

MENENTUKAN KELAYAKAN PENERIMA KIP MENGGUNAKAN KLASIFIKASI DENGAN METODE ALGORITMA NAIVE BAYES

Afdul Hadi, Irfan Ali

Program Studi Komputerisasi Akuntansi STMIK IKMI Cirebon

Program Studi Rekayasa Perangkat Lunak STMIK IKMI Cirebon

Jl. Perjuangan No. 10B Majasem Kec. Kesambi Kota Cirebon Tlp. 0231-490480-490481

afdulhadi13@gmail.com

ABSTRAK

Pemerintah berupaya meningkatkan pendidikan dengan meluncurkan Program Indonesia Pintar (PIP) yang mendukung Kartu Indonesia Pintar (KIP). KIP memberikan dana berupa beasiswa pendidikan kepada anak usia 6-21 tahun. Sejak akhir tahun 2014, program KIP menjadi bagian dari Program Bantuan Siswa Miskin (BSM). Ketidaktepatan target dalam penyaluran KIP di SMAN 1 Lemahabang mengakibatkan target tidak merata dan tidak sesuai harapan. Begitu banyak siswa tanpa dukungan itu harus memilih pendidikan, memilih pekerjaan dan memotong biaya. keluarga Tujuan dari penelitian ini adalah menggunakan algoritma Naive Bayes untuk membidik dan memeringkat SMAN 1 Lemahabang sesuai harapan sehingga sekolah dapat dengan mudah menentukan kecocokan penerima KIP-nya. Oleh karena itu, penulis menggunakan algoritma Naive Bayes sebagai metode yang dapat mengklasifikasikan calon penerima KIP dengan akurasi dan presisi yang tinggi. Naive Bayes mencapai beberapa hasil. Hasil Precision 60,63%, Hasil Precision 92,62%, Hasil Recall 57,50%, dan Hasil AUC 0,730. Menurut tingkat diagnostik saat ini, nilai AUC sebesar 0,730 termasuk dalam klasifikasi wajar. Artinya algoritma Naive Bayes untuk pemilihan metode sudah cukup.

Kata Kunci: KIP, Naive Bayes, Rapid Miner

1. PENDAHULUAN

Ada berbagai jenis hibah yang dikeluarkan oleh pemerintah. Kartu Pintar Indonesia (KIP) adalah salah satunya. Program KIP merupakan dana bantuan pemerintah, dan program tersebut digunakan untuk memberikan kemudahan bagi masyarakat miskin dan tidak dapat melanjutkan pendidikan anaknya ke jenjang yang lebih tinggi. Program ini akan meningkatkan pendidikan dan kesejahteraan masyarakat miskin.

Bagi masyarakat Indonesia, pendidikan adalah hal terpenting untuk masa depan mereka. Naive Bayes adalah metode peramalan perhitungan probabilitas sederhana berdasarkan teorema Bayes dengan asumsi independensi yang kuat. Metode Naif Bayes membutuhkan sedikit pelatihan data, sehingga mudah untuk memperkirakan parameter yang dibutuhkan untuk klasifikasi[1].

Data mining adalah metode pengolahan data yang bertujuan untuk mendapatkan informasi penting tentang data. Menentukan objek data mana yang berada di kelas tertentu dari kelas lain disebut klasifikasi. Klasifikasi adalah proses menemukan kesamaan sifat dalam kelompok atau kelas. Proses klasifikasi itu sendiri memiliki satu tujuan yaitu memprediksi bagian tertentu dari suatu objek yang tidak diketahui. Metode Naif Bayes adalah metode klasifikasi yang membutuhkan sekumpulan data training multi atribut untuk menentukan kelayakan penerima KIP.

Kurangnya koordinasi dalam pendistribusian KIP di SMAN 1 Lemahabang mengakibatkan pemilihan calon penerima KIP yang tidak konsisten

dan mengecewakan, sehingga banyak masyarakat miskin dan kurang mampu yang kehilangan bantuan tersebut. Prediksi yang mengidentifikasi calon penerima KIP dapat membantu sekolah mengidentifikasi secara akurat penerima KIP dan non-penerima KIP dengan algoritma naif bayes..

