

Klasifikasi Kelayakan Penerima Program Indonesia Pintar di SMPN 2 Tangerang Selatan Menggunakan Metode Naïve Bayes

Sabrina Febrian Komala Dewi ¹, Bani ²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, 15310, Indonesia

¹ sabrfeb22@gmail.com, ² dosen02381@unpam.ac.id

Info Artikel

Riwayat Artikel:

Received July 18, 2026

Revised July 23, 2026

Accepted August 23, 2026

Abstract – The process of identifying potential beneficiaries of the Indonesia Smart Program (PIP) at SMPN 2 South Tangerang is still carried out manually and faces limitations in accessing the Integrated Social Welfare Data (DTKS) and verifying possession of the Indonesia Pintar Card (KIP), which could potentially lead to inaccuracies in determining the target recipients of assistance. This study aims to apply the Naïve Bayes method to classify the eligibility of Indonesia Pintar Program (PIP) beneficiaries and to develop a web-based system that utilizes the Naïve Bayes method as a classification tool. The data used were collected from seventh-grade students at SMPN 2 South Tangerang and included attributes such as parental income, parental occupation, number of siblings, possession of a Social Protection Card (KPS), and possession of an Indonesia Pintar Card (KIP). The model was evaluated using RapidMiner with 333 training samples and 83 test samples, yielding an accuracy of 85.54%, a precision of 86.49%, a recall of 82.05%, and an F1 score of 84.21%. Test results show that all system functions have been validated through black-box testing. This developed web-based classification method enables schools to identify potential PIP recipients more objectively, systematically, and efficiently during the PIP recipient registration process.

Keywords: Classification, Naïve Bayes, Smart Indonesia Program, RapidMiner, Website

Corresponding Author:

Sabrina Febrian Komala Dewi

Email: sabrfeb22@gmail.com



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

Abstrak Indonesia – Proses identifikasi calon penerima manfaat Program Indonesia Pintar (PIP) di SMPN 2 Tangerang Selatan masih dilakukan secara manual dan menghadapi keterbatasan akses terhadap Data Terpadu Kesejahteraan Sosial (DTKS) serta kepemilikan Kartu Indonesia Pintar (KIP), yang berpotensi menyebabkan ketidakakuratan dalam penentuan penerima bantuan. Penelitian ini bertujuan untuk menerapkan metode Naïve Bayes dalam mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP) serta mengembangkan sistem berbasis web yang memanfaatkan metode Naïve Bayes sebagai alat bantu klasifikasi. Data yang digunakan berasal dari data siswa kelas VII SMPN 2 Tangerang Selatan, yang mencakup atribut seperti penghasilan orang tua, pekerjaan orang tua, jumlah saudara kandung, kepemilikan Kartu Perlindungan Sosial (KPS), dan kepemilikan Kartu Indonesia Pintar (KIP). Model dievaluasi menggunakan RapidMiner dengan 333 sampel data latih dan 83 sampel data uji, menghasilkan akurasi sebesar 85,54%, presisi sebesar 86,49%, recall sebesar 82,05%, dan skor F1 sebesar 84,21%. Hasil pengujian menunjukkan bahwa seluruh fungsi sistem telah dinyatakan valid melalui pengujian black box. Metode klasifikasi berbasis web yang dikembangkan ini memungkinkan sekolah untuk mengidentifikasi calon penerima PIP secara lebih objektif, sistematis, dan efisien selama proses pendaftaran penerima PIP.

Kata Kunci: Klasifikasi, Naïve Bayes, Program Indonesia Pintar, RapidMiner, Website

1. PENDAHULUAN

Pendidikan merupakan hak dasar setiap warga negara sekaligus pilar utama dalam pengembangan sumber daya manusia di Indonesia. Namun, kemiskinan menjadi tantangan besar, terutama bagi anak-anak

dari keluarga dengan kondisi ekonomi kurang mampu, karena kondisi tersebut membatasi akses dan keberlanjutan pendidikan mereka. Data dari Badan Pusat Statistik [1] menunjukkan bahwa pada Maret 2025, tingkat kemiskinan di Indonesia mencapai 8,47% atau sekitar 23,85 juta jiwa. Untuk mengatasi ketimpangan akses pendidikan yang disebabkan oleh kendala ekonomi, pemerintah, melalui Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek), meluncurkan Program Indonesia Pintar (PIP). Program ini merupakan bantuan pendidikan nasional yang bertujuan untuk mendukung siswa dari keluarga miskin dan hampir miskin di tingkat sekolah dasar maupun menengah. Meskipun secara nasional Program Indonesia Pintar (PIP) menunjukkan kontribusi positif dalam menurunkan angka putus sekolah dan mendukung keberlanjutan pendidikan peserta didik, implementasi program ini di tingkat satuan pendidikan masih menghadapi sejumlah permasalahan. Salah satu persoalan utama adalah ketidaktepatan sasaran dalam penentuan penerima bantuan. Permasalahan ketidaktepatan sasaran umumnya dipengaruhi oleh kualitas data sosial ekonomi siswa yang belum optimal, proses identifikasi yang masih dilakukan secara manual, serta ketergantungan pada dokumen administratif tanpa memastikan kelengkapan dan kebenaran informasi, serta belum maksimalnya pemanfaatan sistem pendataan terintegrasi. Oleh karena itu, kesesuaian dan sinkronisasi data antara DTKS dan Dapodik sangat penting untuk memastikan keakuratan identifikasi penerima bantuan pendidikan. Perkembangan teknologi informasi memungkinkan penerapan teknik data mining untuk meningkatkan kualitas pengambilan keputusan dalam proses identifikasi penerima bantuan pendidikan dan didukung oleh sistem klasifikasi berbasis data yang terstandar dan terukur. Salah satu metode yang umum digunakan adalah algoritma *Naïve Bayes*, yang terbukti efektif dalam mengolah dan mengklasifikasikan data. Oleh karena itu, penelitian ini dilakukan untuk membantu proses klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) menggunakan metode *Naïve Bayes* yang difokuskan pada siswa kelas VII di SMPN 2 Tangerang Selatan.

