IMPLEMENTASI ALGORITMA *NAIVE BAYES* DENGAN METODE KLASIFIKASI DALAM MENENTUKAN SISWA PENERIMA BANTUAN PROGRAM INDONESIA PINTAR (STUDI KASUS : SMPN 3 CIHAMPELAS)

Irma Purnama Oktavia [1], Novita Lestari Anggreini [2]
[1] Teknik Informatika, Politeknik TEDC Bandung
[2] Teknik Komputer, Politeknik TEDC Bandung
Jl.Pasantren KM 2 Cimahi
irmapurnamaoktavia@gmail.com

ABSTRAK

Pendidikan memiliki peran penting dalam meningkatkan kualitas sumber daya manusia (SDM) di Indonesia. Namun, faktor ekonomi sering kali menjadi hambatan dalam mendapatkan pendidikan yang layak. Pemerintah telah mengatasi hal ini dengan Program Indonesia Pintar (PIP), sebuah beasiswa untuk siswa kurang mampu. SMPN 3 Cihampelas merupakan salah satu penerima program ini, namun penentuan siswa penerima PIP masih menemui kendala, terutama dalam memastikan bantuan tepat sasaran. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Naive Bayes* dalam mengklasifikasikan kelayakan siswa penerima bantuan PIP berdasarkan beberapa atribut seperti Jenis Tinggal, Alat Transportasi, Pekerjaan dan Penghasilan Orang Tua, serta penerima KPS dan KIP. Pengujian dilakukan menggunakan perangkat lunak *RapidMiner* dan teknik *Cross Validation*. Hasil penelitian menunjukkan tingkat akurasi sebesar 97.24%, *precision* 72.73%, *recall* 88.89%, dan nilai *AUC* 0.886, yang dikategorikan sebagai *Good Classification*. Sehingga dapat disimpulkan, bahwa algoritma ini dapat diimplementasikan dengan sangat baik untuk mendukung pengambilan keputusan terkait penerima bantuan PIP di SMPN 3 Cihampelas.

Kata kunci : PIP, *Naive Bayes*, *RapidMiner*

1. PENDAHULUAN

Pendidikan memiliki peranan yang sangat penting dalam meningkatkan kualitas sumber daya manusia (SDM). Melalui pendidikan seseorang tidak hanya mendapatkan pengajaran ilmu pengetahuan tetapi juga sesuatu yang lebih mendalam yaitu pengembangan sikap, keterampilan serta kecerdasan intelektual. Setiap warga negara Indonesia berhak untuk mendapatkan pendidikan sebagaimana tertuang dalam UUD 1945 pasal 31 ayat 1-2, namun pada pelaksanaannya tidak semua warga negara Indonesia mendapatkan kesempatan yang sama dalam memperoleh pendidikan yang layak[1].

Salah satu faktor penghambat dalam terwujudnya pendidikan yang layak adalah faktor ekonomi. Pemerintah telah mengambil peranan penting dalam menyikapi permasalahan tersebut dengan mengeluarkan Program Indonesia Pintar (PIP). Tujuan program ini adalah untuk meningkatkan mutu pendidikan bagi anak usia sekolah (6-21 tahun), mencegah dan meminimalisir siswa dari kemungkinan putus sekolah serta memenuhi kebutuhan siswa dalam pembelajaran. Program juga ditujukan agar dapat dimanfaatkan siswa dalam memenuhi kebutuhan sekolah seperti biaya perlengkapan sekolah, biaya transportasi siswa dan uang saku [2].

Namun dalam implementasinya, proses pengambilan keputusan terkait penentuan siswa penerima bantuan PIP di SMPN 3 Cihampelas masih mengalami kendala, antara lain masih kurang tepatnya sasaran dalam pemberian bantuan kepada siswa yang berhak menerima karena proses penilaiannya melibatkan kriteria atau atribut yang saling berkaitan serta tidak

hanya diputuskan berdasarkan kebijakan dari pembuat keputusan. Hal ini juga dikarenakan belum diimplementasikannya sebuah metode yang dapat menentukan kelayakan siswa penerima bantuan PIP berdasarkan kriteria tersebut.

