

IMPLEMENTASI DATA MINING UNTUK MENGLASIFIKASI PENJUALAN PRODUK TERLARIS PADA KOPERASI AL-BAEDLOWI MENGUNAKAN ALGORITMA NAÏVE BAYES

Muhammad Irfan, Ahmad Aftah Syukron

Teknik Informatika, Universitas Ma'arif Nahdlatul Ulama Kebumen
Jalan Kutoarjo Km. 5 Jatisari, Kebumen, Jawa Tengah, Indonesia
irfaniyyah97@gmail.com

ABSTRAK

Koperasi Pondok Pesantren (Kopontren) Manbaul Ihsan Al-Baedlowi di Purbalingga didirikan untuk memenuhi kebutuhan santri dan masyarakat, seperti alat tulis dan perlengkapan mandi. Proses identifikasi pola pembelian konsumen masih dilakukan secara manual, yang kurang efisien dan memakan waktu. Untuk itu, diperlukan pendekatan modern seperti *data mining* dengan algoritma *Naïve Bayes* untuk mengidentifikasi pola, menganalisis data, dan memberikan rekomendasi berbasis probabilitas. Penelitian ini memiliki tujuan untuk melakukan klasifikasi data penjualan di Koperasi Al-Baedlowi menggunakan algoritma *Naïve Bayes* untuk mengetahui hasil dari model yang dibuat dalam mengklasifikasi produk yang memiliki penjualan yang laris dan yang kurang laris. Setelah dilakukan proses pengujian didapatkan bahwa hasil pengujian model mendapat *accuracy* sebesar 80.00%, *precision* kelas positif Kurang Laris sebesar 83.33%, dan *recall* kelas positif Kurang Laris sebesar 83.33%. Selain itu didapatkan nilai *AUC* sebesar 0.917 untuk kelas positif Kurang Laris yang dapat digolongkan sebagai klasifikasi yang sangat baik (*Excelent Classification*).

Kata kunci : *Data Mining, Klasifikasi, Naïve Bayes, Koperasi, Produk Terlaris*

1. PENDAHULUAN

Koperasi Pondok Pesantren (Kopontren) Manbaul Ihsan Al-Baedlowi, yang berlokasi di Kabupaten Purbalingga, didirikan dengan tujuan utama untuk memenuhi kebutuhan para santri dan masyarakat sekitar. Koperasi ini menyediakan berbagai barang kebutuhan sehari-hari, seperti alat tulis, perlengkapan mandi, dan kebutuhan lain yang dibutuhkan oleh santri maupun masyarakat umum. Tidak hanya menjadi tempat untuk melakukan transaksi jual beli, koperasi ini juga memiliki peran strategis dalam memberdayakan ekonomi di lingkungan pesantren dan komunitas sekitarnya.

Sebagai pusat aktivitas ekonomi, Kopontren Manbaul Ihsan berfungsi lebih dari sekadar penyedia barang kebutuhan. Koperasi ini menjadi motor penggerak dalam mendorong terbentuknya berbagai usaha baru yang memberikan manfaat ekonomi sekaligus membantu meningkatkan kesejahteraan masyarakat di sekitarnya. Selain itu, koperasi ini juga menjadi sarana untuk mengembangkan keterampilan kewirausahaan bagi para anggotanya, memperluas relasi, serta mendukung pengembangan diri mereka dalam menghadapi tantangan ekonomi modern. Dengan demikian, keberadaan koperasi ini tidak hanya memenuhi kebutuhan material, tetapi juga berkontribusi terhadap penguatan kemandirian ekonomi santri dan masyarakat.

Namun, meskipun koperasi ini telah memberikan dampak positif yang signifikan, terdapat kendala mendasar dalam proses pengelolaan data, khususnya dalam mengidentifikasi pola pembelian konsumen. Saat ini, proses tersebut masih dilakukan secara manual, yang memerlukan waktu dan tenaga yang tidak sedikit. Pendekatan tradisional ini menjadi hambatan dalam pengambilan keputusan strategis

yang berbasis data, seperti analisis tren penjualan, pengelolaan stok barang, dan perencanaan promosi. Hal ini menyebabkan koperasi menghadapi risiko *overstock* atau kekurangan stok barang tertentu, yang pada akhirnya dapat memengaruhi tingkat kepuasan konsumen dan efisiensi operasional.