2. PENELITIAN TERKAIT

Berbagai penelitian telah dilakukan mengenai klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) dengan menggunakan metode *Naïve Bayes*, yang menunjukkan hasil cukup baik dalam mendukung pengambilan keputusan secara objektif dan berbasis data. Adapun beberapa penelitian terkait yang menjadi rujukan dalam penelitian ini adalah sebagai berikut. Penelitian yang dilakukan oleh Oktavia & Anggreini [2] berfokus pada penerapan metode *Naïve Bayes* untuk mengklasifikasikan kelayakan siswa berdasarkan data kondisi sosial ekonomi sebagai variabel utama. Proses klasifikasi mempertimbangkan beberapa faktor, termasuk jenis tinggal, alat transportasi, pekerjaan dan penghasilan orang tua, serta penerima KPS dan KIP. Pengujian model dilakukan menggunakan RapidMiner dengan metode *Cross Validation*. Metode *Naïve Bayes* digunakan untuk menghitung peluang setiap siswa diklasifikasikan dalam kategori layak atau tidak layak. Hasil penelitian menunjukkan bahwa model tersebut mencapai akurasi sebesar 97,24%, *precision* 72.73%, *recall* 88.89%, dan nilai *AUC* 0.886, yang menunjukkan bahwa metode ini efektif dalam mengklasifikasikan data siswa dan mendukung pengambilan keputusan yang lebih objektif dan berbasis data. Selanjutnya, penelitian yang dilakukan oleh Savitri dkk. [3] mengevaluasi kinerja algoritma *Naïve Bayes* dalam mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP) berdasarkan data sosial ekonomi siswa. Penelitian ini menggunakan dataset yang terdiri dari 172 siswa di SDN 1 Sindang Marga dengan atribut yang meliputi pekerjaan dan penghasilan orang tua, jumlah saudara kandung, urutan anak dalam keluarga, kepemilikan KIP/KPS, serta jenis transportasi ke sekolah. Atribut-atribut tersebut dipilih karena dianggap memiliki hubungan langsung dengan kelayakan siswa dalam menerima bantuan pendidikan. Metode *Naïve Bayes* diterapkan untuk menghitung probabilitas setiap atribut terhadap kelas layak dan tidak layak dengan pengujian model dilakukan dengan *confusion matrix* untuk mengukur nilai akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa kombinasi atribut terbaik adalah pekerjaan orang tua, penghasilan orang tua, kepemilikan KPS, dan kepemilikan KIP, dengan tingkat akurasi sebesar 85,3%, *precision* 92,0%, *recall* 88,5%, dan *F1-score* 90,2%. Ini menunjukkan bahwa pemilihan atribut yang relevan dapat meningkatkan kinerja model dibandingkan menggunakan seluruh atribut yang tersedia.

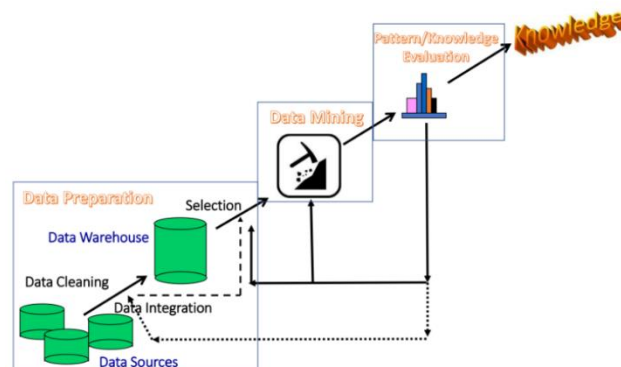
Penelitian yang dilakukan oleh Pratama dkk. [4] menerapkan pendekatan data mining menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan kelayakan berdasarkan beberapa kriteria utama sebagai variabel penilaian, meliputi SKTM (Surat Keterangan Tidak Mampu), KKS (Kartu Keluarga Sejahtera),