Oleh karena itu, penelitian ini menggunakan algoritma *Naive Bayes* dalam mengklasifikasikan kelayakan siswa penerima bantuan Program Indonesia Pintar (PIP) untuk memperoleh hasil penerimaan bantuan yang lebih akurat. Algoritma *Naive Bayes* merupakan sebuah pengklasifikasian dengan metode probabilitas dan statistik yang dapat digunakan untuk memprediksi peluang pada setiap anggota kelas tertentu dan menghitung peluang yang akan terjadi pada atribut yang ada. Algoritma ini dikenal dengan *teorema Bayes* yaitu dapat memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya [3]. *Naive Bayes* juga dapat mencapai nilai akurasi yang tinggi dengan hanya membutuhkan sedikit data latih [4].

2. TINJAUAN PUSTAKA

2.1. Program Indonesia Pintar

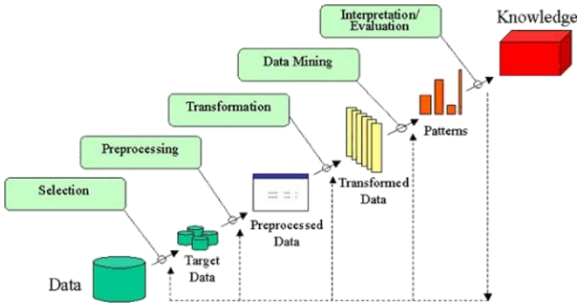
Program Indonesia Pintar (PIP) merupakan program beasiswa yang dikeluarkan oleh pemerintah untuk membantu anak usia sekolah yang berasal dari keluarga kurang mampu agar tetap mendapatkan layanan pendidikan sampai tamat pendidikan

menengah/sederajat. Program ini dikeluarkan dalam wujud bantuan tunai pendidikan sebagai bagian dari penyempurnaan Program Bantuan Siswa Miskin (BSM) yang ditandai dengan adanya Kartu Indonesia Pintar (KIP)[5].

2.2. *Data Mining*

Data Mining merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan atau informasi penting yang selama ini tidak diketahui dalam suatu basis data (*Knowledge Discovery in Database*). *Data mining* melibatkan penerapan teknik statistik, matematika, *machine learning* serta kecerdasan buatan dalam mengekstraksi dan mengidentifikasi informasi dan pengetahuan yang terkait dari berbagai *database* besar. *Data mining* merupakan kegiatan untuk mencari pola yang menarik dari berbagai data dalam jumlah besar dimana data tersebut disimpan dalam *database*, data *warehouse* atau penyimpanan informasi lainnya [6].

2.3. Tahapan *Data Mining*



Gambar 1. Tahapan *Data Mining*

Berikut ini merupakan penjelasan dari gambar 1 :

a. *Data Selection*
   Sekumpulan data operasional perlu dilakukan pemilihan (seleksi) sebelum dimulainya tahap penggalian informasi pada KDD (*Knowledge Discovery in Databases*). Data hasil seleksi disimpan dalam satu halaman terpisah dan akan digunakan untuk proses *data mining*.
b. *Pre-processing*
   Pre-processing merupakan tahap pembersihan data yang dilakukan untuk memperbaiki kualitas data. Tahap ini mencakup pembuangan data yang duplikasi, pemeriksaan data yang tidak konsisten, dan memperbaiki kesalahan pada data.
c. *Transformation*
   Transformasi data yaitu merubah atau menggabungkan data ke dalam format yang sesuai untuk proses *data mining*. Transformasi sangat bergantung pada jenis data atau model yang diambil dari set data.
d. *Data Mining*
   Data mining merupakan tahap dimana diterapkannya suatu teknik atau metode untuk mencari pola atau informasi menarik dari dalam data terpilih. Teknik, metode, atau algoritma yang dapat digunakan bervariasi dan pemilihan metode atau algoritma yang tepat bergantung pada tujuan dan proses KDD secara keseluruhan.
e. *Interpretation/Evaluation*
   Pola informasi yang dihasilkan pada proses *data mining* perlu dievaluasi dengan tujuan untuk menguji pola yang ditemukan dengan hipotesis yang ada sebelumnya. Setelah teruji pola tersebut perlu ditampilkan ke dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan seperti informasi ditampilkan dalam bentuk grafik, pohon keputusan ataupun dalam bentuk rule.
f. *Knowledge*
   Tujuan utama dari Proses Penambangan Data (KDD) adalah untuk mendapatkan pengetahuan yang bermanfaat dan dapat dipahami dengan mudah. Pengetahuan yang dihasilkan diimplementasikan sesuai dengan manfaat atau kegunaan dari informasi tersebut.