2. TINJAUAN PUSTAKA

2.1. Kartu Indonesia Pintar

Kartu Indonesia Pintar (KIP) menjamin semua anak usia sekolah berusia 6-21 tahun dari keluarga pemegang Program Kartu Kesejahteraan Sosial (KKS) Indonesia Pintar (PIP) dapat bersekolah. Ini adalah hibah pendidikan. Papan untuk sekolah, madrasah, kelompok belajar (paket latihan A/B/C) [2].

Siswa yang kecewa atau kurang beruntung sangat rentan terhadap krisis putus sekolah saat ini dan karenanya membutuhkan kartu Indonesia pintar. Ini karena situasi keuangan keluarga siswa yang sulit. Dukungan untuk membantu orang-orang ini memutuskan untuk berhenti kuliah dan mendapatkan pekerjaan. Tim Nasional Penanggulangan Kemiskinan yang dipimpin Kementerian Pendidikan dan Kebudayaan (Kemendikbud) meluncurkan Kebijakan Kartu Indonesia Pintar (TNP2K). Tujuan dari program ini adalah untuk memberikan akses pendidikan siswa yang berkualitas kepada masyarakat kurang mampu, mencegah anak putus sekolah dan memenuhi kebutuhan pendidikan mereka. Siswa diharapkan menggunakan hak ini untuk memenuhi kebutuhan sekolah [3].

2.2. Penambangan Data

Penambangan data adalah kegiatan mengekstraksi pola menarik dari sejumlah besar data yang dapat disimpan dalam basis data, gudang data, atau penyimpanan informasi lainnya. Sistem basis data, gudang data, statistik, pembelajaran mesin, pencarian informasi, dan komputasi tingkat tinggi hanyalah beberapa area yang terkait dengan penambangan data. Selain itu, area seperti jaringan saraf, pengenalan pola, analisis data spasial, database gambar, pemrosesan sinyal mendukung penambangan data [1].

Penambangan data adalah proses penyulingan sejumlah besar data yang disimpan dalam arsip dan menggunakan teknik pengenalan pola dan alat statistik dan matematika. untuk menemukan koneksi, pola, dan tren baru yang penting. Menemukan pola dan wawasan menarik dalam data dalam jumlah besar adalah proses yang disebut penambangan data[4].

2.3. Klasifikasi

Klasifikasi adalah metode pembelajaran data yang digunakan untuk prediksi suatu nilai yang ada pada sekumpulan atribut. Klasifikasi biasanya memperoleh seperangkat *rules* biasa disebut aturan. Aturan ini dipakai untuk metrik guna mendapatkan prediksi suatu objek dari data yang akan dilakukan prediksi[5].

2.4. Naive Bayes

Pengklasifikasi probabilistik, yang disebut Naive Bayes, meringkas frekuensi dan kombinasi nilai dalam kumpulan data yang disediakan. Ilmuwan yang berasal dari england yaitu T. Bayes mengusulkan Klasifikasi Naif Bayes, yang menggunakan metode probabilistik dan statistik untuk memprediksi peluang masa depan berdasarkan kinerja dari masa-masa yang lampau.. T. Bayes disatukan dengan 'naif', dengan asumsi kondisi saling eksklusif antara atribut[6]

2.5. Rapidminer

Rapidminer adalah software untuk mengolah data. RapidMiner menggabungkan statistik, kecerdasan buatan, dan teknologi basis data untuk mengungkap pola dari kumpulan data besar menggunakan prinsip dan algoritme penambangan data. Operator mengizinkan pengguna Rapid Miner untuk melakukan perhitungan data besar dengan cepat dan mudah. Gunakan operator ini untuk mengubah data. Setelah data terhubung ke node operator, kita perlu terhubung ke simpul hasil untuk menampilkan hasilnya. Keluaran Rapidminer juga dapat dilihat secara visual sebagai grafik. Rapidminer telah menjadi salah satu perangkat lunak paling populer untuk mengekstraksi data menggunakan teknik penambangan data yang inovatif Rahmat C.T.I[7].

