EXPERT JURNAL MANAJEMEN SISTEM INFORMASI DAN TEKNOLOGI

Penerapan Naive Bayes Classifier untuk Prediksi Kelayakan Penerima Bantuan Program Indonesia Pintar (PIP)

Zulfy Hidayat*[1], Rakhmat Kurniawan. R[2]

[1] Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, UIN Sumatera Utara, Medan.
[1] zulfyhidayat@gmail.com*[1] , rakhmat.kr@uinsu.ac.id[2]

ABSTRACT – *This research is motivated by the inaccuracy in the manual selection process for recipients of the Smart Indonesia Program (PIP) assistance at MTs Al-Hasanah, which is not yet based on objective data so that it has the potential to distribute assistance that is not on target. Based on this problem, the formulation of this research is how to apply the Naïve Bayes algorithm in predicting student eligibility as PIP assistance recipients, and how accurate this method is in the prediction process. The purpose of this study is to build a Naïve Bayes-based classification model that is able to predict the eligibility of students receiving assistance based on attributes such as family status, KIP ownership, number of dependents, and parental income. The model was developed using a data mining approach with the Naïve Bayes algorithm on the Google Colab platform, with preprocessing using the Label Encoding technique and data division with a ratio of 80:20, resulting in 84 training data and 21 test data. The evaluation results show accuracy, precision, recall, and F1-score values of 1.0 or 100%. This perfect performance indicates the success of the model in classifying all test data correctly. However, very high metric values also indicate potential overfitting, given the limited dataset size and homogeneity of the data used. Therefore, the generalization of the model to broader data cannot be ascertained. This research is useful in supporting data-based decision making in the distribution of PIP assistance more objectively and fairly, and is an initial contribution in the development of artificial intelligence-based systems for the education sector.*

Keywords: Data Mining; Machine Learning Naive Bayes Algorithm; Smart Indonesia Program (PIP).

ABSTRAK – Penelitian ini dilatarbelakangi oleh ketidaktepatan dalam proses seleksi manual penerima bantuan Program Indonesia Pintar (PIP) di MTs Al-Hasanah, yang belum berbasis data objektif sehingga berpotensi menyalurkan bantuan tidak tepat sasaran. Berdasarkan masalah tersebut, rumusan penelitian ini adalah bagaimana penerapan algoritma Naïve Bayes dalam meramalkan kelayakan siswa sebagai penerima bantuan PIP, serta seberapa akurat metode ini dalam proses prediksi. Tujuan dari penelitian ini adalah untuk membangun model klasifikasi berbasis Naïve Bayes yang mampu memprediksi kelayakan siswa penerima bantuan berdasarkan atribut seperti status keluarga, kepemilikan KIP, jumlah tanggungan, dan penghasilan orang tua. Model dikembangkan menggunakan pendekatan data mining dengan algoritma Naïve Bayes pada platform Google Colab, dengan preprocessing menggunakan teknik Label Encoding dan pembagian data dengan rasio 80:20, menghasilkan 84 data latih dan 21 data uji. Hasil evaluasi menunjukkan nilai akurasi, presisi, recall, dan F1-score sebesar 1.0 atau 100%. Kinerja sempurna ini menunjukkan keberhasilan model dalam mengklasifikasikan seluruh data uji secara tepat. Namun demikian, nilai metrik yang sangat tinggi juga mengindikasikan potensi overfitting, mengingat ukuran dataset yang terbatas dan homogenitas data yang digunakan. Oleh karena itu, generalisasi model terhadap data yang lebih luas belum dapat dipastikan. Penelitian ini bermanfaat dalam mendukung pengambilan keputusan berbasis data dalam distribusi bantuan PIP secara lebih objektif dan adil, serta menjadi kontribusi awal dalam pengembangan sistem berbasis kecerdasan buatan untuk sektor pendidikan.

Kata Kunci: Algoritma Naive Bayes; Data Mining; Machine Learning; Program Indonesia Pintar (PIP).

1. PENDAHULUAN

Pendidikan memiliki peranan penting dalam membentuk dan mengembangkan kepribadian seseorang secara mental maupun fisik. Melalui pendidikan, seseorang dapat mengatasi buta huruf, mengasah keterampilan, serta meningkatkan kecakapan berpikir dan sosial [1]. Pendidikan juga merupakan hak asasi manusia yang paling mendasar dan menjadi kewajiban negara dalam mempersiapkan serta mengembangkan sumber daya manusia (SDM) untuk kemajuan bangsa [2].

Namun, proses pendidikan di Indonesia masih menghadapi berbagai tantangan, salah satunya adalah kendala ekonomi. Banyak anak terpaksa menghentikan pendidikan mereka karena keterbatasan biaya. Sehingga berkurangnya generasi yang berkualitas dan menyebabkan banyak anak yang minim literasi. Sebagai tanggapan terhadap kesulitan ekonomi yang memaksa

This work is licensed under a Creative Commons Attribution 4.0 International License

banyak anak untuk berhenti sekolah sehingga pemerintah membuat kebijakan dengan menciptakan program PIP untuk menyelesaikan permasalahan terkait peristiwa putus sekolah kepada generasi anak muda [3].

Dalam hal mengatasi permasalahan peristiwa putus sekolah karena kesulitan ekonomi pemerintah menawarkan program dalam wujud pembiayaan pendidikan yang langsung diberikan kepada siswa, program tersebut disahkan pada tahun 2015 lewat Permendikbud No. 12 Tahun 2015 tanggal 12 Mei 2015 tentang Program Indonesia Pintar (PIP). Program ini diperuntukkan bagi individu dari keluarga kurang sanggup secara ekonomi yang ditandai menggunakan Kartu Indonesia Pintar (KIP)[4]. Adapun kriteria yang mendapatkan beasiswa tersebut diukur melaui, penghasilan orang tua, pekerjaan orang tua, total tanggungan orang tua dan status kepemilikan rumah.

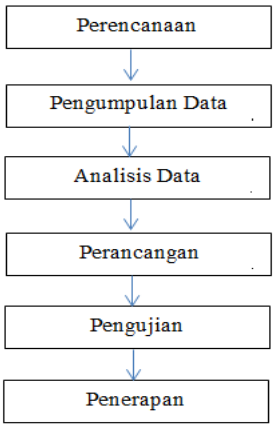
Program Beasiswa PIP sudah dicanangkan pemerintah sejak tahun 2014 dan mulai terlaksana di kecamatan Tanjung Leidong sejak tahun 2019, dan terlaksana di beberapa sekolah yang ada di kabupaten Labuhan Batu Utara sampai saat ini. MTs Al-Hasanah Tj. Leidong merupakan salah satu sekolah yang merasakan dampak program tersebut. Proses seleksi penyaluran beasiswa dapat dilakukan bervariasi disetiap sekolah, namun pada sekolah MTs Al-Hasanah, proses seleksi penerima bantuan masih dilakukan secara manual masih terdapat kendala dalam proses pengambilan keputusan terkait identifikasi peserta didik yang berhak menerima bantuan pemerintah secara langsung. Hal ini disebabkan oleh ketidaktepatan cara penilaian yang digunakan saat ini, karena belum ada model yang dapat dengan tepat memprediksi penerima PIP sesuai dengan target yang ditetapkan. Oleh karena itu, diperlukan suatu sistem yang dapat memberikan dukungan kepada sekolah dalam mengambil keputusan terkait penerimaan bantuan PIP.