2.4. Dataset

Dataset merupakan kumpulan data yang biasanya direpresentasikan dalam bentuk satu tabel *database* atau matriks data di mana setiap kolom mewakili variabel tertentu, sementara setiap baris menunjukkan entri data yang berkaitan dengan variabel tersebut [7].

2.5. Algoritma *Naive Bayes Classifier*

Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan teknik prediksi berbasis probabilitas sederhana yang didasarkan pada penerapan Teorema Bayes dengan mengasumsikan independensi yang kuat [8].

2.6. Klasifikasi

Klasifikasi merupakan teknik untuk menemukan model atau memprediksi kelas dari suatu objek yang labelnya belum diketahui. Untuk mencapai hal tersebut, proses klasifikasi mengembangkan suatu model yang mampu memisahkan data ke dalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu [9].

2.7. *RapidMiner*

RapidMiner merupakan *software* atau perangkat lunak *open source* yang digunakan untuk pengolahan data dalam data mining. *RapidMiner* mengekstrak pola-pola dari dataset yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan *database*. *RapidMiner* merupakan sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi[10].

RapidMiner mulai dikembangkan sejak tahun 2001, sebelumnya disebut dengan nama YALE (*Yet Another Learning Environment*) oleh Unit Artifical Intelligence dari Technical University of Dortmund. *Software* ini digunakan untuk mengukur kinerja algoritma serta mencari algoritma terbaik yang dapat digunakan untuk klasifikasi, prediksi, dan teknik

lainya dalam *data mining*. Dengan fitur-fitur yang ramah pengguna dan GUI (*Graphic User Interface*) yang efektif, *RapidMine*r menjadi pilihan yang umum digunakan di kalangan praktisi dan peneliti dalam bidang ilmu data dan analisis.
Terdapat sifat-sifat yang dimiliki oleh *RapidMiner* sebagai berikut :

a. Penulisan menggunakan bahasa Java memungkinkan *RapidMiner* untuk beroperasi pada berbagai sistem operasi yang berbeda.
b. Proses penemuan pengetahuan diwujudkan dalam bentuk model operator trees.
c. Representasi XML internal digunakan untuk memfasilitasi pertukaran data dalam format standar.
d. Pemanfaatan bahasa skrip memungkinkan untuk melakukan eksperimen dalam skala besar dan otomatisasi eksperimen.
e. Konsep multi-layer memastikan efisiensi tampilan data dan penanganan data.
f. Dilengkapi dengan GUI (*Graphic User Interface*), *command line mode*, dan Java API yang dapat dipanggil oleh program lain

3. METODE PENELITIAN

3.1. Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini yaitu wawancara, observasi dan studi kepustakaan.

a. Wawancara
   Wawancara merupakan teknik pengumpulan data yang digunakan untuk mendapatkan informasi secara langsung dari responden atau subjek penelitian. Melakukan wawancara dengan staf bendahara terkait sistem beasiswa PIP di SMPN 3 Cihampelas, meliputi proses penentuan kelayakan siswa penerima bantuan, kriteria yang digunakan, hingga proses penyaluran beasiswa.
b. Observasi
   Merupakan teknik pengumpulan data dengan cara melakukan pengamatan langsung terhadap kejadian atau fenomena di lapangan. Melakukan observasi di SMPN 3 Cihampelas dan kegiatan yang dilakukan adalah mengambil data siswa penerima beasiswa PIP.
c. Studi Kepustakaan
   Studi kepustakaan merupakan serangkaian kegiatan yang berkaitan dengan metode pengumpulan data pustaka, membaca dan mencatat serta mengolah bahan penelitian. Melakukan studi kepustakaan mengenai teori–teori beasiswa PIP, data mining, algoritma *Naive Bayes*, klasifikasi, dan *RapidMiner*.