3. METODE PENELITIAN

3.1. Pengumpulan Data

Ada beberapa teknik dalam pengumpulan data, yaitu sebagai berikut :

a. Observasi

Observasi adalah suatu cara untuk memperoleh informasi tentang suatu peristiwa yang telah atau sedang terjadi. Dengan menggunakan metode observasi, peneliti melakukan observasi untuk mendapatkan bahan penelitian. .

b. Wawancara

Wawancara adalah metode pengumpulan data yang memungkinkan peneliti memperoleh informasi langsung dari tempat penelitian. Tujuan dari kegiatan ini adalah untuk berdiskusi dan bertanya dengan subjek data atau orang yang lebih mengetahui tentang subjek investigasi. Wawancara ini dilakukan dengan Bimbingan Konseling (BK) dan Tata Usaha (TU) di SMAN 1 Lemahabang.

3.2. Sumber Data

Sumber data dapat dikumpulkan dengan berbagai cara, yaitu sebagai berikut:

a) Penelitian sumber data

SMAN 1 Lemahabang merupakan lokasi penelitian yang terletak di Jl. Kanci-Sindanglaut, Cipeujeuh Wetan, Kec. Lemahabang, Kabupaten Cirebon, Jawa Barat 45183. Fieldwork dilaksanakan mulai tanggal 19 Januari 2023 sampai dengan 19 Februari 2023. Total data yang diperoleh dari SMAN 1 Lemahabang berjumlah 957 data calon penerima smart card Indonesia, dan data tersebut dipilih berdasarkan berbagai atribut, seperti nama, jenis kelamin, nisn, tempat lahir, tanggal, dll. Hanya beberapa atribut yang digunakan. Kelahiran, detail ayah, detail ibu, dan detail penerima kip.

b) Sumber data primer

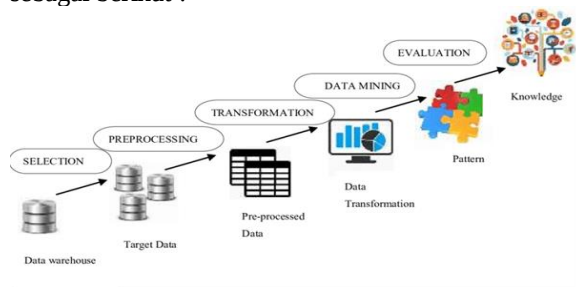
Data primer adalah informasi yang didapatkan melalui pengamatan langsung terhadap pelaksanaan program Kartu Indonesia Pintar atau dari wawancara dengan orang-orang yang terkait langsung dan terlibat dalam topik penelitian. Data ini berkaitan dengan program, prosedur administrasinya, dan keterbatasan yang dihadapi dalam mencapai tujuan program di sektor smart card Indonesia.

c) Sumber data sekunder

Data sekunder, disebut juga informasi yang sebelumnya dikumpulkan atau diperoleh secara tidak langsung, namun masih relevan dengan topik penelitian. Data sekunder ini biasanya meliputi struktur organisasi, dokumen profil perusahaan, dan data yang diperoleh dari websitenya di SMAN 1 Lemahabang.

3.3. Tahap Perancangan

Ada beberapa tahapan pada proses KDD yaitu sebagai berikut :



Gambar 1. Langkah-langkah KDD

1. Pemilihan data

Pemilihan data dimulai dengan sekumpulan kumpulan data fungsional. Hal ini dapat dilakukan dengan mencari informasi di database (KDD) sebelum melanjutkan ke tahap information retrieval. Data terpilih yang digunakan dalam proses penambangan data disimpan dalam file terpisah dari basis data operasional.

2. Pembersihan

Sebelum memasuki proses data mining menjadi *Knowledge Discovery in Database (KDD)*, data harus dibersihkan. Pembersihan ini dilakukan untuk menghapus data serupa, meninjau data yang terus berubah, dan memperbaiki kesalahan.

3. Transformasi

Langkah ini mengubah data menjadi format yang sesuai sehingga proses penambangan data dapat

dilakukan. Langkah ini bergantung pada model atau tipe data yang diambil dari database.