kondisi rumah (Papan, Semi Batu, atau Batu), penghasilan orang tua, dan jumlah tanggungan keluarga. Proses penelitian mencakup tahapan transformasi data, perhitungan nilai probabilitas untuk setiap kriteria, hingga penentuan hasil klasifikasi. Selain itu, sistem berbasis web dikembangkan untuk membantu identifikasi, yang dilengkapi dengan fitur login, pengolahan data siswa, serta tampilan hasil klasifikasi. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* dapat mengklasifikasikan data siswa ke dalam kategori layak dan tidak layak secara efektif. Sistem yang dikembangkan juga dapat membantu pihak sekolah dalam mengambil keputusan secara lebih objektif, sistematis, dan berbasis data. Penelitian yang dilakukan oleh Nur dkk. [5] bertujuan untuk mengoptimalkan penerima bantuan PIP di SMAN 1 Sukamulia yang masih dilakukan secara manual dan berpotensi menghasilkan keputusan yang tidak akurat. Penelitian ini menggunakan metode *Naïve Bayes* dengan bantuan RapidMiner serta teknik K-Fold Cross Validation dengan 9 kali iterasi (K-Fold 2 hingga 10) untuk mengevaluasi performa model. Variabel yang digunakan meliputi nama siswa, pekerjaan dan penghasilan orang tua, tanggungan orang tua, kepemilikan KIP (Kartu Indonesia Pintar), dan KPS (Kartu Perlindungan Sosial), serta status dan ranking siswa. Hasil penelitian menunjukkan tingkat akurasi sebesar 92,81% dengan nilai *AUC* 0,973 yang termasuk kategori *excellent classification*, yang menunjukkan bahwa metode ini efektif dalam mendukung proses pengambilan keputusan yang lebih objektif dan berbasis data.

Penelitian lainnya dilakukan oleh Lubis dkk. [6] berfokus pada pengembangan model klasifikasi untuk mentransformasi proses seleksi manual menjadi pendekatan yang lebih sistematis dan objektif. Algoritma *Naïve Bayes* dipilih sebagai instrumen utama untuk mengklasifikasikan data karena efektivitasnya dalam menangani kumpulan data yang besar secara cepat. Variabel penilaian yang dilibatkan dalam perhitungan mencakup alat transportasi, pendidikan ayah, pekerjaan ayah, penghasilan ayah, pendidikan ibu, pekerjaan ibu, penghasilan ibu, serta alasan layak PIP yang mencakup kategori PKH, yatim, miskin atau rentan miskin. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* memberikan kinerja sangat baik dengan tingkat akurasi mencapai 95%. Hal ini membuktikan bahwa pendekatan berbasis data efektif dalam mengurangi subjektivitas dan meningkatkan ketepatan sasaran dalam pemilihan penerima manfaat PIP. Berdasarkan penelitian-penelitian tersebut menunjukkan bahwa metode *Naïve Bayes* efektif dalam mengklasifikasikan kelayakan penerima bantuan pendidikan, khususnya dalam Program Indonesia Pintar (PIP). Meskipun demikian, masih terdapat beberapa keterbatasan, seperti ketergantungan pada kualitas data dan belum adanya penerapan pada konteks dan lokasi tertentu.

3. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah *Naïve Bayes*, yaitu metode klasifikasi probabilistik untuk menentukan kemungkinan suatu data termasuk ke dalam kelas tertentu berdasarkan data sebelumnya, yang diterapkan menggunakan pendekatan data mining dengan tahapan *Knowledge Discovery in Databases* (KDD).



Gambar 1. Tahapan KDD [7]

a. Pengumpulan Data

Mengumpulkan data siswa kelas VII tahun ajaran 2025/2026 yang berkaitan dengan kondisi sosial ekonomi.

- b. Pra-Pemrosesan Data
Melakukan pembersihan data, seleksi atribut, serta transformasi data agar siap digunakan dalam proses klasifikasi.
- c. Pembentukan Model *Naïve Bayes*
Menghitung probabilitas setiap atribut terhadap kelas (layak/tidak layak) menggunakan metode *Naïve Bayes* menggunakan RapidMiner.
- d. Pengujian Model
Menguji model klasifikasi menggunakan data uji untuk mengetahui tingkat akurasi.
- e. Evaluasi Model
Mengevaluasi hasil kinerja model klasifikasi menggunakan parameter seperti akurasi, presisi, *recall*, dan *f1-score*.

Tabel I. Atribut Penelitian

No	Atribut	Keterangan	Nilai
1	Pekerjaan Orang Tua	Pekerjaan yang saat ini dilakukan oleh orang tua	Buruh; Karyawan BUMN; Karyawan Swasta; Pedagang Besar; Pedagang Kecil; Pensiunan; Petani; Peternak; PNS/TNI/Polri; Wiraswasta; Wirausaha; Lainnya; Sudah Meninggal; Tidak bekerja; Tidak dapat diterapkan
2	Penghasilan Orang Tua	Kondisi keuangan orang tua	Kurang dari Rp. 500.000; Rp. 500.000 - Rp. 999.999; Rp. 1.000.000 - Rp. 1.999.999; Rp. 2.000.000 - Rp. 4.999.999; Rp. 5.000.000 - Rp. 20.000.000; Lebih dari Rp. 20.000.00
3	Jumlah Saudara Kandung	Jumlah Saudara Kandung Siswa	Menyesuaikan anak dalam keluarga
4	Penerima KPS	Siswa Penerima Kartu Perlindungan Sosial (KPS)	Ya; Tidak
5	Penerima KIP	Siswa Penerima Kartu Indonesia Pintar (KIP)	Ya; Tidak
6	Status Kelayakan (Label)	Kelayakan siswa untuk mendapatkan bantuan pendidikan ditentukan berdasarkan penilaian terhadap kondisi sosial ekonomi	Layak; Tidak Layak

3.1 Perhitungan Manual Naive Bayes

a. Probabilitas Prior

Probabilitas prior digunakan untuk menghitung probabilitas awal setiap kelas berdasarkan jumlah data pada data training.

