	Tidak
7	1300000007	Aura Salwa Azhira	Ya	Ya
8	1300000008	Bunga Alisyah	Ya	Ya
9	1300000009	Chania Fill	Ya	Ya
10	1300000010	Chayla Divina	Tidak	Tidak
…	…	…	…	…
212	1300000212	Vikky Violla Zalogo	Ya	Ya
213	1300000213	Wafiq Azzahra Hamdila Putri	Ya	Ya

Hasil probabilitas yang didapatkan dari perhitungan P(X|Ci) * P(Ci) selanjutnya dilakukan perbandingan nilai. Perbandingan nilai probabilitas P(X|Tidak Menerima) dan P(X|Menerima) adalah Tidak Menerima > Menerima sehingga data tersebut diklasifikasikan kedalam class Tidak Menerima. Tabel evaluasi model dapat ditemukan pada tabel berikut

Tabel 8. *confusion matrix*

Confusion matrix	Nilai aktual	
	Positif	**Negatif**

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

Nilai Prediksi	**Positif**	87	2
	Negatif	8	116
Jumlah		95	118
	Total		213

Tabel 9. rumus *confusion matrix*

No	Pengukuran	Rumus	Rumus	Hasil
1	Akurasi	$\frac{TP + TN}{TP + FP + FN + TN}$	(87 + 116)/ (87 + 2 + 8 + 116)	95%
2	Presisi	$\frac{TP}{TP+FP}$	87 / (87 + 2)	98%
3	Recall	$\frac{TP}{TP + FN}$	87 / (87 + 8)	92%
4	F1-Score	$\frac{2xRecall\ x\ Precision}{Recall + Precision}$	2 * 98 * 92 / (92 + 98)	95%

Pembahasan penelitian ini menyoroti kemampuan Naive Bayes dalam mengklasifikasikan kelayakan penerima PIP dengan akurasi tinggi. Model mengenali pola penting, terutama pada atribut penghasilan orang tua, kepemilikan PKH/KKS/KIP, dan jumlah tanggungan, karena perbedaan probabilitasnya jelas antara siswa layak dan tidak layak. Naive Bayes menghitung kontribusi tiap atribut secara terpisah melalui asumsi independensi, dan meski tidak selalu sepenuhnya sesuai kondisi nyata, metode ini tetap efektif secara empiris. Selain itu, model membantu mengurangi bias subjektif penilaian manual di sekolah sehingga keputusan lebih berbasis data.

3.3.1 Menghitung nilai probabilitas pada setiap kelas data *Baru*

Visualisasi metrik ini menjadi alat penting dalam validasi sistem, dan bisa digunakan untuk mendukung justifikasi penggunaan algoritma *Naive Bayes* dalam studi kasus ini. Berikut ini merupakan hasil perhitungan confusion matriks untuk data baru :

Tabel 10. Data Masyarakat Baru

No.	**NIS & Nama**	**Layak PIP (Aktual)**	**Layak PIP (Evaluasi)**
1	88950005 : HAIRANI	Ya	Ya
2	88950006 : MAFZUL	Ya	Ya
3	88950007 : SURIADI	Tidak	Tidak
4	88950008 : ZAHRIL QAIRI	Ya	Ya
5	88950009 : NUR AKMALIA	Ya	Tidak
6	88950010 : MARTUNIS	Ya	Ya
7	88950011: NURSABARINI	Ya	Ya
8	88950012 : T ZIKRA	Tidak	Tidak
9	88950013 : SRI NOVIANI	Ya	Ya
10	88950014 : MULIADI AB	Tidak	Tidak
…	…	…	…
30	88950034 : HERI FAISANDRA	Ya	Ya

Tabel 11. Data Perhitungan Data

Confusion matrix		Nilai Aktual	
		Positif	Negatif
Nilai Prediksi	Positif	16	2
	Negatif	1	11
Jumlah		17	13

Tabel 12. Data Perhitungan Confusion Matriks

No	Pengukuran	Rumus	Hasil
1	Akurasi	16 + 11 / (16 + 2 + 1 + 11)	90 %
2	Presisi	16 / (16 + 2)	89 %
3	Recall	16 / (16 + 1)	94 %
4	F1-Score	2 * 94 * 89 / (94 + 89)	91 %