Dalam era digital yang semakin berkembang, teknologi menjadi solusi potensial untuk mengatasi keterbatasan tersebut. Salah satu pendekatan modern yang relevan adalah penggunaan *data mining*. *Data mining* adalah proses penggalian informasi berharga dari kumpulan data besar dengan memanfaatkan metode dan algoritma tertentu. Dengan *data mining*, koperasi dapat mengidentifikasi pola, tren, dan hubungan yang sebelumnya tersembunyi dalam data penjualan mereka. Salah satu algoritma yang dapat digunakan dalam konteks ini adalah algoritma *Naïve Bayes*.

Algoritma *Naïve Bayes* merupakan teknik dalam *machine learning* yang menggunakan pendekatan probabilistik untuk melakukan analisis data dan klasifikasi. Algoritma ini memiliki keunggulan dalam kecepatan dan efisiensi, khususnya dalam menangani dataset yang besar. Dengan algoritma ini, koperasi dapat melakukan analisis penting, seperti mengklasifikasikan produk-produk mana yang memiliki tingkat penjualan tinggi (laris) dan yang kurang diminati. Selain itu, penerapan algoritma ini memungkinkan koperasi untuk mengoptimalkan pengelolaan stok barang, menetapkan harga yang lebih kompetitif, serta menawarkan produk yang sesuai dengan kebutuhan konsumen.

Berdasarkan permasalahan dan potensi yang telah diuraikan, penelitian ini bertujuan untuk mengklasifikasi data penjualan di Koperasi Manbaul

Ihsan Al-Baedlowi menggunakan algoritma *Naïve Bayes*. Penelitian ini diharapkan dapat menghasilkan model klasifikasi yang memberikan wawasan mengenai produk-produk dengan tingkat penjualan laris maupun kurang laris. Dengan wawasan ini, koperasi dapat mengambil keputusan yang lebih tepat dan strategis dalam pengelolaan usaha, sehingga meningkatkan efisiensi dan efektivitas operasionalnya serta memberikan manfaat yang lebih besar bagi para santri dan masyarakat sekitar.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining adalah proses penjelajahan data dalam *database* untuk mengidentifikasi bentuk atau hubungan yang signifikan dan bermakna [1]. *Data Mining* mengacu pada serangkaian teknik yang dipakai dalam mengidentifikasi pola-pola yang belum diketahui dalam data yang sudah terkumpul [2]. *Data Mining* adalah proses yang memanfaatkan teknik dari berbagai disiplin ilmu seperti matematika, statistika, *artificial intelligence*, dan *machine learning* untuk menggali serta mengidentifikasi data yang relevan dan informasi penting dari basis data berukuran besar [3].

2.2. Klasifikasi

Klasifikasi adalah proses menempatkan data yang baru ke dalam kategori atau kelas yang sudah ditetapkan sebelumnya [4]. Proses klasifikasi melibatkan penggunaan variabel kategori untuk menentukan target klasifikasi. Misalnya, dalam konteks klasifikasi keuntungan, data dapat dikelompokkan ke dalam tiga kategori, yaitu kecil, menengah, dan besar [5]. Klasifikasi adalah teknik untuk mengkategorikan atau mengelompokkan data berdasarkan kelas atau label yang dimilikinya. Algoritma yang digunakan dalam klasifikasi termasuk dalam kategori *supervised learning*. Teknik ini juga dapat dipakai untuk memprediksi kecenderungan atau pola data di masa depan [6].

2.3. Naïve Bayes

Naïve Bayes atau *Naïve Bayes Classifier* adalah metode yang efektif untuk klasifikasi dua kelas maupun lebih. Metode ini menggunakan teknik klasifikasi yang terawasi untuk memprediksi kategori suatu hal di masa depan dengan memberikan label kelas pada setiap *instance* atau catatan berdasarkan probabilitas bersyarat. [7]. *Naïve Bayes* adalah algoritma yang populer dalam *data mining* karena kemampuannya untuk memberikan hasil klasifikasi yang efektif, meskipun tidak ada hubungan langsung antara atribut-atribut yang ada. Algoritma ini tetap menunjukkan kinerja yang baik dan kompetitif dalam memproses klasifikasi. Metode *Naïve Bayes* dapat diterapkan secara efektif pada sistem rekomendasi. Kesederhanaan dan efisiensinya memungkinkannya untuk melakukan penskalaan yang baik dengan kumpulan data yang besar, menjadikannya pilihan yang tepat untuk rekomendasi [8].