$$P(C) = \frac{N_C}{N}$$

Keterangan

$P(C)$ = probabilitas prior kelas C

N_C = jumlah data pada kelas C

N = total data training

b. Probabilitas Likelihood

Likelihood mengukur probabilitas kemunculan suatu atribut dalam suatu kelas tertentu dan digunakan untuk menentukan seberapa kuat atribut tersebut terkait dengan kelas tersebut.

$$P(N_i|C) = \frac{N(X_i, C)}{N_C}$$

Keterangan

$P(N_i|C)$ = probabilitas atribut X_i pada kelas C

$N(X_i, C)$ = jumlah data yang memiliki atribut X_i pada kelas C

N_C = jumlah data pada kelas C

c. Probabilitas Posterior

Probabilitas posterior dihitung dengan mengalikan probabilitas prior dengan probabilitas likelihood dari semua atribut dalam data. Nilai ini merepresentasikan probabilitas data baru diklasifikasikan ke dalam kelas tertentu setelah semua atribut relevan dipertimbangkan.

$$P(C|X) = P(C) \times \prod_{i=1}^5 P(X_i|C)$$

Keterangan

$P(C|X)$ = probabilitas akhir suatu kelas

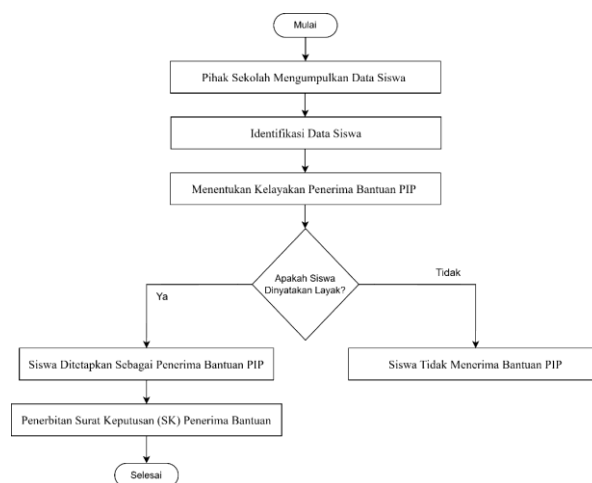
$P(C)$ = probabilitas prior

$P(X_i|C)$ = likelihood setiap atribut

5 = jumlah atribut yang digunakan dalam penelitian

3.2 Analisis Sistem Berjalan

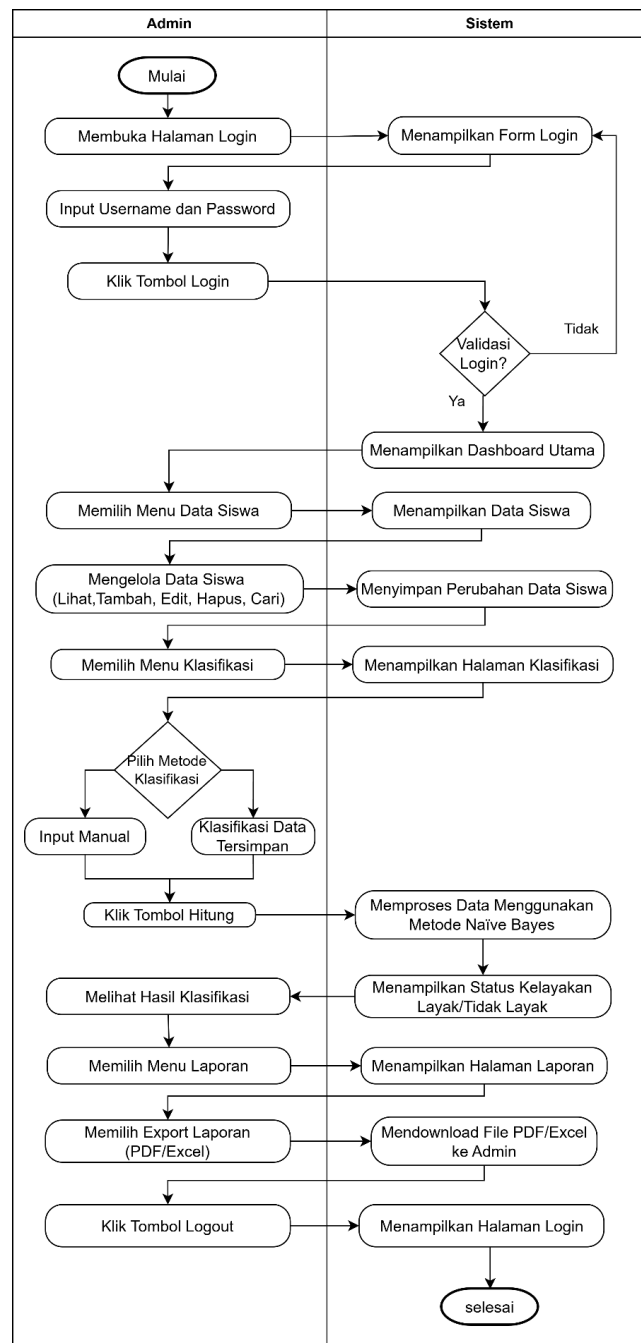
Identifikasi penerima manfaat Program Indonesia Pintar (PIP) di SMPN 2 Tangerang Selatan masih dilakukan secara manual. Prosedur ini dimulai dengan pengumpulan data siswa oleh pihak sekolah, dilanjutkan dengan proses identifikasi kondisi sosial ekonomi siswa. Siswa yang layak akan ditetapkan sebagai penerima bantuan Program Indonesia Pintar (PIP) dan memperoleh Surat Keputusan (SK) penerima bantuan. Namun, karena proses tersebut masih mengandalkan evaluasi manual, berpotensi menyebabkan ketidaktepatan sasaran dalam pengambilan keputusan, serta memakan waktu karena data siswa harus ditinjau satu per satu.