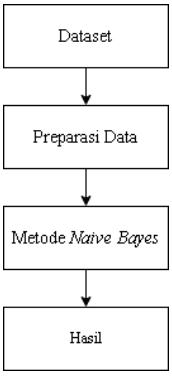
3.2. Sumber Data

Sumber data yang digunakan penulis berasal dari data primer dan data sekunder.

a. Data primer merupakan data yang dikumpulkan secara langsung dari sumber terkait, diperoleh dari Staf Bendahara SMPN 3 Cihampelas yaitu 180 data siswa penerima bantuan Program Indonesia Pintar (PIP) pada tahun 2022/2023.
b. Data sekunder merupakan data yang diperoleh secara tidak langsung atau data yang telah dikumpulkan dan diproses sebelumnya oleh pihak lain atau untuk tujuan lain, diperoleh dari beberapa sumber data seperti jurnal, artikel, *e-book* dan *website* yang berkaitan dengan SMPN 3 Cihampelas.

3.3. Metode yang digunakan

Pada penelitian ini menggunakan metode klasifikasi algoritma *Naive Bayes Classifier*.



Gambar 2. Metode *Naive Bayes*

Pada gambar 2 merupakan alur proses pengolahan data hingga menemukan hasil, berikut penjelasannya :

a. Dataset
   Tahap awal dimulai dengan pengumpulan dataset yang akan digunakan dalam peneltian. Dataset ini berisi data siswa penerima bantuan PIP yang akan dianalisis.
b. Preparasi Data
   Preparasi data dilakukan setelah diperoleh data secara lengkap, preparasi merupakan proses pembersihan data (*data cleansing*) dimana informasi data yang tidak lengkap, tidak diperlukan serta informasi data yang kurang jelas akan dibuang. Tahap ini juga meliputi transformasi data untuk memastikan data berada dalam format yang sesuai untuk analisis.
c. Metode *Naive Bayes*
   Pada tahap ini, algoritma *Naive Bayes Classifier* diterapkan untuk mengklasifikasikan data siswa penerima bantuan PIP. Proses pengujian dilakukan dengan teknik *Cross Validation* untuk memastikan keakuratan model dalam mengklasifikasi data.
d. Hasil
   Tahap akhir adalah evaluasi dari model yang telah dibangun yang menghasilkan sebuah klasifikasi dari dataset siswa penerima bantuan PIP. Kesimpulan akan ditarik berdasarkan hasil klasifikasi tersebut.

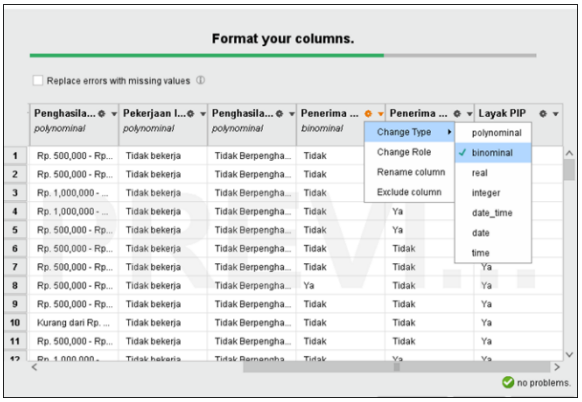
4. HASIL DAN PEMBAHASAN

4.1. Dataset

Penelitian ini menggunakan dataset yang terdiri dari 180 data siswa penerima Program Indonesia Pintar (PIP) pada tahun ajaran 2022/2023 di SMPN 3 Cihampelas. Dataset ini merupakan kumpulan data yang akan diolah lebih lanjut sebelum digunakan dalam proses analisis. Dataset akan dipreparasi dan diubah menjadi format yang sesuai untuk digunakan dalam model klasifikasi *Naive Bayes*.