Untuk menghasilkan informasi yang tepat mengenai peramalan penerima beasiswa yang sudah diprogramkan pemerintah, perlu diolah dengan teknik data mining. Data mining merupakan suatu proses untuk mengumpulkan informasi yang berharga dari sekumpulan data yang besar. Dalam proses ini, metode statistika dan matematika sering digunakan untuk menganalisis data tersebut [5]. Salah satu metode dalam data mining yang efektif untuk prediksi adalah algoritma Naïve Bayes. Naïve Bayes disebut juga multinomial yang merupakan metode klasifikasi yang dapat memprediksi dengan cara memanfaatkan probabilitas dan statistic[6][7].

Dalam hal ini, algoritma Naive Bayes digunakan untuk mengklasifikasi siswa penerima beasiswa Indonesia Pintar di MTs Al-Hasanah. Dengan menerapkan algoritma Naive Bayes, sistem prediksi penerima beasiswa di MTs Al-Hasanah dapat memberikan informasi awal yang akurat kepada pihak sekolah. Algoritma ini mampu mengklasifikasikan siswa berdasarkan data historis serta kriteria seperti kondisi ekonomi dan prestasi akademik. Hasil klasifikasi ini diharapkan membantu sekolah dalam merancang strategi penyaluran bantuan yang lebih tepat sasaran dan adil, sehingga beasiswa benar-benar diterima oleh siswa yang paling membutuhkan.. Dengan uraian dan latar belakang masalah di atas, maka peneliti menjelaskan tentang "Analisis Algoritma Naive Bayes Untuk Klasifikasi Calon Penerima Bantuan Program Indonesia Pintar".

2. METODOLOGI

Tahapan dari metode yang diusulkan ditunjukkan pada Gambar 1. Proses dimulai dari Perencanaan, pengumpulan data, Analisis Data, Pengujian, Penerapan.



Gambar 1. Alur Metode Penelitian

Perencanaan

Pada tahapan ini merupakan tahapan awal dari penelitian yang dilakukan oleh penulis. Di mana tahapan ini mengacu pada permasalahan umum yang berasal dari fenomena yang terjadi khususnya di MTs Al-Hasanah. Fenomena yang diamati selama ini adalah ketidaktepatan dalam proses seleksi penerima beasiswa Program Indonesia Pintar (PIP), yang masih dilakukan secara manual tanpa analisis data yang sistematis. Hal ini menyebabkan adanya potensi ketidaktepatan sasaran penerima beasiswa, sehingga bantuan tidak sepenuhnya diterima oleh siswa yang benar-benar membutuhkan. Pada tahap perencanaan ini peneliti melakukan Pengumpulan Data Siswa sesuai latar belakang ekonomi keluarga, prestasi Akademik, kepemilikan KIP/PKH serta keaktifan siswa di sekolah, Mengkaji fenomena dengan merujuk kepada penelitian terdahulu untuk menemukan metode apa yang digunakan dalam penelitian. Merumuskan penyelesaian permasalahan dengan menggunakan metode yang dipandang sesuai, berdasarkan kepada riset-riset yang pernah dilakukan. Menyusun kerangka penelitian., kemudian Menentukan tempat dan waktu penelitian [8].

Pengumpulan Data

Pada tahapan ini penulis mengumpulkan data melalui observasi dan wawancara. Observasi dilakukan untuk mengambil data calon penerima PIP melalui tata usaha di kelas IX MTs Al-Hasanah Tanjung Leidong. wawancara dengan wakil kepala sekolah yang bertanggungjawab dalam pengelolaan PIP. Selain itu peneliti mengumpulkan data terkait jurnal maupun dari website yang relevan dengan penelitian terkait Analisis

This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 23

Algoritma Naive Bayes Untuk Klasifikasi Calon Penerima Bantuan Program Indonesia Pintar .

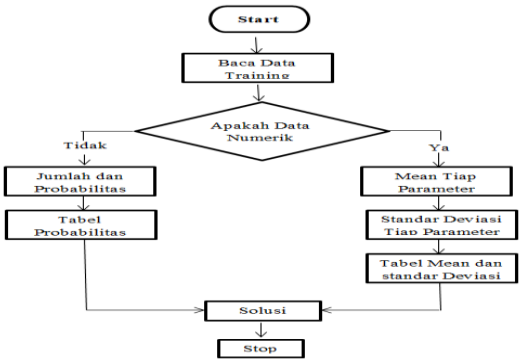
Analisis Data

Dalam penyusunan penelitian ini terdapat beberapa tahapan yang perlu dilakukan, dimulai dari penyusunan atribut yang digunakan, yaitu kepemilikan bantuan sosial seperti Program Keluarga Harapan (PKH) dan Kartu Indonesia Pintar (KIP), jumlah tanggungan orang tua dan pendapatan orang tua. Data yang digunakan merupakan data pribadi siswa MTs Al-Hasanah yang memiliki atau tidak memiliki bantuan sosial tersebut. Setelah data terkumpul, dilakukan proses pembersihan data (data cleaning) untuk memastikan bahwa data bebas dari nilai kosong, duplikasi, atau ketidaksesuaian format. Selanjutnya, pada tahap klasifikasi, dataset akan diproses menggunakan algoritma Naive Bayes untuk memprediksi kelayakan siswa sebagai penerima beasiswa PIP berdasarkan kepemilikan bantuan sosial.

Pada tahap akhir, model prediksi akan dievaluasi menggunakan Confusion Matrix untuk mendapatkan nilai accuracy, precision, dan recall, sehingga dapat mengukur seberapa akurat dan efektif model dalam memprediksi siswa penerima beasiswa.

Perancangan

Berikut adalah flowchart naive bayes yang bertujuan untuk menyelesaikan permasalahan pada penyaluran beasiswa PIP di Mts Al-Hasanah dapat dilihat pada gambar 2.



Gambar 2. Flowchart Algoritma Naive Bayes (Sumber: Alfa Saleh, 2015)[9]

Pengujian

Pada tahap pengujian ini, dilakukan proses untuk mengetahui seberapa baik model Naive Bayes yang dibangun dalam memprediksi kelayakan siswa penerima beasiswa PIP berdasarkan atribut kepemilikan bantuan sosial (PKH dan KIP). Pengujian dilakukan dengan metode hold-out validation, yaitu membagi data menjadi data latih sebesar 80% dan data uji sebesar 20%. Seluruh proses dilakukan menggunakan Google Colab sebagai alat bantu untuk pengolahan data dan pembangunan model [2].