Gambar 2. Sistem Berjalan

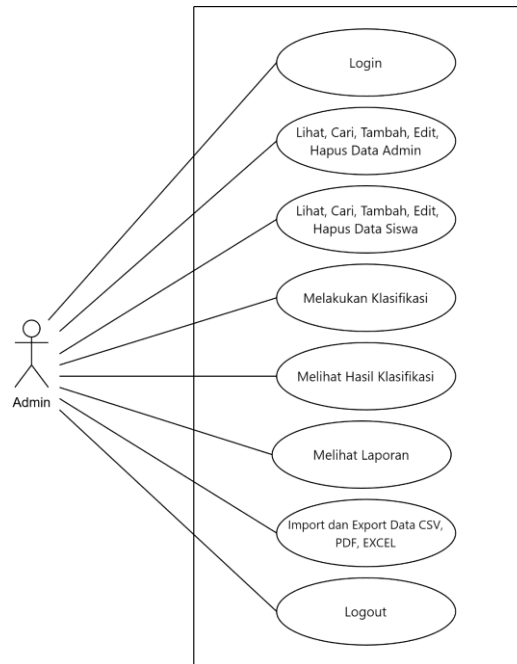
3.3 Analisis Sistem Usulan

Berdasarkan permasalahan pada sistem berjalan, penelitian ini mengusulkan sistem klasifikasi penerima Program Indonesia Pintar (PIP) berbasis web dengan menggunakan metode Naïve Bayes guna mengatasi keterbatasan sistem yang ada saat ini. Sistem yang diusulkan bertujuan untuk membantu proses identifikasi penerima bantuan agar lebih objektif, cepat, dan akurat. Sistem ini mengklasifikasikan siswa berdasarkan data sosial ekonomi yang dimasukkan oleh administrator, kemudian menghasilkan keputusan berupa kategori layak atau tidak layak menerima bantuan PIP. Selain itu, sistem juga dapat membantu pihak sekolah mengelola data siswa serta mendukung proses pengambilan keputusan dalam penentuan penerima bantuan PIP.



Gambar 3. Sistem Usulan

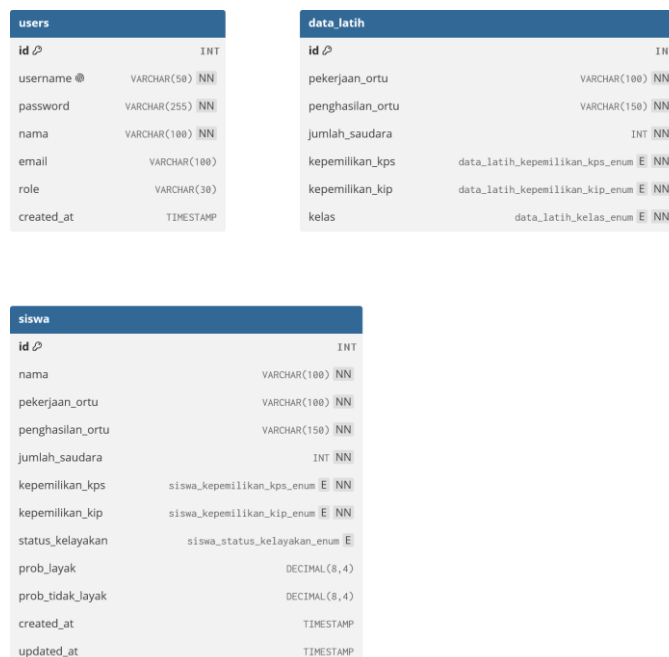
3.4 Perancangan Use Case Diagram



Gambar 4. Use Case Diagram

Pada sistem ini, admin berperan sebagai aktor tunggal yang memiliki akses penuh. Delapan hak akses utama yang diberikan kepada admin, yaitu melakukan login dan logout, mengelola data admin (melihat, mencari, menambahkan, mengedit, dan menghapus), mengelola data siswa dengan hak akses yang sama, melakukan proses klasifikasi kelayakan penerima bantuan, melihat hasil klasifikasi, melihat laporan, serta melakukan import dan export data dalam format CSV, PDF, dan Excel.