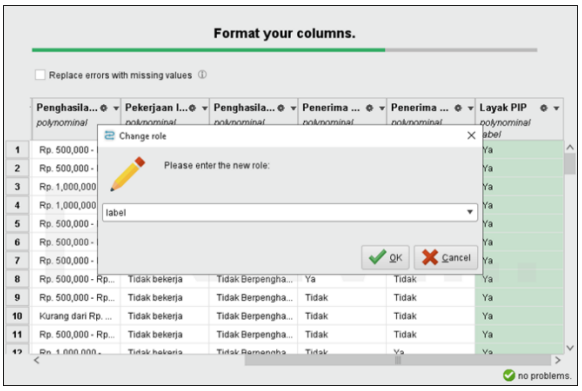
4.2. Preparasi Data

Atribut yang digunakan untuk klasifikasi kelayakan siswa penerima bantuan PIP yaitu Jenis Tinggal, Alat Transportasi, Pekerjaan Ayah, Penghasilan Ayah, Pekerjaan Ibu, Penghasilan Ibu, Penerima KPS, dan Penerima KIP serta Layak PIP yang digunakan sebagai label. Setelah data di preprasai, data diubah menjadi model analisis data dan memodelkan data agar sesuai dengan analisis *data mining* yang diharapkan, tujuan transformasi adalah mengubah data yang dipilih ke dalam bentuk prosedur penambangan.



Gambar 3. *Change Type*

Pada gambar 3 merupakan proses pengubahan tipe data dari atribut yang ada dalam dataset, *binomial* diterapkan untuk kolom yang memuat dua jenis data, dan *polynomial* diterapkan untuk kolom yang mengandung tiga jenis data atau lebih.

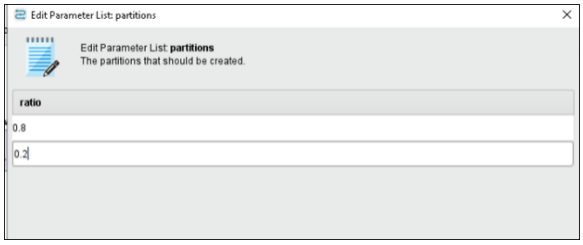


Gambar 4. *Change Role*

Selanjutnya pada gambar 4 ialah mengubah role pada kolom Layak PIP sebagai label. Label berperan sebagai atribut target untuk operator pembelajaran, atau dapat disebut dengan variabel atau kelas.

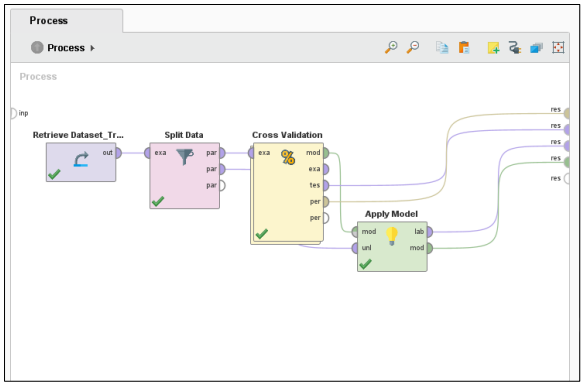
4.3. Proses Pengujian *Naive Bayes Classifier*

Pada tahap ini adalah memulai proses pengujian pada data yang sudah siap diolah menggunakan *RapidMiner*, menambahkan operator *Retrieve* kemudian *drag & drop* data pada layar *process* dan menambahkan operator *Split Data*.



Gambar 5. *Split Data*

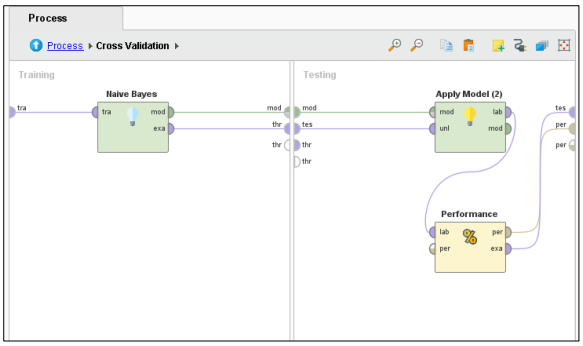
Pada gambar 5 merupakan proses *Split Data* untuk membagi data dengan perbadingan 80% untuk data *training* dan 20% untuk data *testing*..