3. HASIL DAN PEMBAHASAN

Pengumpulan Data

Pada penelitian ini, penulis melakukan proses pengumpulan data melalui dokumentasi dan observasi terhadap data peserta didik yang berpotensi menerima bantuan Program Indonesia Pintar (PIP). Data diperoleh dari MTs Al-Hasanah Tj. Leidong, Kecamatan Kualuh Leidong, Kabupaten Labuhan Batu Utara, Sumatera Utara dimana mencakup 105 data siswa dengan berbagai atribut, yaitu nama lengkap, NISN, NIK, tempat lahir, tanggal lahir, tingkat-rombel, umur, status keluarga, jenis kelamin, alamat, kebutuhan khusus, status kepemilikan Kartu Indonesia Pintar (KIP), nama ayah kandung, nama ibu kandung, nama wali, Pekerjaan Orang Tua, Jumlah Tanggungan Orang Tua, dan Penghasilan Orang Tua.

Data tersebut digunakan sebagai dasar dalam membangun model klasifikasi menggunakan algoritma Naive Bayes, yang bertujuan untuk memprediksi apakah seorang siswa layak atau tidak layak menerima bantuan PIP. Sebelum digunakan dalam proses pelatihan model, dilakukan tahap *pre-processing* untuk menyiapkan data. Tahapan ini meliputi penanganan nilai kosong, serta transformasi atribut non-numerik ke dalam bentuk numerik menggunakan metode *LabelEncoder*, sehingga seluruh data dapat dibaca oleh algoritma klasifikasi [1][2].

Hasil dari proses pre-processing kemudian disimpan dalam format CSV dan dibagi menjadi data latih dan data uji dengan rasio 80:20. Model yang dibangun diuji untuk mengukur kinerja menggunakan beberapa metrik evaluasi, seperti *accuracy*, *precision*, *recall*, dan F1-*Score*. Tujuan dari evaluasi ini adalah untuk mengetahui tingkat efektivitas algoritma Naive Bayes dalam mengklasifikasikan siswa sebagai penerima atau bukan penerima bantuan PIP. Visualisasi hasil klasifikasi juga digunakan untuk mendukung analisis dan interpretasi temuan penelitian.

Dataset Penelitian

Data yang digunakan dalam penelitian ini terdiri dari 105 data siswa yang diperoleh dari MTs Al-Hasanah Tj. Leidong, Kecamatan Kualuh Leidong, Kabupaten Labuhan Batu Utara. Atribut dalam dataset mencakup Nama Lengkap, NISN, NIK, Tempat Lahir, Tanggal Lahir, Tingkat-Rombel, Umur, Status (Yatim, Piatu, Yatim Piatu, Lengkap), Jenis Kelamin, Alamat, Kebutuhan Khusus, Status KIP (Ada/Tidak Ada), Nama Ayah Kandung, Nama Ibu Kandung, Nama Wali, Pekerjaan Orang Tua, Jumlah Tanggungan Orang Tua, dan Penghasilan Orang Tua. Untuk kebutuhan pemodelan, tidak semua atribut digunakan, hanya atribut yang dianggap berpengaruh terhadap klasifikasi penerima bantuan PIP.

Untuk keperluan pelatihan dan pengujian model, data dibagi dengan rasio 80:20. Meskipun jumlah data relatif kecil, pembagian ini dipilih karena merupakan praktik umum dalam pembelajaran mesin untuk menghindari overfitting dan tetap memungkinkan evaluasi model secara objektif, dengan mempertahankan proporsi yang cukup untuk uji validitas performa model.

Model prediksi status penerima PIP dibangun berdasarkan beberapa informasi kunci dari atribut-atribut tersebut. Misalnya, atribut "Status" menggambarkan

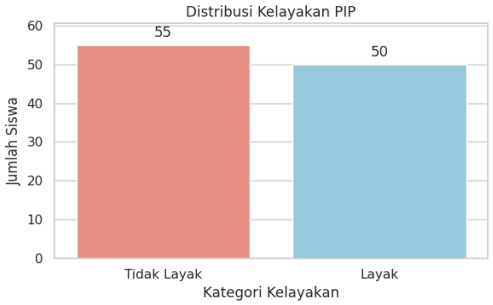
This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 24

kondisi keluarga siswa, "Status KIP" menunjukkan apakah siswa telah memiliki Kartu Indonesia Pintar, dan atribut "Jenis Kelamin" serta "Umur" digunakan untuk mengelompokkan karakteristik siswa. Atribut-atribut ini digunakan sebagai dasar untuk memprediksi apakah seorang siswa layak menerima bantuan PIP.

Dalam proses penelitian ini, penulis menggunakan *platform Google Colab* untuk pengolahan data. Google Colab dipilih karena kemampuannya dalam menjalankan kode Python secara efisien serta integrasinya dengan pustaka-pustaka machine learning seperti pandas, sklearn, dan numpy. Dataset yang telah dipreprocessing disimpan dalam format .csv, dan algoritma Naive Bayes digunakan untuk membangun serta mengevaluasi model klasifikasi berdasarkan data tersebut.

Selanjutnya melakukan pembersihan data dengan menghapus beberapa kolom yang dianggap tidak relevan atau tidak memberikan kontribusi langsung terhadap proses klasifikasi status kelayakan penerima bantuan Program Indonesia Pintar (PIP). Kolom-kolom yang dihapus meliputi: Nama Lengkap, NISN, NIK, Tempat Lahir, Tanggal Lahir, Tingkat - Rombel, Alamat, Nama Ayah Kandung, Nama Ibu Kandung, dan Nama Wali. Penghapusan kolom-kolom ini bertujuan untuk menyederhanakan struktur data dan hanya menyisakan fitur-fitur yang berpotensi berpengaruh terhadap hasil klasifikasi.

Setelah itu menambahkan satu kolom baru yang dinamakan "layak_pip" sebagai label klasifikasi. Kolom ini menentukan apakah seorang siswa termasuk dalam kategori "Layak" atau "Tidak Layak" menerima bantuan PIP berdasarkan beberapa kriteria. Kriteria yang digunakan yaitu: siswa yang berstatus yatim, piatu, atau yatim piatu; siswa yang memiliki kebutuhan khusus; serta siswa yang memiliki status kepemilikan KIP (Kartu Indonesia Pintar). Apabila salah satu dari kondisi tersebut terpenuhi, maka siswa tersebut akan diberi label "Layak". Sebaliknya, jika tidak memenuhi satupun kriteria, maka diberi label "Tidak Layak". Proses penentuan ini dilakukan secara otomatis menggunakan fungsi logika pada Python dan diterapkan ke seluruh baris data dengan metode *apply*().



Gambar 3. Distribusi Kelayakan

Gambar di atas menunjukkan distribusi jumlah siswa berdasarkan hasil klasifikasi kelayakan menerima Program Indonesia Pintar (PIP). Dari total data yang dianalisis, sebanyak 55 siswa diklasifikasikan sebagai "Tidak Layak" menerima bantuan, sedangkan 50 siswa diklasifikasikan sebagai "Layak". Visualisasi ini memperlihatkan bahwa jumlah siswa yang tidak layak sedikit lebih banyak dibandingkan yang layak, meskipun perbedaannya tidak terlalu signifikan.