4. KESIMPULAN

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

Pengujian Naive Bayes pada data calon penerima PIP menunjukkan performa yang baik dengan akurasi 90%, precision 89%, recall 94%, dan F1-score 91%. Hasil ini menandakan model cukup andal dalam mengklasifikasikan kelayakan siswa. Precision yang tinggi menunjukkan model relatif jarang memberi label "layak" kepada siswa yang sebenarnya tidak layak, sehingga membantu mencegah bantuan salah sasaran. Recall 94% juga menunjukkan sebagian besar siswa yang layak sudah berhasil terdeteksi, meskipun masih ada kemungkinan sebagian kecil terlewat karena data tumpang tindih, ketidakseimbangan kelas, atau variabel yang kurang informatif. Berdasarkan probabilitas posterior, atribut yang paling berpengaruh adalah pekerjaan orang tua, penghasilan, jumlah tanggungan keluarga, dan kepemilikan KIP sebelumnya. Sementara itu, atribut seperti nilai rapor dan jarak tempat tinggal memiliki kontribusi lebih kecil. Temuan ini sejalan dengan kebijakan bantuan sosial yang menitikberatkan pada kondisi ekonomi keluarga. Naive Bayes bekerja efektif karena data PIP didominasi fitur kategorikal dan pola distribusinya relatif sederhana. Performa model masih dapat ditingkatkan melalui preprocessing, misalnya pengelompokan kategori yang lebih tepat dan penyeimbangan kelas (contoh: SMOTE). Dibanding metode C4.5 atau K-NN, Naive Bayes memberikan akurasi yang kompetitif dan cenderung stabil. Bagi SMA Laksamana Martadinata, model ini dapat membantu seleksi awal penerima PIP secara lebih objektif, cepat, dan terstandarisasi, serta mengurangi bias subjektif dalam penentuan penerima.

REFERENCES

[1] I. M. B. Adnyana, "Implementasi Naïve Bayes Untuk Memprediksi Waktu Tunggu Alumni Dalam Memperoleh Pekerjaan," *Semin. Nas. Teknol. Komput. Sains*, vol. 1, no. 1, pp. 131–134, 2020.

[2] F. D. Pratama, I. Zufria, and T. Triase, "Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 1, pp. 77–84, 2022, doi: 10.36341/rabit.v7i1.2217.

[3] Gagan Suganda, Marsani Asfi, Ridho Taufiq Subagio, and Ricky Perdana Kusuma, "Penentuan Penerima Bantuan Beasiswa Kartu Indonesia Pintar (Kip) Kuliah Menggunakan Naïve Bayes Classifier," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 193–199, 2022, doi: 10.30656/jsii.v9i2.4376.

[4] A. Pebdika, R. Herdiana, and D. Solihudin, "Klasifikasi Menggunakan Metode *Naive Bayes* Untuk Menentukan Calon Penerima Pip," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 452–458, 2023, doi: 10.36040/jati.v7i1.6303.

[5] W. R. Sari Oktapia Ningse, S. Sumarno, and Z. M. Nasution, "C4.5 Algorithm Classification for Determining Smart Indonesia Program Recipients at MIS Al-Khoirot," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 65–76, 2022, doi: 10.55123/jomlai.v1i1.165.

[6] Y. Shino, Y. Durachman, and N. Sutisna, "Implementation of Data Mining with *Naive Bayes* Algorithm for Eligibility Classification of Basic Food Aid Recipients," *Int. J. Cyber IT Serv. Manag.*, vol. 2, no. 2, pp. 154–162, 2022, doi: 10.34306/ijcitsm.v2i2.114.

[7] A. Amalia, A. Irma Purnamasari, and I. Ali, "Implementasi Algoritma C4.5 Dan Naïve Bayes Dalam Pengambilan Keputusan Untuk Program Indonesia Pintar (Pip) Di Sekolah Dasar Negeri 04 Majalangu," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 1889–1896, 2024, doi: 10.36040/jati.v8i2.8311.

[8] N. Aprilyani, I. Zulfa, and H. Syahputra, "Penerapan Algoritma Decision Tree C4.5 Untuk Model Penentuan Penerima Beasiswa Program Indonesia Pintar (Pip) Studi Kasus Sma Negeri 3 Timang Gajah," *J. Tek. Inform. dan Elektro*, vol. 5, no. 1, pp. 96–109, 2022, doi: 10.55542/jurtie.v5i1.452.