Dalam konteks klasifikasi, algoritma *Naïve Bayes* mengklasifikasikan data berdasarkan probabilitas kondisi:

Di mana:

$$P(C|X) = \frac{P(X|C).P(C)}{P(X)} \quad (1)$$

- $P(C|X)$ merupakan probabilitas kelas C diberikan data X.
- $P(X|C)$ merupakan probabilitas data X diberikan kelas C.
- $P(C)$ merupakan probabilitas apriori kelas C.
- $P(X)$ merupakan probabilitas apriori data X.

2.4. RapidMiner

RapidMiner adalah perangkat lunak *open source* atau dapat digunakan oleh siapa saja. Perangkat lunak ini digunakan sebagai solusi untuk menganalisis pengolahan data, dengan menerapkan berbagai teknik seperti analisis deskriptif dan prediktif [9]. *RapidMiner* adalah perangkat lunak yang dirancang khusus untuk keperluan *data mining*. Perangkat lunak ini menyediakan berbagai model yang cukup banyak, seperti *Model Bayesian*, *Tree Induction*, *Neural Network*, *Naïve Bayes*, dan sebagainya. *RapidMiner* juga mendukung beragam metode, termasuk klasifikasi, klastering, dan asosiasi [10].

2.5. Confusion Matrix

Confusion Matrix merupakan matriks yang digunakan dalam mengevaluasi tingkat akurasi pada proses *data mining*. Matriks ini memiliki dua kelas, yaitu kelas positif dan kelas negatif, yang masing-masing memiliki empat elemen yaitu:

Tabel 1. *Confusion Matrix*

Predicted Values	Actual Values	
	Positive	Negative
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

- *True Positive* (TP): Jumlah prediksi positif yang benar.
- *True Negative* (TN): Jumlah prediksi negatif yang benar.
- *False Positive* (FP): Jumlah prediksi positif yang salah.
- *False Negative* (FN): Jumlah prediksi negatif yang salah.

Akurasi dihitung dengan membandingkan jumlah data yang berhasil diklasifikasikan dengan benar terhadap total jumlah sampel data yang diuji [11]. Selain *Accuracy*, *Confusion Matrix* juga mengukur metrik lainnya seperti *Precision*, dan *Recall*. Dengan rumus sebagai berikut [12].

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (3)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (4)$$

3. METODE PENELITIAN

Metodologi penelitian yang dipakai yaitu *Knowledge Discovery in Database* (KDD). KDD adalah proses menemukan informasi berguna dari database besar yang dianalisis untuk memaksimalkan fungsinya sebagai tempat penyimpanan data yang bermanfaat [13]. Tahapan pada metodologi KDD yaitu *data selection*, *data preprocessing*, *transformation*, *data mining*, *evaluation*. Data yang digunakan adalah data penjualan Alat Tulis Kantor dari Koperasi Al-Baedlowi.

Penelitian dilakukan pada November 2023 hingga Maret 2024 di Koperasi Al-Baedlowi, yang beralamat di Desa Karangpucung, RT 03 RW 02, Kecamatan Kertanegara, Kabupaten Purbalingga. Waktu penelitian dibagi ke dalam tiga tahap pelaksanaan, yaitu: Tahap persiapan mencakup pengajuan judul, penyusunan proposal, penyusunan instrumen penelitian, pembuatan rancangan *data mining* pada alat tulis kantor di Koperasi Al-Baedlowi, dan seminar proposal. Tahap penelitian melibatkan pemilihan data yang diperoleh dari wawancara untuk diproses lebih lanjut dalam pemilihan atribut yang akan digunakan dalam *preprocessing*. Tahap penyelesaian meliputi penyusunan laporan hasil penelitian.