Gambar 6. Proses Pengujian Awal

Pada gambar 6 merupakan proses pengujian awal dengan menghubungkan operator *Cross Validation* pada data *training* untuk melatih dan mengevaluasi model serta Operator *Apply Model* untuk menerapkan model akhir yang telah dilatih pada data *testing* yang terpisah.

Dalam *Cross Validation*, data akan dibagi secara acak menjadi 10 bagian atau *folds*. Model dilatih menggunakan 9 bagian data, dan bagian yang ke-10 digunakan untuk menguji model, ini dilakukan sebanyak 10 kali, dengan setiap bagian mendapatkan giliran untuk menjadi data uji, dan yang lainnya digunakan untuk pelatihan. Kemudian *error rate* dihitung pada *holdout set*. Setelah 10 kali pengujian, *error rate* dirata-rata untuk mendapatkan gambaran seberapa baik model ini dalam memprediksi secara keseluruhan. *Cross Validation* memastikan bahwa model tidak *overfitting*, yaitu bekerja sangat baik pada data pelatihan tetapi buruk pada data baru atau data yang belum pernah dilihat sebelumnya.



Gambar 7. Desain Pengujian *Naive Bayes*

Sebagaimana dapat dilihat pada gambar 7, pemodelan *Cross Validation* digunakan untuk melatih model yang di dalamnya terdapat 3 bagian yaitu operator *Naive Bayes* yang digunakan pada bagian *training*, operator *Apply Model* untuk mengaplikasikan model pada data *testing* dan *Performance* untuk menampilkan *Confusion Table* yang digunakan untuk menampilkan hasil dari *Accuracy*, *Precision*, *Recall*, dan grafik *AUC*.

4.4. Analisis Hasil Akurasi dan Prediksi

	true Ya	true Tidak	class precision
pred. Ya	132	1	99.25%
pred. Tidak	3	8	72.73%
class recall	97.78%	88.89%	

Gambar 8. Confusion Matrix

Pada gambar 8 merupakan hasil pengujian dari 144 data *training*, jumlah prediksi Ya yang diklasifikasikan sebagai True Ya (merupakan data layak yang diprediksi sesuai) oleh *classifier* yaitu 132 data, prediksi Ya yang diklasifikasikan sebagai True Tidak (merupakan data yang diprediksi layak tetapi ternyata tidak layak) yaitu 1 data, prediksi Tidak yang diklasifikasikan sebagai True Ya (merupakan data yang diprediksi tidak layak ternyata layak) yaitu 3 data, dan prediksi Tidak yang diklasifikasikan sebagai True Tidak (merupakan data tidak layak yang diprediksi sesuai) yaitu 8 data. Perhitungan tabel *Confusion Matrix* dapat dihitung sebagai berikut :

a. *Accuracy*

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$
$$= \frac{132+8}{132+8+1+3} \times 100\%$$
$$= \frac{140}{144} = 0.972 = 97.2\%$$

b. *Precision*

$$Precision = \frac{TP}{TP+FP} \times 100\%$$
$$= \frac{132}{132+1} \times 100\%$$
$$= \frac{132}{133} = 0.992 = 99.2\%$$

c. *Recall*

$$Recall = \frac{TP}{TP+FN} \times 100\%$$
$$= \frac{132}{132+3} \times 100\%$$
$$= \frac{132}{135} = 0.977 = 97.7\%$$

Keterangan :

a. TP (*True Positive*) adalah jumlah *record positif* yang diklasifikasikan sebagai *positif*.
b. FP (*False Positive*) adalah jumlah *record negatif* yang diklasifikasikan sebagai *positif*.
c. FN (*False Negative*) adalah jumlah *record positif* yang diklasifikasikan sebagai *negative*.
d. TN (*True Negative*) adalah jumah *record negative* yang diklasifikasi sebagai *negative*.