Model Naïve Bayes Dan Melakukan Prediksi

Pada penelitian ini, algoritma yang digunakan untuk melakukan klasifikasi adalah *Categorical* Naive Bayes [3]. Pemilihan algoritma ini didasarkan pada karakteristik dataset yang seluruhnya terdiri dari data kategorikal, seperti jenis kelamin, status keluarga, kebutuhan khusus, status KIP, pekerjaan orang tua, jumlah tanggungan, dan penghasilan orang tua.

Categorical Naive Bayes merupakan varian dari *algoritma Naive Bayes* yang dirancang khusus untuk menangani fitur-fitur kategorikal secara langsung. Algoritma ini menghitung *probabilitas* masing-masing kelas berdasarkan frekuensi kemunculan nilai dari setiap fitur dalam kelas tersebut. *Teorema Bayes* digunakan sebagai dasar dalam menghitung probabilitas posterior, dengan asumsi bahwa antar fitur saling bebas (independen) [4]. Dalam konteks penelitian ini, model bertugas untuk memprediksi apakah seorang siswa layak atau tidak layak menerima bantuan Program Indonesia Pintar (PIP), berdasarkan kombinasi dari kondisi sosial dan ekonomi yang direpresentasikan oleh atribut-atribut yang telah dipilih. Untuk menjelaskan cara kerja algoritma ini secara lebih konkret, dilakukan simulasi perhitungan secara manual dengan menggunakan 10 data latih dan 2 data uji sebagai contoh. Langkah-langkah perhitungan manual meliputi:

Tabel 1. Data Uji

Umur	Status	Jenis Kelamin	Kebutuhan Khusus	Status KIP	pekerjaan orang tua	jumlah tanggungan orang tua	Penghasilan orang tua	layak_pip
15	3	0	1	1	2	2	3800000	?
14	0	0	1	0	6	4	2450000	?

Dari data latih yang diberikan, setiap atribut (seperti status, jenis kelamin, kebutuhan khusus, status KIP, pekerjaan orang tua, jumlah tanggungan orang tua, dan penghasilan orang tua) memiliki frekuensi kemunculan yang digunakan untuk menghitung probabilitas kondisional berdasarkan kelas target yaitu kelayakan menerima PIP (layak/tidak layak). Probabilitas prior dari kelas target juga dihitung berdasarkan proporsi jumlah data untuk masing-masing kelas. Tujuan utama dari data latih ini adalah untuk membangun model probabilistik yang dapat menghitung kemungkinan (probabilitas posterior) suatu data termasuk dalam kelas tertentu berdasarkan kombinasi nilai-nilai atributnya. Data uji kemudian digunakan untuk memprediksi kelas kelayakan PIP berdasarkan model yang telah dilatih tersebut.

1. Untuk data uji ke-1 dimana diketahui :
   Umur : 15

This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 25

Status : 3 (yatim piatu)
JK : 0 (laki-laki)
Kebutuhan Khusus : 1 (Kesulitan Belajar)
Status KIP : 1 (Ada)
Pekerjaan Orang Tua : 2 (nelayan)
Tanggungan : 2
Penghasilan : 3800000

a. Perhitungan nilai probabilitas prior $(P(Class))$:

$$P\ (Y) = \frac{Jumlah\ Data\ dengan\ Y}{Total\ Jumlah\ Data\ Latih}$$

Dengan menggunakan persamaan di atas, kita akan memperoleh probabilitas prior untuk 5 kelas pertama :
Total data latih : 10

1. P (Layak = 5) = $\frac{5}{10}$= 0.5
2. P (Tidak Layak = 5) = $\frac{5}{10}$ = 0.5

b. Perhitungan Nilai Probabilitas Kondisional Dengan Laplace Smoothing :

$$P\ (C|\ X) \propto P(C) * \prod_{i=1}^{n} P(x_i|C)$$

Dimana :
C : kelas (Layak atau Tidak Layak)
x_i : fitur ke-i
$P(C)$: probabilitas awal (prior)
$P(x_i|C)$: probabilitas fitur ke-i diberikan kelas C
Menggunakan laplace smoothing :

$$P(x_i|C) = \frac{jumlah\ nilai\ x\ pada\ kelas\ C+1}{jumlah\ data\ kelas\ C+jumlah\ nilai\ unik\ pada\ fitur\ tersebut}$$

Dengan menggunakan persamaan di atas, kita akan memperoleh probabilitas kondisional

1. Probabilitas Kondisional untuk kelas 1 (Layak) dimana jumlah data layak = 5
   Fitur Status :
   Nilai : 3
   Frekuensi di Layak : 2
   Nilai Unik : 4
   $$P(x_i|C) = \frac{2+1}{5+4} = \frac{3}{9} = 0.333$$
   Fitur JK :
   Nilai : 0
   Frekuensi di Layak : 2
   Nilai Unik : 2
   $$P(x_i|C) = \frac{2+1}{5+2} = \frac{3}{7} = 0.429$$
   Fitur Kebutuhan Khusus :
   Nilai : 1
   Frekuensi di Layak : 3
   Nilai Unik : 2
   $$P(x_i|C) = \frac{3+1}{5+2} = \frac{4}{7} = 0.571$$
   Fitur Status KIP:
   Nilai : 1
   Frekuensi di Layak : 3
   Nilai Unik : 2
   $$P(x_i|C) = \frac{3+1}{5+2} = \frac{4}{7} = 0.5714$$
   Fitur Pekerjaan Orang Tua :
   Nilai : 2
   Frekuensi di Layak : 1
   Nilai Unik : 11
   $$P(x_i|C) = \frac{1+1}{5+11} = \frac{2}{16} = 0.125$$
   Fitur Jumlah Tanggungan :
   Nilai : 2
   Frekuensi di Layak : 1
   Nilai Unik : 4
   $$P(x_i|C) = \frac{1+1}{5+4} = \frac{2}{9} = 0.222$$
   Fitur Penghasilan :
   Nilai : 3800000
   Frekuensi di Layak : 1
   Nilai Unik : 6
   $$P(x_i|C) = \frac{1+1}{5+6} = \frac{2}{11} = 0.182$$
2. Probabilitas Kondisional untuk kelas 0 (Tidak Layak)
   Fitur Status :
   Nilai : 3
   Frekuensi di Layak : 1
   Nilai Unik : 4
   $$P(x_i|C) = \frac{1+1}{5+4} = \frac{2}{9} = 0.222$$
   Fitur JK :
   Nilai : 0
   Frekuensi di Layak : 3
   Nilai Unik : 2
   $$P(x_i|C) = \frac{3+1}{5+2} = \frac{4}{7} = 0.571$$
   Fitur Kebutuhan Khusus :
   Nilai : 1
   Frekuensi di Layak : 1
   Nilai Unik : 2
   $$P(x_i|C) = \frac{1+1}{5+2} = \frac{2}{7} = 0.286$$
   Fitur Status KIP:
   Nilai : 1
   Frekuensi di Layak : 1
   Nilai Unik : 2
   $$P(x_i|C) = \frac{1+1}{5+2} = \frac{2}{7} = 0.286$$
   Fitur Pekerjaan Orang Tua :
   Nilai : 2
   Frekuensi di Layak : 0
   Nilai Unik : 11
   $$P(x_i|C) = \frac{0+1}{5+11} = \frac{1}{16} = 0.0625$$
   Fitur Jumlah Tanggungan :
   Nilai : 2
   Frekuensi di Layak : 1
   Nilai Unik : 4
   $$P(x_i|C) = \frac{1+1}{5+4} = \frac{2}{9} = 0.222$$
   Fitur Penghasilan :
   Nilai : 3800000