3.5 Perancangan Basis Data

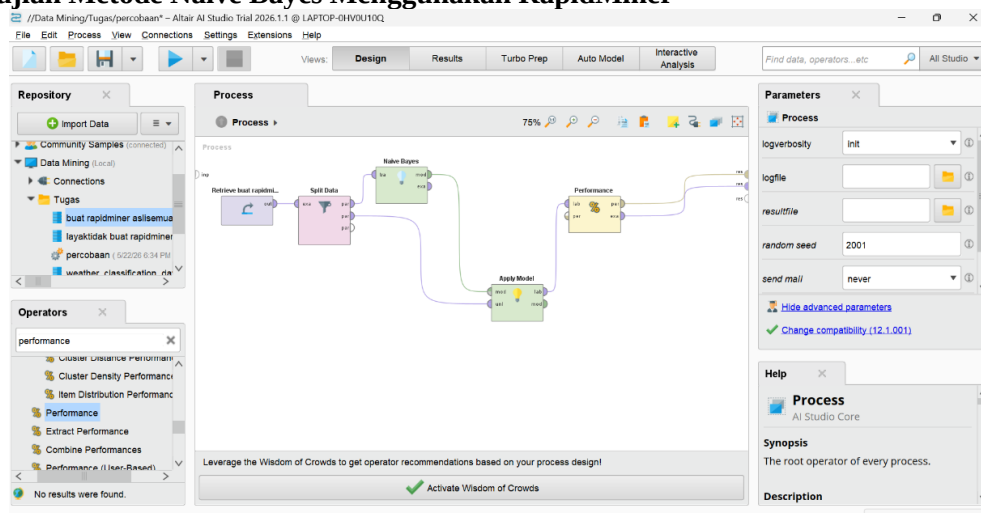


Gambar 5. Entity Relationship Diagram

Entity Relationship Diagram (ERD) pada sistem klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) menggambarkan struktur basis data yang digunakan untuk mendukung proses pengelolaan data dan klasifikasi menggunakan metode *Naïve Bayes*. ERD ini terdiri dari tiga entitas utama, yaitu users, siswa, dan data latih.

4 HASIL DAN PEMBAHASAN

4.1 Pengujian Metode *Naïve Bayes* Menggunakan RapidMiner



Gambar 6. Rancangan Proses Klasifikasi *Naïve Bayes* pada RapidMiner

Pemrosesan data dilakukan dengan membangun alur kerja (workflow) metode *Naïve Bayes* di RapidMiner untuk mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP). Proses dimulai dengan operator Retrieve untuk mengambil data penelitian dari repository RapidMiner. Selanjutnya, operator Split Data diterapkan untuk membagi data menjadi data latih (training data) dan data uji (testing data) dengan rasio 80:20, di mana 333 data digunakan sebagai data latih dan 83 data digunakan sebagai data uji. Algoritma *Naïve Bayes* kemudian dilatih menggunakan data latih untuk membuat model klasifikasi yang berasal dari atribut prediktor. Model ini kemudian diterapkan pada data uji menggunakan operator Apply Model untuk memprediksi kelayakan siswa. Pada tahap akhir operator Performance mengevaluasi kinerja model secara keseluruhan melalui pengukuran *accuracy*, *precision*, *recall*, *F1-Score* serta menampilkan confusion matrix. Seluruh operator dihubungkan secara berurutan sehingga membentuk alur kerja yang terintegrasi.

	true Layak	true Tidak Layak
pred. Layak	32	5
pred. Tidak Layak	7	39

Gambar 7. Confusion Matrix Model *Naive Bayes*

Hasil evaluasi dari matriks kebingungan (confusion matrix) menunjukkan kinerja model *Naïve Bayes* dalam mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP) pada 83 data uji.

accuracy: 85.54%

	true Layak	true Tidak Layak	class precision
pred. Layak	32	5	86.49%
pred. Tidak Layak	7	39	84.78%
class recall	82.05%	88.64%	

Gambar 8. Hasil Pengujian Model *Naive Bayes*

Berdasarkan hasil evaluasi model *Naïve Bayes* terhadap 83 data uji menghasilkan akurasi sebesar 85,54%, yang berarti model tersebut berhasil mengklasifikasikan 71 data dengan benar, sementara 12 data lainnya salah diklasifikasikan. Model mencapai presisi sebesar 86,49% untuk kelas Layak dan 84,78% untuk kelas Tidak Layak, yang menunjukkan akurasi prediksi untuk setiap kategori. Nilai *recall* untuk kelas Layak sebesar 82,05% menunjukkan kemampuan model dalam mengidentifikasi data siswa yang sebenarnya Layak, sedangkan *recall* untuk kelas Tidak Layak sebesar 88,64% menunjukkan bahwa model mampu mengidentifikasi data siswa yang sebenarnya Tidak Layak. Selain itu, model memperoleh nilai *F1-Score* sebesar 84,21% mencerminkan keseimbangan yang baik antara presisi dan *recall*. Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa model *Naïve Bayes* memiliki kinerja baik dalam mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP).

4.2 Perhitungan Manual *Naïve Bayes*

a. Probabilitas Prior

Perhitungan probabilitas prior untuk kelas Layak dihitung dengan membagi jumlah data yang diklasifikasikan sebagai kelas Layak dengan total data latih.

$$\begin{aligned} P(\text{Layak}) &= \frac{N_{\text{Layak}}}{N} \\ &= \frac{158}{333} \\ &= 0,475 \end{aligned}$$

Setelah menghitung probabilitas prior kelas Layak, perhitungan yang sama dilakukan untuk kelas Tidak Layak dengan membandingkan jumlah data Tidak Layak dengan jumlah total data dalam data latih.

$$\begin{aligned} P(\text{Tidak Layak}) &= \frac{N_{\text{Tidak Layak}}}{N} \\ &= \frac{175}{333} \\ &= 0,525 \end{aligned}$$

b. Probabilitas Likelihood

1. Atribut Pekerjaan Orang Tua (Buruh)

$$\begin{aligned} P(\text{Buruh}|\text{Layak}) &= \frac{23}{158} \\ &= 0,146 \\ P(\text{Buruh}|\text{Tidak Layak}) &= \frac{8}{175} \\ &= 0,046 \end{aligned}$$