Adapun hasil *Perfomance Vector* pada *RapidMiner* adalah sebagai berikut.

a. *Accuracy*

accuracy: 97.24% +/- 4.91% (micro average: 97.22%)

	true Ya	true Tidak	class precision
pred. Ya	132	1	99.25%
pred. Tidak	3	8	72.73%
class recall	97.78%	88.89%	

Gambar 9. *Accuracy*

Pada gambar 9 dapat dilihat dengan mengetahui jumlah data yang diklasifikasi secara benar, maka dapat diketahui akurasi dari hasil *RapidMiner* adalah sebesar 97.24% dari hasil 144 data *training*.

b. *Precision*

precision: 72.73% (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	132	1	99.25%
pred. Tidak	3	8	72.73%
class recall	97.78%	88.89%	

Gambar 10. *Precision*

Pada gambar 10 menunjukan Hasil dari nilai *precision* adalah sebesar 99.25% untuk *class* ya (layak) dan 72.73% untuk *class* tidak (tidak layak).

c. *Recall*

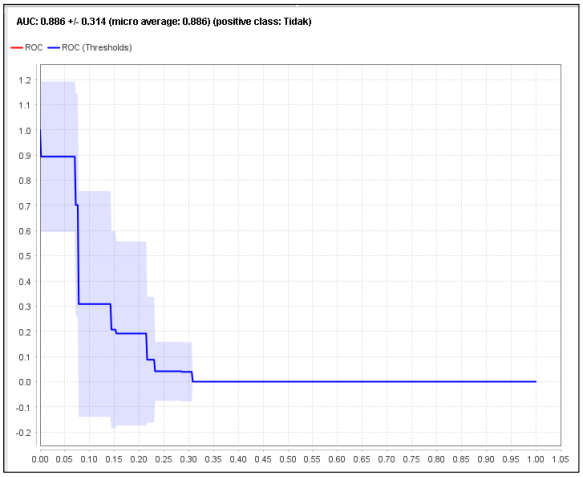
recall: 88.89% (positive class: Tidak)

	true Ya	true Tidak	class precision
pred. Ya	132	1	99.25%
pred. Tidak	3	8	72.73%
class recall	97.78%	88.89%	

Gambar 11. *Recall*

Gambar 11 menunjukan hasil dari nilai *recall* adalah sebesar 97.78% untuk *class* ya (layak) dan 88.89% untuk *class* tidak (tidak layak).

d. *AUC*



Gambar 12. *AUC*

Tingkat keakurasian kurva *ROC/AUC* dapat diklasifikasikan menjadi 5 kelompok, yaitu :

1. 0.90 – 1.00 = *Excellent Clasification*
2. 0.80 – 0.90 = *Good Classification*
3. 0.70 – 0.80 = *Fair Classification*
4. 0.60 – 0.70 = *Poor Classification*
5. 0.50 – 0.60 = *Failure*

Berdasarkan gambar 12 menunjukan klasifikasi yang dihasilkan termasuk ke dalam kelompok *Good Classification* (klasifikasi baik) karena nilai *AUC* yang didapatkan menggunakan algoritma *Naive Bayes* sebesar 0.886.

Hasil pengujian yang telah dilakukan menggunakan algoritma *Naive Bayes* dengan teknik *Cross Validation*, didapatkan hasil kesalahan prediksi untuk siswa penerima bantuan PIP. Kesalahan ini terjadi pada beberapa siswa yang seharusnya layak menerima bantuan namun diprediksi tidak layak, serta siswa yang diprediksi layak namun sebenarnya tidak layak sebagai berikut.

Tabel 1. Analisa Kesalahan Prediksi

Nama	Layak PIP	Prediction Layak PIP
Nurhalimah	Ya	Tidak
Hayatu Ramadhani	Ya	Tidak
Farhan Purwandi	Tidak	Ya
Intan Adawiyah	Ya	Tidak
Risti Nur Amelia	Ya	Tidak
Siti Rodiah	Tidak	Ya
Milan	Tidak	Ya

Tabel 1 di atas merupakan hasil kesalahan prediksi siswa penerima bantuan PIP secara keseluruhan, yang diperoleh dari pengujian *Cross Validation* pada 144 data *training* dan pengujian model akhir pada 36 data *testing* dengan jumlah keseluruhan data yang diuji adalah 180 data, 173 data diprediksi benar dan 7 data diprediksi salah.