This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 26

Frekuensi di Layak : 1
Nilai Unik : 6

$$P(x_i|C) = \frac{1+1}{5+6} = \frac{2}{11} = 0.182$$

Selanjutnya adalah menghitung probabilitas posterior dari data tersebut.

c. Perhitungan nilai probabilitas posterior $(P(Class|Data)$:
Untuk kelas **Layak** :

$$P\ (\ Layak|\ X) = 0.5 \times 0.333 \times 0.429 \times 0.571 \times 0.571 \times 0.125 \times 0.222 \times 0.182$$
$$= 0.5 \times 0.000225$$
$$= 0.0001125$$

Untuk kelas **Tidak Layak**:

$$P\ (\ Tidak\ Layak|\ X) = 0.5 \times 0.222 \times 0.571 \times 0.286 \times 0.286 \times 0.0625 \times 0.222 \times 0.182$$
$$= 0.5 \times 0.000111$$
$$= 0.0000555$$

Prediksi model menyimpulkan dari hasil perhitungan, kelas dengan probabilitas terbesar adalah **Layak**. Jadi, prediksi untuk data uji pertama adalah **layak** mendapatkan PIP. Selanjutnya proses perhitungan untuk data uji berikutnya menggunakan metode yang sama.

2. Untuk data uji ke-2 dimana diketahui :
Umur : 14
Status : 0 (lengkap)
JK : 0 (laki-laki)
Kebutuhan Khusus : 1 (Kesulitan Belajar)
Status KIP : 0 (Tidak Ada)
Pekerjaan Orang Tua : 6 (guru honorer)
Tanggungan : 4
Penghasilan : 2450000

a. Perhitungan nilai probabilitas prior $(P(Class))$:

$$P\ (Y) = \frac{Jumlah\ Data\ dengan\ Y}{Total\ Jumlah\ Data\ Latih}$$

Dengan menggunakan persamaan di atas, kita akan memperoleh probabilitas prior untuk 5 kelas pertama :

Total data latih : 10

3. P (Layak = 5) = $\frac{5}{10}$= 0.5
4. P (Tidak Layak = 5) = $\frac{5}{10}$ = 0.5

b. Perhitungan Nilai Probabilitas Kondisional Dengan Laplace Smoothing :

Probabilitas Kondisional untuk kelas 1 (Layak) dimana jumlah data layak = 5

Fitur Status :
Nilai : 0
Frekuensi di Layak : 4
Nilai Unik : 2

$$P(x_i|C) = \frac{4+1}{5+2} = \frac{5}{7} = 0.714$$

Fitur JK :
Nilai : 0
Frekuensi di Layak : 3
Nilai Unik : 2

$$P(x_i|C) = \frac{3+1}{5+2} = \frac{4}{7} = 0.571$$

Fitur Kebutuhan Khusus :
Nilai : 1
Frekuensi di Layak : 3
Nilai Unik : 2

$$P(x_i|C) = \frac{3+1}{5+2} = \frac{4}{7} = 0.571$$

Fitur Status KIP:
Nilai : 0
Frekuensi di Layak : 1
Nilai Unik : 2

$$P(x_i|C) = \frac{1+1}{5+2} = \frac{2}{7} = 0.286$$

Fitur Pekerjaan Orang Tua :
Nilai : 6
Frekuensi di Layak : 0
Nilai Unik : 6

$$P(x_i|C) = \frac{0+1}{5+6} = \frac{1}{11} = 0.091$$

Fitur Jumlah Tanggungan :
Nilai : 4
Frekuensi di Layak : 1
Nilai Unik : 4

$$P(x_i|C) = \frac{1+1}{5+4} = \frac{2}{9} = 0.222$$

Fitur Penghasilan :
Nilai : 2450000
Frekuensi di Layak : 0
Nilai Unik : 5

$$P(x_i|C) = \frac{0+1}{5+5} = \frac{1}{10} = 0.1$$

3. Probabilitas Kondisional untuk kelas 0 (Tidak Layak)

Fitur Status :
Nilai : 0
Frekuensi di Layak : 4
Nilai Unik : 2

$$P(x_i|C) = \frac{4+1}{5+2} = \frac{5}{7} = 0.714$$

Fitur JK :
Nilai : 0
Frekuensi di Layak : 2
Nilai Unik : 2

$$P(x_i|C) = \frac{2+1}{5+2} = \frac{3}{7} = 0.429$$

Fitur Kebutuhan Khusus :
Nilai : 1
Frekuensi di Layak : 1
Nilai Unik : 2

$$P(x_i|C) = \frac{1+1}{5+2} = \frac{2}{7} = 0.286$$

Fitur Status KIP:
Nilai : 0

This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 27

Frekuensi di Layak : 4
Nilai Unik : 2

$$P(x_i|C) = \frac{4+1}{5+2} = \frac{5}{7} = 0.714$$

Fitur Pekerjaan Orang Tua :
Nilai : 6
Frekuensi di Layak : 2
Nilai Unik : 6

$$P(x_i|C) = \frac{2+1}{5+6} = \frac{3}{11} = 0.273$$

Fitur Jumlah Tanggungan :
Nilai : 4
Frekuensi di Layak : 2
Nilai Unik : 4

$$P(x_i|C) = \frac{2+1}{5+4} = \frac{3}{9} = 0.333$$

Fitur Penghasilan :
Nilai : 2450000
Frekuensi di Layak : 1
Nilai Unik : 5

$$P(x_i|C) = \frac{1+1}{5+5} = \frac{2}{10} = 0.2$$

Selanjutnya adalah menghitung probabilitas posterior dari data tersebut.

c. Perhitungan nilai probabilitas posterior $(P(Class|Data)$:

Untuk kelas **Layak** :

$$P\ (\ Layak|\ X) = 0.5 \times 0.714 \times 0.714 \times 0.571 \times 0.571 \times 0.286 \times 0.091 \times 0.222 \times 0.1$$
$$= 0.5 \times 0.0000958$$
$$= 0.0000479$$