2. Penerima KIP

a. Ya

$$\begin{aligned} P(\text{Ya}|\text{Layak}) &= \frac{74}{158} \\ &= 0,469 \\ P(\text{Ya}|\text{Tidak Layak}) &= \frac{0}{175} \\ &= 0 \end{aligned}$$

b. Tidak

$$\begin{aligned} P(\text{Tidak}|\text{Layak}) &= \frac{84}{158} \\ &= 0,532 \end{aligned}$$

$$\begin{aligned} P(\text{Tidak}|\text{Tidak Layak}) &= \frac{175}{175} \\ &= 1 \end{aligned}$$

c. Probabilitas Posterior

Perhitungan dimulai dari kelas Layak menggunakan nilai prior dan likelihood yang telah diperoleh pada tahap sebelumnya. Proses perhitungan tersebut ditunjukkan sebagai berikut:

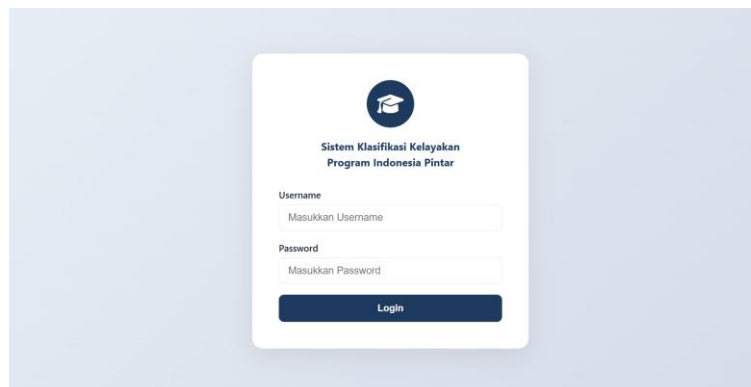
$$\begin{aligned} P(\text{Layak}|X) &= P(\text{Layak}) \times \prod_{i=1}^5 P(X_i|\text{Layak}) \\ P(\text{Layak}|X) &= 0,475 \times 0,146 \times 0,627 \times 0,703 \times 0,968 \times 0,532 \\ &= 0,016 \end{aligned}$$

Selanjutnya, perhitungan probabilitas posterior dilakukan pada kelas Tidak Layak dengan memanfaatkan nilai prior dan likelihood pada kelas tersebut.

$$\begin{aligned} P(\text{Tidak Layak}|X) &= P(\text{Tidak Layak}) \times \prod_{i=1}^5 P(X_i|\text{Tidak Layak}) \\ P(\text{Tidak Layak}|X) &= 0,525 \times 0,046 \times 0,274 \times 0,623 \times 1 \times 1 \\ &= 0,004 \end{aligned}$$

Nilai probabilitas kelas Layak sebesar 0,016 lebih tinggi daripada nilai probabilitas kelas Tidak Layak sebesar 0,004, maka siswa tersebut diklasifikasikan sebagai Layak menerima bantuan Program Indonesia Pintar (PIP).

4.3 Implementasi Sistem



Gambar 9. Tampilan Halaman Login



Gambar 10. Tampilan Halaman Dashboard

Gambar 11. Tampilan Halaman Klasifikasi Input Manual

No	Nama	Pekerjaan Orang Tua	Penghasilan Orang Tua	Jumlah Saudara	Penerima KPS	Penerima KIP
1	sab	Buruh	Rp. 500.000 - Rp. 999.999	2	Tidak	Tidak
2	a	Wiraswasta	Rp. 2.000.000 - Rp. 4.999.999	1	Ya	Tidak
3	a	Wiraswasta	Rp. 2.000.000 - Rp. 4.999.999	1	Tidak	Tidak
4	a	Wiraswasta	Rp. 2.000.000 - Rp. 4.999.999	1	Tidak	Tidak

Gambar 12. Tampilan Halaman Klasifikasi Data Tersimpan

No	Nama	Penghasilan Orang Tua	Jumlah Saudara Kandung	Penerima KPS	Penerima KIP	Hasil
11	S-042	Rp. 1.000.000 - Rp. 1.999.999	1	Tidak	Ya	Layak
12	S-046	Rp. 2.000.000 - Rp. 4.999.999	1	Tidak	Tidak	Tidak Layak
13	S-053	Rp. 1.000.000 - Rp. 1.999.999	1	Tidak	Tidak	Layak
14	S-054	Tidak Berpenghasilan	1	Tidak	Tidak	Layak
15	S-060	Rp. 5.000.000 - Rp. 20.000.000	1	Tidak	Ya	Tidak Layak