Untuk mengatasi masalah yang berkaitan dengan penerapan algoritma *Naïve Bayes*, dilakukan langkah-langkah dalam *Knowledge Discovery in Database* (KDD) [14]:

a. Data Selection

Memilih data yang relevan dari hasil wawancara langsung dengan pihak Koperasi Al-Baedlowi untuk digunakan dalam *preprocessing*.

b. Data Preprocessing

Data yang masih mentah diubah menjadi data berkualitas yang siap untuk diproses pada tahap selanjutnya. Proses ini terdiri dari beberapa langkah seperti membersihkan data dengan menghapus atau mengganti nilai yang hilang (*missing values*), memperbaiki data yang tidak sesuai, serta melakukan normalisasi atau standarisasi data sesuai kebutuhan.

c. Transformation

Mempersiapkan data dalam format yang sesuai dengan kebutuhan penelitian. Proses dimulai dengan mengimpor data ke dalam proyek baru menggunakan berbagai format seperti CSV, Excel, atau database. Setelah itu, dilakukan pratinjau data untuk memahami struktur dan kontennya, termasuk pemeriksaan atribut serta tipe data. Selanjutnya, dilakukan seleksi fitur untuk memilih atribut yang relevan bagi model *Naïve Bayes*, dengan memanfaatkan operator seperti *Select Attributes* untuk memastikan hanya fitur penting yang digunakan dalam analisis.

d. Data Mining

Data yang telah diproses dipakai untuk melakukan proses klasifikasi data penjualan dengan algoritma *Naïve Bayes* pada aplikasi *RapidMiner*. Proses ini melibatkan pembagian data menjadi dua set, yaitu data pelatihan (*training data*) dan data pengujian

(*testing data*). Model *Naïve Bayes* kemudian dilatih menggunakan data latih untuk melakukan klasifikasi produk berdasarkan pola yang teridentifikasi dari data tersebut.

e. Evaluation

Evaluasi dilakukan untuk mengukur validitas model klasifikasi menggunakan *confusion matrix*. Dengan hasil yaitu nilai *accuracy*, *precision*, dan *recall*.

Selain *confusion matrix*, dari hasil pengujian juga akan didapatkan nilai *AUC* yang kemudian dikategorikan sesuai tingkatannya seperti pada Tabel 2 [5].

Tabel 2. Kategori Nilai AUC

Nilai AUC	Kategori Klasifikasi
0.90 – 1.00	<i>Excellent Classification</i>
0.80 – 0.90	<i>Good Classification</i>
0.70 – 0.80	<i>Fair Classification</i>
0.60 – 0.70	<i>Poor Classification</i>
0.50 – 0.60	<i>Failure</i>

4. HASIL DAN PEMBAHASAN

Penelitian ini dimulai dengan tahap pengumpulan data dengan melakukan proses wawancara untuk mendapatkan data penjualan produk di Koperasi Al-Baedlowi, yang beralamat di Desa Karangpucung, RT 03 RW 02, Kecamatan Kertanegara, Kabupaten Purbalingga. Data yang didapatkan yaitu data penjualan produk pada bulan November – Desember tahun 2023.

4.1. Data Selection

Data yang telah didapatkan dilakukan proses seleksi dengan menghapus data yang tidak lengkap dan sebagainya dengan menyisakan data sebanyak 50 macam produk yang diubah menjadi bentuk Excel seperti pada Tabel 3.

Tabel 3. Data Penjualan Produk

No	Nama Barang	Jumlah Terjual	Stok Barang
1	Bolpoin H Tech C Biru	20	24
2	Bolpoin H Tech C Hitam	24	24
3	Bolpoin H Tech H	48	60
4	Bolpoin Joyko	12	24
5	Bolpoin Safari Brand	36	36
6	Bolpoin Standar	15	36
7	Buku Fiqih Wanita	3	5
8	Buku Gambar	15	20
9	Buku Kwarto	35	50
10	Buku Risalah Hayawan	2	5

4.2. Data Preprocessing

Data penjualan selanjutnya dilakukan pra-pemrosesan dengan menambahkan kolom Keterangan

yang dijadikan sebagai label dengan parameter jika Jumlah Terjual lebih dari 80% dari Stok Produk maka

Keterangan berisi “Laris” dan jika kurang maka berisi “Kurang Laris” seperti pada Tabel 4.