Transformasi merupakan fase dimana akan menggabungkan atau mengubah data menjadi bentuk yang dapat diproses. Beberapa teknik konversi data memerlukan format data tertentu sebelum dapat digunakan. beberapa teknik umum. B. Analisis Asosiasi dan Klaster hanya dapat menerima data input kategorikal. Oleh karena itu, saya perlu membagi data menjadi interval yang berbeda dalam bentuk angka berurutan. Transformasi data adalah nama lain untuk teknik ini[8].

4. Data mining

Pada fase ini, kami menggunakan metode khusus untuk mencari model dan informasi yang menarik dari data yang dipilih. Ada banyak metode berbeda, salah satunya digunakan oleh para peneliti - algoritma Naive Bayes.

5. Evaluasi

Pada fase ini, proses data mining dibuat dan disajikan dalam bentuk yang mudah dipahami oleh yang terlibat. Langkah ini juga merupakan bagian dari proses *knowledge discovery in database (KDD)* di dalam database. Hasil teknik data mining menghasilkan pola model prediksi yang dievaluasi untuk menilai apakah hipotesis yang ada terpenuhi.

4. HASIL DAN PEMBAHASAN

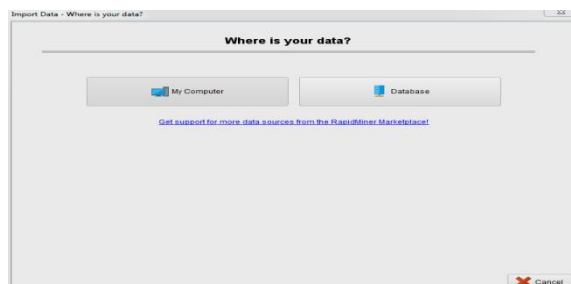
4.1. Mengetahui data penerima KIP

a. Seleksi data

Tabel 1 Dataset sebelum di seleksi

Nama	Kelas	No. KPS	Nomor KIP	Penerima KIP	Tempat Lahir	Nama	Pekerjaan	Penghasilan
Adila nuraini	X.1			Tidak	Cirebon	HARUN	Buruh	Rp. 500,000 - Rp. 999,999
Ahmad taufiq ramadhan	X.1			Tidak	Cirebon	BAHRUDIN	Pedagang Kecil	Rp. 500,000 - Rp. 999,999
Anindya Hafizh	X.1		T18ZG6	Ya	Cirebon	WARIDI SUPRIYADI	Karyawan Swasta	Rp. 1,000,000 - Rp. 1,999,999

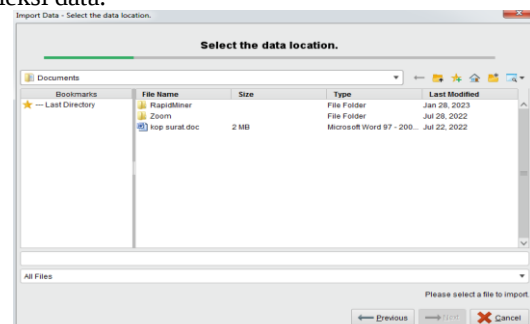
Tabel 1 menunjukkan kumpulan data berikutnya sebelum seleksi di rapid miner. Data tersebut masih berisi data yang tidak dibutuhkan pembuatnya. Dengan kata lain, data memenuhi keinginan penulis dan memenuhi kebutuhan penulis.



Gambar 2. Import Data

Sebelum memilih data di Rapidminer, penulis harus terlebih dahulu mengimpor data yang akan digunakan untuk pemilihan data.

Pada gambar 2 merupakan langkah awal untuk mengimpor data untuk dilakukan seleksi data.



Gambar 3. Mencari Lokasi Data

Pada gambar 3 penulis mencari letak atau lokasi penyimpanan dataset untuk menghasilkan sebuah data yang nantinya sudah diseleksi oleh sistem atau rapidminer.