Gambar 13. Tampilan Halaman Laporan

5 KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi kelayakan penerima Program Indonesia Pintar (PIP) di SMPN 2 Tangerang Selatan dengan menggunakan metode *Naïve Bayes* menghasilkan kesimpulan bahwa metode *Naïve Bayes* telah berhasil diterapkan secara efektif untuk mengklasifikasikan kelayakan penerima Program Indonesia Pintar (PIP) di kalangan siswa kelas VII SMPN 2 Tangerang Selatan, berdasarkan penghasilan orang tua, pekerjaan orang tua, jumlah saudara kandung, kepemilikan Kartu Perlindungan Sosial (KPS), dan kepemilikan Kartu Indonesia Pintar (KIP). Pengujian model menggunakan RapidMiner menunjukkan bahwa metode *Naïve Bayes* mampu mengklasifikasikan kelayakan penerima PIP dengan akurasi sebesar 85,54%, presisi sebesar 86,49%, *recall* sebesar 82,05%,

dan skor F1 sebesar 84,21%.. Hasil tersebut menunjukkan bahwa metode *Naïve Bayes* dapat membantu mengklasifikasikan kelayakan penerima PIP berdasarkan data yang tersedia. Sistem berbasis web ini menerapkan pendekatan *Naïve Bayes* untuk mengkategorikan kelayakan penerima PIP secara otomatis, sehingga menghasilkan kategori “layak” atau “tidak layak” dari hasil klasifikasi tersebut. Hasil pengujian black box menunjukkan bahwa seluruh fungsi sistem, termasuk proses masuk (login), pengelolaan data admin, pengelolaan data siswa, proses klasifikasi, laporan, dan proses keluar (logout), beroperasi sesuai dengan yang diharapkan.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik, “Tingkat Kemiskinan Kembali Menurun,” bps.go.id. [Online]. Available: <https://www.bps.go.id/id/news/2025/07/25/731/poverty-rate-continues-to-decline.html>
- [2] I. P. Oktavia and N. L. Anggreini, “IMPLEMENTASI ALGORITMA NAIVE BAYES DENGAN METODE KLASIFIKASI DALAM MENENTUKAN SISWA PENERIMA BANTUAN PROGRAM INDONESIA PINTAR (STUDI KASUS : SMPN 3 CIHAMPÉLAS),” vol. 8, no. 6, pp. 11152–11158, 2024.
- [3] V. K. Savitri, H. Setiawan, and Z. R. Mair, “Classification Of Eligibility For Assistance Recipients Program Indonesia Pintar Using The Naïve Bayes Method,” *bit-Tech*, vol. 8, no. 2, pp. 1571–1582, 2025, doi: 10.32877/bt.v8i2.3002.
- [4] F. D. Pratama, I. Zufria, and Triase, “Implementasi Data Mining Menggunakan Algoritma Naïve Bayes untuk Klasifikasi Penerima Program Indonesia Pintar,” vol. 7, no. 1, pp. 77–84, 2022.
- [5] A. M. Nur, W. Nursali, Nurhidayati, and I. Fathurrahman, “Penerapan Metode Naïve Bayes Untuk Penentuan Penerima Beasiswa Program Indonesia Pintar,” vol. 7, no. 1, pp. 93–102, 2024.
- [6] A. S. Lubis, L. Tanti, and R. Puspasari, “Klasifikasi Kelayakan Penerima Program Indonesia Pintar (PIP) Menggunakan Teknik Data Mining Naive Bayes,” vol. 7, pp. 568–580, 2025, doi: 10.30865/json.v7i2.9160.
- [7] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*. Morgan kaufmann, 2022.
- [8] Pusat Layanan Pembiayaan Pendidikan, “Program Indonesia Pintar,” Kementerian Pendidikan Dasar dan Menengah. [Online]. Available: https://pip.kemendikdasmen.go.id/home_v1
- [9] Yanuar, “Beberapa Penyebab Siswa Miskin dan Rentan Miskin Gagal Memeroleh PIP,” Pusat Layanan Pembiayaan Pendidikan (Puslapdik) Kementerian Pendidikan Dasar dan Menengah. Accessed: May 30, 2026. [Online]. Available: <https://puslapdik.kemendikdasmen.go.id/beberapa-penyebab-siswa-miskin-dan-rentan-miskin-gagal-memeroleh-pip/>
- [10] Y. Jatnika, “9,7 Juta Siswa Telah Menerima Bantuan PIP Tahun 2024.” [Online]. Available: <https://puslapdik.kemendikdasmen.go.id/97-juta-siswa-telah-menerima-bantuan-pip-tahun-2024/>
- [11] O. Peretz, M. Koren, and O. Koren, “Engineering Applications of Artificial Intelligence Naive Bayes classifier – An ensemble procedure for recall and precision enrichment,” *Eng. Appl. Artif. Intell.*, vol. 136, no. PB, p. 108972, 2024, doi: 10.1016/j.engappai.2024.108972.
- [12] S. C. Pratama and Y. Suprayogi, “Optimalisasi Program Indonesia Pintar (PIP) Dalam Upaya Peningkatan Kesetaraan Pendidikan Di Provinsi Jawa Tengah Sofi,” vol. 10, no. 11, pp. 799–811, 2024.
- [13] N. Sharma, R. Boghey, and P. R. Prasad, “A Review on Data Mining Issues, Solution & Techniques,” *Int. J. Multidiscip. Res.*, vol. 6, no. 4, pp. 1–9, 2024, doi: 10.36948/ijfmr.2024.v06i04.26654.
- [14] X. Shu and Y. Ye, “Knowledge Discovery: Methods from data mining and machine learning,” *Soc. Sci. Res.*, vol. 110, no. April 2022, p. 102817, 2023, doi: 10.1016/j.ssresearch.2022.102817.
- [15] A. Jeyaraj, “Integrating RapidMiner in Business Analytics Education: An Instructional Approach for Skill Development,” *Am. J. Educ. Learn.*, vol. 10, no. 1, pp. 104–116, 2025, doi: 10.55284/ajel.v10i1.1458.