Tabel 4. Data Penjualan Produk dengan Keterangan

No	Nama Barang	Jumlah Terjual	Stok Barang	Keterangan
1	Bolpoin H Tech C Biru	20	24	Laris
2	Bolpoin H Tech C Hitam	24	24	Laris
3	Bolpoin H Tech H	48	60	Laris
4	Bolpoin Joyko	12	24	Kurang Laris
5	Bolpoin Safari Brand	36	36	Laris
6	Bolpoin Standar	15	36	Kurang Laris
7	Buku Fiqih Wanita	3	5	Kurang Laris
8	Buku Gambar	15	20	Kurang Laris
9	Buku Kwarto	35	50	Kurang Laris
10	Buku Risalah Hayawan	2	5	Kurang Laris

4.3. Transformation

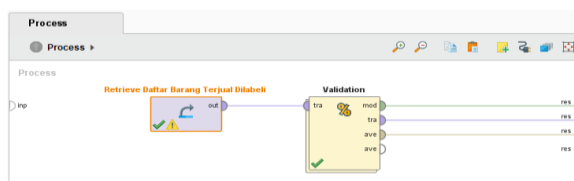
Pada tahap ini, setelah data sudah siap maka langkah selanjutnya dilakukan proses *input data* pada *RapidMiner* yang di mana Nama Produk dijadikan sebagai id dan Keterangan sebagai label seperti pada Gambar 1.

Row No.	Nama Barang	Keterangan	Jumlah Terjual	Stok Barang
1	Bolpoin H Te...	Laris	20	24
2	Bolpoin H Te...	Laris	24	24
3	Bolpoin H Te...	Laris	48	60
4	Bolpoin Joyko	Kurang Laris	12	24
5	Bolpoin Safar...	Laris	36	36
6	Bolpoin Stan...	Kurang Laris	15	36
7	Buku Fiqih W...	Kurang Laris	3	5
8	Buku Gambar	Kurang Laris	15	20
9	Buku Kwarto	Kurang Laris	35	50
10	Buku Risalah...	Kurang Laris	2	5

Gambar 1. Data Penjualan Produk di *RapidMiner*

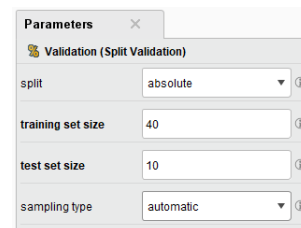
4.4. Data Mining

Tahap selanjutnya yaitu melakukan proses *data mining* untuk mengklasifikasi data penjualan menggunakan aplikasi *RapidMiner*. Seperti pada Gambar 2 proses klasifikasi ini menggunakan data penjualan yang telah di-*input* sebelumnya yang kemudian dihubungkan pada operator *Validation* yang digunakan untuk proses validasi model.



Gambar 2. Proses Klasifikasi

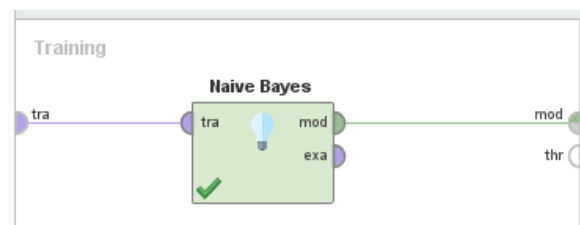
Sebagai pada Gambar 3 pada operator *Validation* dilakukan perubahan pada parameter dengan mengubah *split* yang berfungsi membagi data menjadi dua yaitu *training* dan *test* yang menggunakan *absolute*, *training set size* atau jumlah data pelatihan yang berjumlah 40, *test set size* atau jumlah data pengujian yang berjumlah 10, dan *sampling type* menggunakan *automatic*.



Gambar 3. Parameter Klasifikasi

Di dalam operator *Validation*, terdapat dua bagian utama yang saling melengkapi, yaitu *Training* dan *Testing*, yang masing-masing berperan penting dalam membangun dan mengevaluasi model *Naïve Bayes*. Bagian ini dirancang untuk memastikan bahwa model tidak hanya dilatih dengan baik tetapi juga diuji secara komprehensif guna menilai kinerjanya pada data baru.

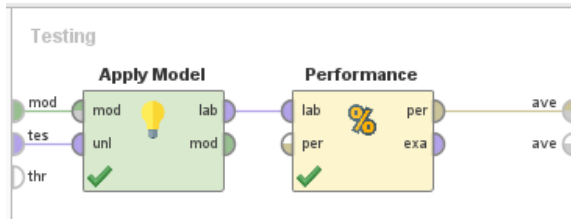
Pada bagian *Training* seperti yang ditampilkan pada Gambar 4, operator *Naïve Bayes* ditambahkan untuk melakukan proses pelatihan model. Operator ini menggunakan data pelatihan (*training data*) untuk mempelajari pola-pola atau hubungan di antara fitur-fitur dalam *dataset*. *Output* dari proses ini adalah model yang telah dilatih, yang mencerminkan hasil pembelajaran dari data pelatihan. Model ini nantinya akan diaplikasikan pada data uji untuk memprediksi hasil dan mengevaluasi performanya.