5. KESIMPULAN DAN SARAN

Hasil penelitian dan pengujian menunjukkan bahwa algoritma *Naive Bayes* dapat diimplementasikan dengan sangat baik dalam mengklasifikasikan kelayakan siswa penerima bantuan PIP, dengan tingkat akurasi 97,24%, *precision* 72,73%, *recall* 88,89%, dan nilai *AUC* 0,886 yang dikategorikan sebagai *good classification*. Dari keseluruhan 180 data yang diuji, 173 data diprediksi dengan benar dan 7 data diprediksi salah, menjadikan algoritma ini efektif sebagai bahan evaluasi dalam pengambilan keputusan terkait penentuan penerima bantuan PIP berikutnya.

Untuk penelitian selanjutnya, disarankan untuk melakukan perbandingan dengan metode algoritma lain seperti *C4.5*, *Random Forest*, dan *Support Vector Machine* untuk mengevaluasi dan membandingkan tingkat akurasi, *precision*, serta *recall* dari masing-masing algoritma, guna menentukan metode klasifikasi yang paling optimal.

DAFTAR PUSTAKA

[1] N. I. Nurhidayati, Y. Yahya, F. Fathurrahman, L. . Samsu, and W. Amnia, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)," *Infotek J. Inform. dan Teknol.*, vol. 6, no. 1, pp. 177–188, 2023, doi: 10.29408/jit.v6i1.7529.

[2] M. Sari, S. Musdalifah, and E. A. Asfar, "Implementasi Kebijakan Kartu Indonesia Pintar Dalam Upaya Pemerataan Pendidikan di MTsN 1 Watampone 1Mauliana," *Adaara J. Manaj. Pendidik. Islam*, vol. 3, no. 1, pp. 848–870, 2021, doi: 10.35673/ajmpi.v8i1.422.

[3] Bustami, "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi pada BPR Pantura," 2019, [Online]. Available: https://repository.nusamandiri.ac.id/index.php/repo/viewitem/13890

[4] R. Setiawan and A. Triayudi, "Klasifikasi Status Gizi Balita Menggunakan Naïve Bayes dan K-Nearest Neighbor Berbasis Web," *J. Media Inform. Budidarma*, vol. 6, no. 2, p. 777, 2022, doi: 10.30865/mib.v6i2.3566.

[5] I. Randi Purnama, "Analisis Program Indonesia Pintar Dalam Mengurangi Putus," vol. 5, no. 1, pp. 66–73, 2023.

[6] O. Rini and S. O. Kunang, "Implementasi Data Mining Menggunakan Metode Naive Bayes Untuk Penentuan Penerima Bantuan Program Indonesia Pintar (Pip) (Studi Kasus : Sd Negeri 9 Air Kumbang)," *Bina Darma Conf. ...*, pp. 714–722, 2021.

[7] S. Faisal, "Klasifikasi Data Minning Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang," *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2019, doi: 10.36805/technoxplore.v4i1.541.

[8] M. A. W. Pratama, M. Fuad, H. Hazriani, and ..., "Penentuan Status Penerima Bantuan Indonesia Pintar pada SMKN 9 Bulukumba Dengan Metode Naive Bayes," *Pros. SISFOTEK*, 2023, [Online]. Available: http://seminar.iaii.or.id/index.php/SISFOTEK/article/view/387%0Ahttp://seminar.iaii.or.id/index.php/SISFOTEK/article/download/387/319

[9] N. Riyanah and F. Fatmawati, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 206–213, 2021, doi: 10.35746/jtim.v2i4.117.

[10] N. Pramesti, "Klasifikasi Persediaan Barang Menggunakan Naïve Bayes," *J. Data Sci. Inform.*, vol. 1, no. 2, pp. 53–57, 2021, [Online]. Available: http://publikasi.bigdatascience.id