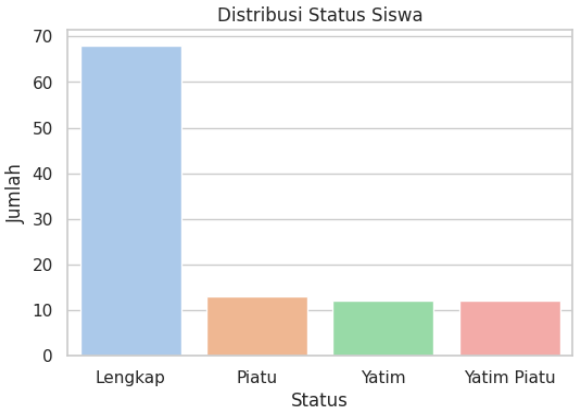
Untuk kelas **Tidak Layak**:

$$P\ (\ Tidak\ Layak|\ X) = 0.5 \times 0.714 \times 0.714 \times 0.429 \times 0.286 \times 0.714 \times 0.273 \times 0.333 \times 0.2$$
$$= 0.5 \times 0.000812$$
$$= 0.000406$$

Berdasarkan hasil perhitungan manual menggunakan algoritma *Categorical Naïve Bayes*, setiap data uji dianalisis dengan menghitung probabilitas posterior untuk masing-masing kelas target, yaitu Layak dan Tidak Layak. Dalam proses ini digunakan probabilitas prior dari masing-masing kelas serta probabilitas kondisional dari fitur-fitur seperti Jenis Kelamin, Status, Kebutuhan Khusus, KIP, Pekerjaan Orang Tua, Jumlah Tanggungan Orang Tua, dan Penghasilan Orang Tua. Karena nilai probabilitas posterior tertinggi terdapat pada kelas Tidak Layak, maka prediksi untuk data uji ke-2 adalah Tidak Layak[5]. Prediksi ini menunjukkan bahwa pola data latih mendukung kecenderungan fitur-fitur yang dimiliki oleh data uji ke-2 lebih banyak muncul pada kategori Tidak Layak, sehingga model memberikan probabilitas tertinggi pada kelas tersebut [6]. Hal ini konsisten dengan distribusi fitur-fitur yang lebih sering muncul pada kelas Tidak Layak dalam data latih dibandingkan kelas Layak.

Visualisasi Hasil dan Hasil Evaluasi Menggunakan Confusion Matriks

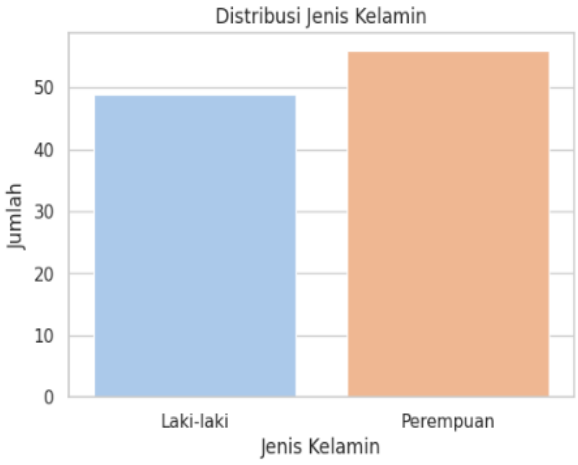
Distribusi Status Siswa



Gambar 4. Kategori Status Siswa

Gambar tersebut menampilkan distribusi status siswa berdasarkan kondisi keberadaan orang tua, baik dalam bentuk grafik batang maupun output data. ada output awal, status siswa masih ditampilkan dalam bentuk angka: 0, 1, 2, dan 3, yang masing-masing merepresentasikan kategori status sebagai berikut: 0 = Lengkap (memiliki ayah dan ibu), 1 = Piatu (kehilangan ibu), 2 = Yatim (kehilangan ayah), dan 3 = Yatim Piatu (kehilangan kedua orang tua). Berdasarkan data, mayoritas siswa berada dalam kategori "Lengkap" (memiliki kedua orang tua), sebanyak 68 siswa. Sementara itu, kategori "Piatu" (kehilangan ibu) berjumlah 13 siswa, "Yatim" (kehilangan ayah) sebanyak 12 siswa, dan "Yatim Piatu" (kehilangan kedua orang tua) juga berjumlah 12 siswa.

Distribusi Jenis Kelamin

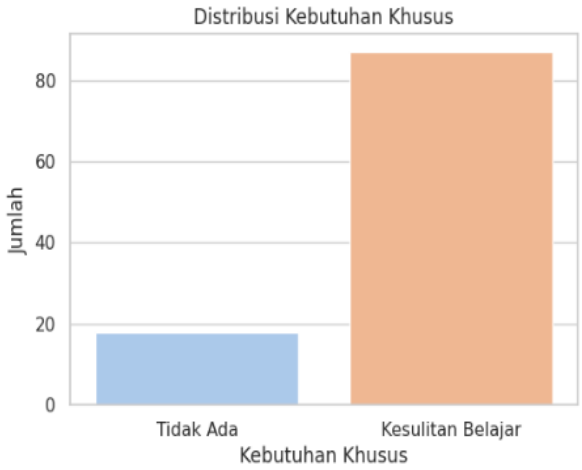


Gambar 5. Distribusi Berdasarkan Jenis Kelamin

This work is licensed under a Creative Commons Attribution 4.0 International License
DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 28

Gambar tersebut menampilkan distribusi jenis kelamin siswa dalam bentuk grafik batang dan output data numerik. Pada output awal, jenis kelamin siswa masih direpresentasikan dalam bentuk angka: 0 dan 1, di mana 0 = Laki-laki dan 1 = Perempuan. Berdasarkan data, jumlah siswa perempuan sedikit lebih banyak, yaitu 56 orang, dibandingkan dengan siswa laki-laki yang berjumlah 49 orang.

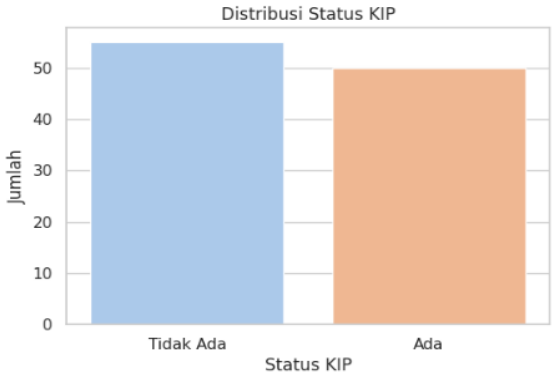
Distribusi Kebutuhan Khusus



Gambar 6. Distribusi Kebutuhan Khusus

Gambar tersebut menampilkan distribusi kebutuhan khusus siswa dalam bentuk grafik batang dan data numerik. Pada output awal, kategori kebutuhan khusus masih ditampilkan dalam bentuk angka: 0 dan 1, di mana 0 = Tidak Ada (tidak memiliki kebutuhan khusus) dan 1 = Kesulitan Belajar. Berdasarkan data, mayoritas siswa, yaitu 87 orang, termasuk dalam kategori Kesulitan Belajar, sedangkan 18 siswa tidak memiliki kebutuhan khusus.