Row No.	Peserta KIP	No	Nama	JK	NISN	Tempat Lahir	Tanggal Lahir	Ayah	Ibu
1	Tidak	1	ADILA NURAL	P	67166719	CIREBON	2006-09-16	HARUN	Bun...
2	Tidak	2	ABIMO TALUF	L	3677036435	CIREBON	2007-09-23	BAHRUDIN	Pes...
3	Ya	3	Aninda Hafid	P	74836380	Cirebon	2007-09-29	WARDI SUP	Kar...
4	Tidak	4	AYESHA NUL	P	71023781	CIREBON	2007-09-26	ROSDIN	Kar...
5	Tidak	5	BIANG BING	L	67828988	CIREBON	2006-10-17	Amsar	Pes...
6	Tidak	6	CHIKA DVA	P	3652545059	KUNINGAN	2006-11-14	ASEP SUGA	Pes...
7	Tidak	7	DINDA MERL	P	77788333	Shimon	2007-05-25	DARSONO	Wra...
8	Tidak	10	Faisal Shihab	L	75992741	Cirebon	2007-02-10	Ahmad Fathur	Wra...
9	Tidak	11	FIRDA NAYLA	P	77888457	Cirebon	2007-06-25	SOLEH	Lan...
10	Tidak	14	NANA DWIN	L	72340348	Cirebon	2007-08-29	Asep Sapri	Bun...
11	Tidak	15	JANJAN ADY	L	77191190	CIREBON	2007-01-13	SAEFUL MUJ	Wra...
12	Tidak	16	KHARUN NISA	P	62105482	CIREBON	2006-10-29	MURDOKANT	Pes...
13	Tidak	17	KHOZMUL L	L	73284290	CIREBON	2007-08-08	Ridwan	Kar...
14	Tidak	18	KREINA AKA	L	64609474	Shimon	2006-10-27	ALHEDYAT	Pes...

Gambar 4. Hasil Seleksi Data

Pada gambar 4 menunjukkan bahwasannya hasil dari seleksi data sudah sesuai keinginan dan sudah sesuai apa yang dibutuhkan oleh penulis.

b. Pembersihan

Pada tahapan reprocessing akan dilakukan sebuah pembersihan data yang memiliki data kosong atau missing values.

Attribute	Type	Missing	Statistics
No	Integer	0	1
Nama	Nominal	0	1
JK	Nominal	0	1
NISN	Real	0	1
Tempat Lahir	Nominal	0	1
Tanggal Lahir	Nominal	0	1
Ayah	Nominal	0	1

Gambar 5. Cek Missing Values

Reprocessing sendiri memiliki arti yaitu pembersihan data, yang dimana nantinya data tersebut akan dilakukan pembersihan data. Apabila data tersebut memiliki missing values, maka akan dilakukan pembersihan menggunakan operator Replace Missing Values. Jika data tersebut tidak memiliki sebuah data kosong atau missing values, maka dapat dilanjutkan ke tahap berikutnya.

Pada gambar 5 menunjukkan bahwasannya data tersebut tidak memiliki sebuah data yang kosong atau missing value, maka dari itu proses tersebut dapat dilanjutkan ke tahap berikutnya.

c. Transformasi

Attribute	Type	Format
Pekerjaan 1	Nominal	polynomial
Penghasilan 1	Nominal	polynomial
Ibu	Nominal	polynomial
Pekerjaan 2	Nominal	polynomial
Penghasilan 2	Nominal	polynomial
Penerima	Nominal	polynomial

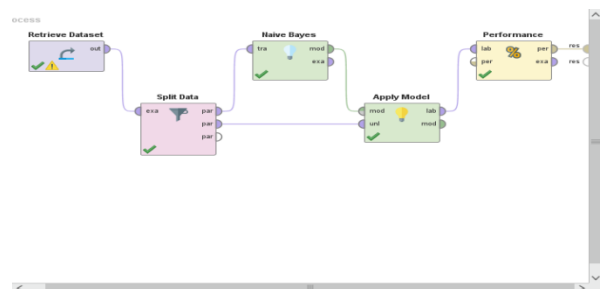
Gambar 6. Transformasi

Pada gambar 6 di tahapan transformasi data akan diberikan sebuah inisialisasi pada sebuah data yang dibutuhkan oleh Naive Bayes, yang nantinya akan dilakukan sebuah prediksi menggunakan label tersebut.