[9] S. Muntari, F. Rahmadayanti, and A. Lovita, "Klasifikasi Kelayakan Penerima Bantuan Sosial Dengan Algoritma Decision Tree," *Escaf*, pp. 1002–1007, 2023, [Online]. Available: https://semnas.univbinainsan.ac.id/index.php/escaf/article/view/469%0Ahttps://semnas.univbinainsan.ac.id/index.php/escaf/article/download/469/303

[10] J. Bramanda, "Klasifikasi Masyarakat Penerima Bantuan Sosial dari Pemerintah dengan Metode Algoritma C4.5," *J. Komput. Antart.*, vol. 3, no. 1, pp. 34–41, 2025, doi: 10.70052/jka.v3i1.234.

[11] I. Bpjs, M. Algoritma, C. Di, and K. Deli, "JDSP Jurnal Data Science Penusa Penerapan Data Mining Untuk Menentukan Kelayakan Penerima Bantuan," vol. 1, no. 1, pp. 166–177, 2024.

[12] C. A. Sugianto and P. N. Sari, "Klasifikasi Kelayakan Penerima Bantuan Langsung Tunai Bagi UMKM di Kota Cimahi Menggunakan Algoritma Naïve Bayes," *J. Informatics Electron. Eng.*, vol. 4, no. 2, p. 64, 2024, doi: 10.70428/jiee.v4i2.1091.

[13] M. Daud, R. Juita, and C. D. Suhendra, "Penerapan Metode Algoritma C4.5 Untuk Klasifikasi Kelayakan Penerima Program Bantuan Pada Dinas Sosial Kabupaten Manokwari," *Decod. J. Pendidik. Teknol. Inf.*, vol. 5, no. 1, pp. 271–278, 2025, doi: 10.51454/decode.v5i1.1057.

[14] A. Data *et al.*, "Analisis Kelayakan Penerima Bantuan Iuran Jaminan Kesehatan Menggunakan Teknik Klasifikasi Data Mining dengan Metode Naïve Bayes," vol. 07, no. 02, pp. 2721–1800, 2025, [Online]. Available: https://journal.cattleyadf.org/index.php/jatilima/index

[15] D. Amalia, P. Lubis, R. A. Putri, and A. M. Harahap, "PENERAPAN DATA MINING UNTUK CLUSTERING KELAYAKAN PENERIMA BPNT MENGGUNAKAN ALGORITMA K-MEANS BERBASIS WEB Kemajuan teknologi informasi sekarang sudah semakin berkembang pesat dan hampir mencangkup di segala bidang kehidupan . Kemajuan tersebut menghasilka," vol. 4307, no. August, pp. 1254–1260, 2024.

[16] D. Subuhanto and L. Tanti, "Model Deteksi Anomali Jaringan Komputer Menggunakan Teknik Machine Learning," *Pros. Semin. Nas. Multi Disiplin Ilmu*, vol. 1, no. 1, pp. 239–259, 2024.

[17] A. Aldiyansyah, A. Irma Purnamasari, and I. Ali, "Perbandingan Tingkat Akurasi Algoritma Decision Tree Dan Random Forest Dalam Mengklasifikasi Penerima Bantuan Sosial Bpnt Di Desa Slangit," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 127–132, 2024, doi: 10.36040/jati.v8i1.8290.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 7, Nomor 2, Desember 2025
e-ISSN 2685-998X
DOI 10.30865/json.v7i2.9160

[18] A. N. A. Rohim, A. I. Purnamasari, and I. Ali, "Komparasi Efektifitas Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Manfaat Program Keluarga Harapan (Studi Kasus : Kecamatan Cicalengka Kabupaten Bandung)," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2355–2362, 2024, [Online]. Available: https://www.ejournal.itn.ac.id/index.php/jati/article/download/8345/5377

[19] D. W. Safitri, E. Fadilah, and S. Rahayu, "BANTUAN PROGRAM KELUARGA PENINJAUAN MENGGUNAKAN METODE," vol. 10, no. 2, pp. 361–368, 2025.

[20] W. Dari and A. Y. Sari, "Implementasi Data Mining Dengan Naïve Bayes Untuk Prediksi Penerima Bantuan Langsung Tunai (Blt) Warga Desa Xyz," *J. Sist. Informasi, dan Teknol. Inf.*, vol. 2, no. 2, pp. 29–40, 2023.