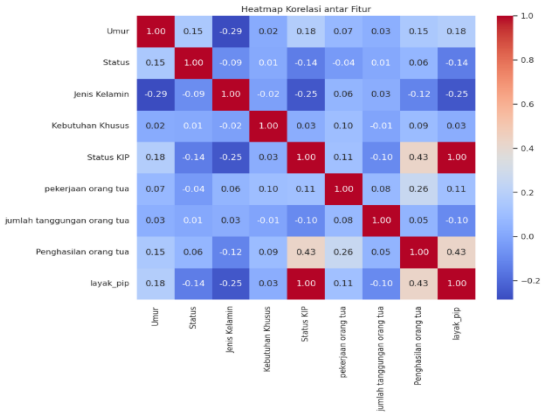
Distribusi Status KIP



Gambar 7. Distribusi Status KIP

Gambar tersebut menampilkan distribusi status Kartu Indonesia Pintar (KIP) siswa dalam bentuk grafik batang dan data numerik. Pada output awal, status KIP direpresentasikan dengan angka 0 dan 1, di mana 0 = Tidak Ada (tidak menerima KIP) dan 1 = Ada (menerima KIP). Berdasarkan data, terdapat 55 siswa yang tidak memiliki KIP, sedangkan 50 siswa tercatat menerima KIP. Informasi ini dapat membantu dalam analisis distribusi bantuan pendidikan di kalangan siswa.

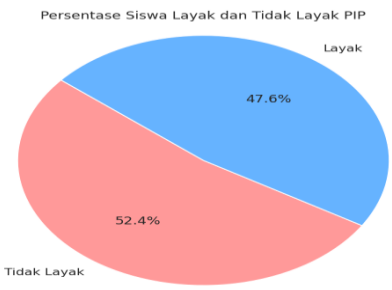
Heatmap Korelasi Antar Fitur



Gambar 8. Heatmap Korelasi Antar Fitur

Gambar tersebut menunjukkan heatmap korelasi antar fitur dalam dataset yang merepresentasikan hubungan linear antara masing-masing variabel. Nilai korelasi berkisar dari -1 hingga 1, di mana nilai positif menunjukkan hubungan searah, nilai negatif menunjukkan hubungan berlawanan arah, dan nilai mendekati nol berarti tidak ada hubungan linear yang kuat. Terlihat bahwa fitur Status KIP memiliki korelasi yang sangat tinggi dengan layak_pip (nilai 1.00), yang mengindikasikan bahwa keberadaan KIP sangat menentukan kelayakan menerima bantuan PIP. Selain itu, Penghasilan orang tua juga memiliki korelasi positif sedang dengan layak_pip (0.43), menandakan bahwa semakin rendah penghasilan, semakin besar kemungkinan siswa memenuhi syarat PIP. Beberapa fitur lain seperti Jenis Kelamin, Kebutuhan Khusus, dan Jumlah Tanggungan Orang Tua menunjukkan korelasi yang lemah dengan variabel lainnya, menandakan bahwa hubungan antar fitur-fitur tersebut cenderung lemah atau tidak signifikan secara linear[7].

Persentase Siswa Layak dan Tidak Layak



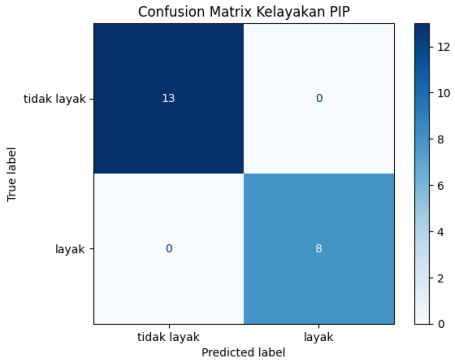
Gambar 9. Persentase Layak dan Tidak Layak

This work is licensed under a
Creative Commons Attribution 4.0 International License

DOI http://dx.doi.org/10.36448/expert.v15i1.4326
e-ISSN 2745-7265 p-ISSN 2088-5555 EXPERT Vol. 12 No. 1
June 5, 2025 – Hal. 29

Gambar *pie chart* tersebut menggambarkan persentase siswa yang layak dan tidak layak menerima bantuan Program Indonesia Pintar (PIP). Berdasarkan visualisasi, terlihat bahwa proporsi siswa yang tidak layak PIP sedikit lebih besar, yaitu sebesar 52.4%, dibandingkan dengan siswa yang layak menerima PIP yang berjumlah 47.6%. Perbedaan ini menunjukkan bahwa hampir setengah dari total siswa memenuhi kriteria untuk menerima bantuan, sedangkan sisanya belum memenuhi atau tidak memenuhi kriteria yang telah ditentukan.

Confusion Matriks



Gambar 10. Confusion Matrix

Pada gambar 10 hasil pengklasifikasian untuk mencari *accuration*, *precision*, recall dan F1-*score* mengikuti rumus berikut, misalnya pada elemen "Klinik Jiwa Dewasa"

Tabel 2. Penentuan Nilai TP, FP, TN dan FP

Label	Nilai
TP (*True Positive*)	8
FP (*False Positive*)	0
FN (*False Negative*)	0
TN (*True Negative*)	13

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{8 + 13}{8 + 13 + 0 + 0} = \frac{21}{21} = 1.0$$

$$precision = \frac{TP}{TP + FP} = \frac{8}{8 + 0} = \frac{8}{8} = 1.0$$

$$recall = \frac{TP}{TP + FN} = \frac{8}{8 + 0} = \frac{8}{8} = 1.0$$

$$F1 - Score = 2\,x\,\frac{Precision\ x\ Recall}{Precision + Recall} = 2\,x\,\frac{1 * 1}{1 + 1} = 2 * \frac{1}{2} = 1.0$$

Tabel 3. Classification Report

Classification Report				
	Precision	Recall	F1-Score	Support
Tidak Layak	1.00	1.00	1.00	13
Layak	1.00	1.00	1.00	8
Accuracy			1.00	21
Macro avg	1.00	1.00	1.00	21
Weighted avg	1.00	1.00	1.00	21

Berdasarkan hasil classification report pada gambar, model klasifikasi berhasil mengidentifikasi data penerima Program Indonesia Pintar (PIP) dengan sangat baik. Seluruh metrik evaluasi, termasuk precision, recall, dan f1-score untuk kedua kelas ("Layak" dan "Tidak Layak"), menunjukkan nilai sempurna yaitu 1.00. Ini berarti model mampu mengklasifikasikan semua data uji (sebanyak 21 data) secara akurat tanpa kesalahan. Nilai accuracy, macro average, dan weighted average yang juga sebesar 1.00 semakin menguatkan bahwa model memiliki kinerja prediksi yang sangat tinggi pada dataset ini. Namun demikian, tingkat akurasi yang sempurna ini juga perlu diwaspadai, karena dapat mengindikasikan kemungkinan terjadinya overfitting—terutama mengingat ukuran data uji yang relatif kecil. Oleh karena itu, diperlukan evaluasi lanjutan seperti cross-validation untuk memastikan generalisasi model terhadap data baru.