4.2. Menerapkan algoritma naive bayes

Terdapat 2 tahapan dalam pemrosesan pada rapidminer sesuai proses KDD yaitu sebagai berikut;

a. Data Mining



Gambar 7 Naive Bayes

Berdasarkan Gambar 7 untuk mengimplementasikan model Naive Bayes, penulis membutuhkan sebuah aplikasi bernama Rapidminer, versi 9.9.003, dan menjalankannya menggunakan operator: retrieve, split data, naive bayes, apply model, dan performance. Pada data *training* diterapkan algoritma klasifikasi yang disebut Naive Bayes, tetapi pada data *testing* ada model yang diterapkan guna menerapkan model naif bayes serta mengukur kinerja model naive Bayes untuk kinerja.

	true Tidak	true Ya	class precision
pred. Tidak	138	11	92.62%
pred. Ya	102	36	26.09%
class recall	57.50%	76.60%	

Gambar 8 Perhitungan Rapidminer

Berdasarkan hasil pada gambar 8 tiap tingkatan akurasi yang memiliki nilai tersendiri dari ketiga tingkatan tersebut dimulai dari accuracy sebesar 60.63%, percision sebesar 26.09%, dan recall 57.50%.

Dibawah ini adalah perhitungan dari tools rapidminer yang menghasilkan sebuah data dengan tingkat keakurasian yang berbeda-beda yaitu sebagai berikut

1. Akurasi

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{138 + 36}{138 + 36 + 11 + 102} = \frac{174}{287} = 60.63\%$$

2. Percision

$$\frac{\frac{TP}{TP + FP}}{\frac{138 + 11}{149}} = \frac{92.62\%}{149}$$

3. Recall

$$\frac{\frac{TP}{TP + FN}}{\frac{138 + 102}{240}} = \frac{57.50\%}{240}$$

Terdapat 4 istilah pada pengukuran menggunakan confusion matrix sebagai hasil klasifikasi. Istilah dari keempat tersebut ialah True positive (TP), true negative (TN), false positive (FP), and false negative (FN). Berikut perhitungan manual pada confusion matrix.

Tabel 1. Confusion Matrix

		actual values		
		positive	Negative	
Predict values	Positive	TP	FP	
	negative	FN	TN	

Berikut keterangan yang ada pada tabel 2 :

1. True Positive (TP) adalah memprediksi siswa penerima Kartu Indonesia Pintar (KIP) dan memang benar siswa tersebut menerima Kartu Indonesia Pintar (KIP).
2. True Negative (TN) adalah memprediksi siswa bukan penerima Kartu Indonesia Pintar (KIP) dan memang benar siswa tersebut bukan penerima Kartu Indonesia Pintar (KIP).
3. False Positive (FP) adalah memprediksi siswa penerima Kartu Indonesia Pintar (KIP) dan prediksi salah, ternyata siswa tersebut memang bukan penerima Kartu Indonesia Pintar (KIP).
4. False Negative (FN) adalah memprediksi siswa bukan penerima KIP dan di prediksi salah, ternyata siswa tersebut penerima Kartu Indonesia Pintar.

Lalu ada perhitungan manual untuk mencari nilai akurasi dan recall seperti berikut:

1. Akurasi

$$\frac{\frac{TP + TN}{TP + TN + FP + FN}}{\frac{138 + 36}{138 + 36 + 11 + 102}} = \frac{60.63\%}{287}$$

2. Percision

$$\frac{TP}{TP + FP}$$

$$\frac{138}{138 + 11} = \frac{92.62\%}{149}$$

3. Recall

$$\frac{\frac{TP}{TP + FN}}{\frac{138 + 102}{240}} = \frac{57.50\%}{240}$$

b. Evaluasi

Pada tahapan ini akan menghasilkan pola-pola dari proses data mining berupa visual seperti sebagai berikut:

1. Akurasi

Process	Results	Process result	Completed	Feb 14, 2023 9:02:49 AM	(execution time: 0 s)
Process (1 results)	Process result	Completed	Feb 14, 2023 11:32:11 AM	(execution time: 0 s)	
Process (1 results)	Process result	Completed	Feb 14, 2023 11:32:28 AM	(execution time: 0 s)	
Process (1 results)	Process result	Completed	Feb 14, 2023 11:33:07 AM	(execution time: 0 s)	
Process (2 results)	Process result	Completed	Feb 14, 2023 11:33:54 AM	(execution time: 0 s)	

Gambar 9. Results

Pada gambar 9 menunjukkan bahwa percobaan menggunakan algoritma naive bayes dengan tools rapidminer yang dimana untuk mengeksekusi hanya membutuhkan waktu 0 second, yang artinya eksekusi tersebut berjalan dengan cepat.

Criterion	Accuracy	Recall	Precision	F1 Score
accuracy	60.63%			
recall		57.50%		
precision			92.62%	
F1 Score				73.83%

Gambar 10. Hasil Akurasi

Pada gambar 10 menunjukkan hasil yang diperoleh dari akurasi sebesar 60.63%, dengan hasil tersebut bahwasannya model yang digunakan dapat dilakukan untuk klasifikasi guna menentukan kelayakan calon penerima kip. Nilai tersebut juga dikarenakan oleh kompleksnya data yang digunakan, sehingga dapat memperhitungkan dengan tepat.

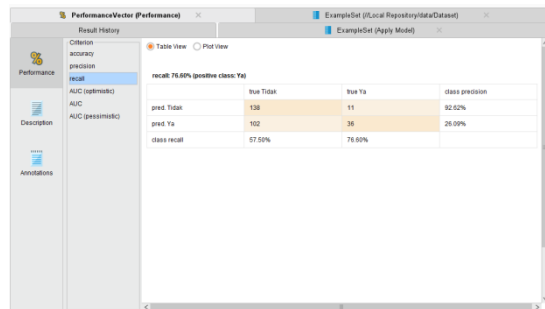
2. Presisi

Criterion	Accuracy	Recall	Precision	F1 Score
accuracy	60.63%			
recall		57.50%		
precision			92.62%	
F1 Score				73.83%

Gambar 11. Hasil Presisi

Hasil precision dapat dilihat pada gambar 11 di tahapan ini menunjukkan bahwa nilai yang diperoleh dari precision sebesar 26.09%, yang dimana artinya akurasi precision sendiri memiliki tingkat akurasi antara informasi yang diminta oleh pengguna dan hasil yang diberikan oleh sistem.

3. Recall

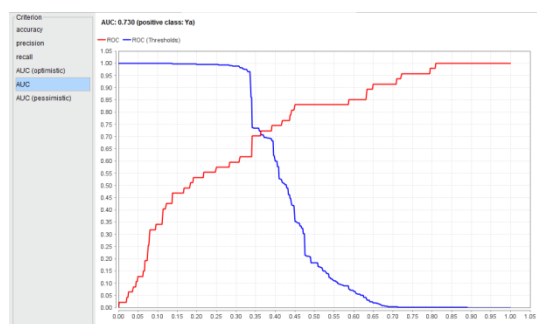


Gambar 12. Hasil Recall

Hasil dari data tersebut dapat dilihat pada gambar 12, yang dimana menunjukkan hasil dari akurasi recall adalah sebesar 76.60%, yang artinya akurasi recall memiliki tingkat keberhasilan sistem dalam layanan data.

Berdasarkan tingkat akurasi yang diperoleh sebesar 60.63% yang sudah dijelaskan diatas dengan rincian. Prediksi Tidak dan true Tidak sebesar 138, Prediksi Tidak dan true Ya sebesar 11, Prediksi Ya dan true Tidak sebesar 102, Prediksi Ya dan true Ya sebesar 36. Dengan memiliki class recall Tidak sebesar 57.50% dan class recall Ya sebesar 76.60%.