4. KESIMPULAN

Penelitian ini menggunakan algoritma Naïve Bayes untuk mengklasifikasikan calon penerima Program Indonesia Pintar (PIP) berdasarkan beberapa atribut kategori yang diperoleh dari MTs Al-Hasanah Tj. Leidong, Kecamatan Kualuh Leidong, Kabupaten Labuhan Batu Utara, Sumatera Utara, seperti pekerjaan orang tua, jumlah tanggungan, dan penghasilan orang tua. Data diperoleh dari 105 siswa, kemudian dilakukan tahap preprocessing data dengan mengubah nilai kategorikal menjadi numerik menggunakan metode Label Encoding. Proses ini bertujuan agar data dapat diterima oleh algoritma Categorical Naïve Bayes secara optimal. Selanjutnya, data dibagi menjadi data latih dan data uji dengan rasio 80:20, menghasilkan 84 data latih dan 21

This work is licensed under a Creative Commons Attribution 4.0 International License

data uji. Evaluasi performa model dilakukan menggunakan confusion matrix dan empat metrik evaluasi, yaitu akurasi, presisi, recall, dan F1-score.Hasil pengujian menunjukkan bahwa model memperoleh nilai akurasi sebesar 1.0 atau 100%, precision sebesar 1.0 atau 100%, recall sebesar 1.0 atau 100%, dan F1-score sebesar 1.0 atau 100%. Nilai-nilai tersebut menunjukkan bahwa model Naïve Bayes mampu mengklasifikasikan kelayakan siswa sebagai penerima bantuan PIP dengan sangat baik. Namun demikian, hasil yang sempurna ini juga perlu dicermati secara kritis. Ukuran dataset yang relatif kecil dan pembagian data yang hanya satu kali (hold-out) dapat meningkatkan risiko overfitting, di mana model terlalu sesuai dengan data uji saat ini namun kurang mampu menggeneralisasi pada data baru. Oleh karena itu, penelitian ini memiliki potensi untuk dikembangkan lebih lanjut dengan menambahkan data yang lebih besar dan beragam, serta menerapkan teknik validasi silang seperti k-fold cross-validation untuk memperoleh estimasi performa yang lebih andal.

DAFTAR PUSTAKA

[1] R. A. Tyas, M. Anggraini, I. A. Sulasiyah, and Q. Aini, "Implementasi Algoritma Naïve Bayes Dalam Penentuan Rating Buku," *Sistemasi*, vol. 9, no. 3, p. 557, 2020, doi: 10.32520/stmsi.v9i3.915.

[2] E. A. Lisangan, A. Gormantara, and R. Y. Carolus, "Implementasi Naive Bayes pada Analisis Sentimen Opini Masyarakat di Twitter Terhadap Kondisi New Normal di Indonesia," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 23–32, 2022, doi: 10.24002/konstelasi.v2i1.5609.

[3] Soebagio, *Manajemen Pendidikan Indonesia*. Jakarta: Ardadizya Jaya, 2005.

[4] K. P. dan K. Indonesia, *Peraturan Menteri Pendidikan dan Kebudayaan Nomor 12 Tahun 2015 tentang Program Indonesia Pintar*. 2015, p. BN.2015/No.724, jdih.kemdikbud : 6 hlm.

[5] W. P. Nurmayanti, "Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak.," *Geodika J. Kaji. Ilmu Dan Pendidik. Geogr.*, vol. 5, no. 1, pp. 123–132, 2021, doi: https://doi.org/10.29408/geodika.v5i1.3430.

[6] A. Wibisono, A. D., Dadi Rizkiono, S., & Wantoro, "Filtering Spam Email Menggunakan Metode Naive Bayes.," *Telefortech J. Telemat. Inf. Technol.*, vol. 1, no. 1, pp. 9–17, 2020, doi: https://doi.org/10.33365/tft.v1i1.685.

[7] D. Nofriansyah, *Konsep Data Mining Vs Sistem Pendukung Keputusan*. Yogyakarta: Deepublish, 2015.

[8] S. Felicia Watratan, A., Puspita, A. B., Moeis, D., Informasi, S., & Profesional Makassar, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia.," *J. 47 Appl. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 12–22, 2020.

[9] A. Saleh, "Implementasi Metode Klasifikasi Naive Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga.," *Citec J.*, vol. 2, no. 3, pp. 207–217, 2015.

[10] S. Hendrian., "ALGORITMA KLASIFIKASI DATA MINING UNTUK MEMPREDIKSI SISWA DALAM MEMPEROLEH BANTUAN DANA PENDIDIKAN," *Fakt. Exacta*, vol. 11, no. 3, pp. 266–274, 2019, doi: 10.30998/faktorexacta.v11i3.2777.

[11] A. Nuraini, A. Faqih, G. Dwilestari, N. Dienwati Nuris, and R. Narasati, "Analisis Sentimen Terhadap Review Aplikasi Brimo Di Google Play Store Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3661–3666, 2023, doi: 10.36040/jati.v7i6.8228.

[12] F. Santoso, H., Armansyah, A., & Siregar, "Implementation of Naïve Bayes Method in Classification of Nutritional Status of Toddlers at Pasar Ujungbatu Sosa Public Health Center.," *IJISTECH (International J. Inf. Syst. Technol.*, vol. 6, no. 3, pp. 392–398, 2022, doi: https://doi.org/10.30645/ijistech.v6i3.254.

[13] Frira Sesilia, Viktor Handrianus Pranatawijaya, and Ressa Priskila, "Machine Learning untuk Memprediksi Jumlah Penjualan, Stok dan Jumlah Tanam Hasil Pertanian Hidroponik," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 222–233, 2024, doi: 10.24002/konstelasi.v4i1.9055.

[14] W. Lestari, R., D, Muhammad, S.,H, Sri, "Penerapan Data Mining Menggunakan Metode Teknik Klasifikasi Untuk Melihat Potensi Kepatuhan Wajib Pajak Kendaraan.," *J. Comput. Sci. Informatics Eng.*, vol. 3, no. 1, pp. 1–14, 2024, doi: 10.55537/cosie.v3i1.710.

[15] F. Marisa, "Educational Data Mining (Konsep dan Penerapan)," *Teknol. Inf.*, vol. 4, no. 2, pp. 90–97, 2018.

[16] R. Armansyah, Rakhmad, K., "Model Prediksi Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes.," *Edumatic J. Pendidik. Inform.*, vol. 6, no. 1, pp. 1–10, 2022, doi: https://doi.org/10.29408/edumatic.v6i1.4789.

This work is licensed under a Creative Commons Attribution 4.0 International License