4. Hasil AUC



Gambar 13 Hasil AUC

Algoritma Naive Bayes untuk memprediksi kelayakan calon penerima Kartu Indonesia Pintar (KIP) dengan perolehan akurasi sebesar 60.63%, dan hasil dari AUC sebesar 0.730. Berikut tingkat keakurasian kurva ROC/AUC dapat diklasifikasikan menjadi 5 kelompok, yaitu;

- 0.90 - 1.00 = Excellent Clasification
- 0.80 - 0.90 = Good Classification
- 0.70 - 0.80 = Fair Classification
- 0.60 - 0.70 = Poor Classification
- 0.50 - 0.60 = Failure.

Berdasarkan nilai AUC pada gambar 13 sebesar 0.730 yang artinya dimana sesuai tingkat diagnosa, maka nilai AUC sebesar 0.730 termasuk ke dalam kelas fair classification untuk melakukan klasifikasi kelayakan calon penerima Kartu Indonesia Pintar (KIP).

5. KESIMPULAN DAN SARAN

Berdasarkan penjelasan di atas, dapat disimpulkan dari penelitian ini bahwa: Berdasarkan hasil dari pengolahan data menggunakan data SMAN 1 Lemahabang, menjelaskan bahwa data yang diambil dari Januari sampai dengan Desember 2022 tentang calon penerima KIP. Dari tabel data sekolah pada SMA Negeri 1 Lemahabang yang menggunakan permodelan Naive Bayes menghasilkan tingkat Accuracy sebesar 60.63% dengan rincian Prediksi Tidak dan True Tidak sebesar 138, Prediksi Tidak dan True Ya sebesar 11, Prediksi Ya dan True Tidak sebesar 102, Prediksi Ya dan True Ya sebesar 36. Dengan class recall Tidak sebesar 57.50% dan class recall Ya sebesar 76.60%.

DAFTAR PUSTAKA

- [1] 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 Dharma Conf.*, pp. 714–722, 2021, [Online]. Available: <https://conference.binadarma.ac.id/index.php/B-DCCS/article/view/2450%0Ahttps://conference.binadarma.ac.id/index.php/B-DCCS/article/download/2450/1100>
- [2] S. I. Stmik and K. Binjai, "Seminar Nasional Informatika (SENATIKA) Prosiding SENATIKA 2021 Sistem Pendukung Keputusan Penerimaan Kartu Indonesia Pintar (KIP) Pada Siswa Menggunakan Metode Moora," 2021.
- [3] S. Sulaiman, "Keterampilan Manajerial Kepala Madrasah Dalam Meningkatkan Kinerja Guru Di Madrasah Aliyah-Falah Arungkeke Kabupatenjeneponto," *Adaara J. Manaj. Pendidik. Islam*, vol. 8, no. 1, pp. 848–870, 2019, doi: 10.35673/ajmpi.v8i1.422.
- [4] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [5] S. Wahyuningsih and D. R. Utari, "Perbandingan Metode K-Nearest Neighbor , Naive Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit," *Konf. Nas. Sist. Inf. 2018 STMIK Atma Luhur Pangkalpinang*, 8 – 9 Maret 2018, pp. 619–623, 2018.
- [6] A. Sitompul, Suhada, and Safulah, "Teknik Data Mining Dalam Prediksi Jumlah Siswa Baru Dengan Algoritma Naive Bayes," *Kesatria J.*

- Penerapan Sist. Inf. (Komputer Manaj.*, vol. 2, no. 2, pp. 108–117, 2021, [Online]. Available: <http://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/65>
- [7] D. Novianti, “Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner,” *Paradig. - J. Komput. dan Inform.*, vol. 21, no. 1, pp. 49–54, 2019, doi: 10.31294/p.v21i1.4979.
- [8] D. Nofriansyah, K. Erwansyah, and M. Ramadhan, “Penerapan Data Mining dengan Algoritma Naive Bayes Clasifier untuk Mengetahui Minat Beli Pelanggan terhadap Kartu Internet XL (Studi Kasus di CV. Sumber Utama Telekomunikasi),” *J. Saintikom*, vol. 15, no. 2, pp. 81–92, 2018.