Gambar 4. Proses Training

Pada bagian *Testing* seperti yang ditampilkan pada Gambar 5, terdapat dua operator utama yang digunakan untuk mengaplikasikan dan mengevaluasi model yaitu *Apply Model* yang bertugas mengaplikasikan model yang telah dilatih pada bagian *Training* ke data uji (*testing data*) dan *Performance* yang berfungsi untuk mengevaluasi kinerja model

berdasarkan hasil prediksi dari *Apply Model*. Evaluasi dilakukan dengan membandingkan label prediksi dengan label aktual dari data uji. Hasil dari evaluasi ini disajikan dalam bentuk *confusion matrix*, yang mencakup metrik-metrik penting seperti *accuracy*, *precision*, *recall*, dan metrik lainnya yang relevan.



Gambar 5. Proses Testing

4.5. Evaluation

Dari tahap evaluasi model didapatkan hasil berupa *confusion matrix* yaitu *accuracy*, *precision*, dan *recall*. Hasil dari model klasifikasi *Naïve Bayes* pada data penjualan produk yaitu sebagai berikut.

a. Accuracy

Accuracy sebesar 80.00% seperti pada Gambar 6, di mana dari sebanyak 10 data pengujian terbagi menjadi sejumlah 4 untuk data berlabel Laris dengan prediksi dari model yaitu 3 benar dan 1 salah, dan sejumlah 6 untuk data berlabel Kurang Laris dengan prediksi yaitu 5 benar dan 1 salah.

accuracy: 80.00%

	true Laris	true Kurang Laris	class precision
pred. Laris	3	1	75.00%
pred. Kurang Laris	1	5	83.33%
class recall	75.00%	83.33%	

Gambar 6. Hasil Accuracy

Atau dengan perhitungan manual seperti berikut.

$$\text{Accuracy} = \frac{(3 + 5)}{(3 + 5 + 1 + 1)} \times 100\% = 80.00\%$$

b. Precision

Precision mendapatkan sebesar 83.33% seperti pada Gambar 7 yang masing-masing kelasnya mendapatkan 75.00% untuk kelas Laris dan 83.33% untuk data Kurang Laris.

precision: 83.33% (positive class: Kurang Laris)

	true Laris	true Kurang Laris	class precision
pred. Laris	3	1	75.00%
pred. Kurang Laris	1	5	83.33%
class recall	75.00%	83.33%	

Gambar 7. Hasil Precision

Atau dengan perhitungan manual untuk *precision* kelas positif Kurang Laris seperti berikut.

$$\text{Precision} = \frac{5}{(5 + 1)} \times 100\% = 83.33\%$$

c. Recall

Recall mendapatkan sebesar 83.33% seperti pada Gambar 8 dengan masing-masing kelas mendapatkan 75.00% untuk kelas Laris dan 83.33% untuk kelas Kurang Laris.

recall: 83.33% (positive class: Kurang Laris)

	true Laris	true Kurang Laris	class precision
pred. Laris	3	1	75.00%
pred. Kurang Laris	1	5	83.33%
class recall	75.00%	83.33%	

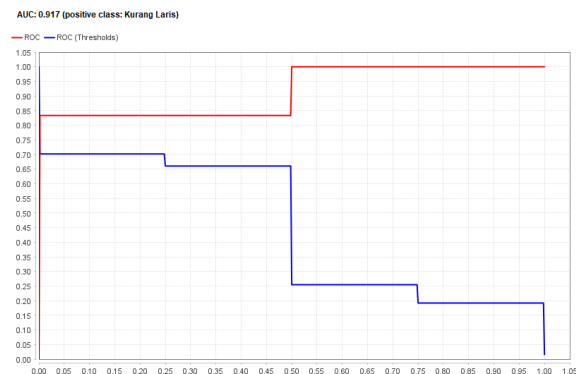
Gambar 8. Hasil Recall

Atau dengan perhitungan manual untuk *recall* kelas positif Kurang Laris seperti berikut.

$$\text{Recall} = \frac{5}{(5 + 1)} \times 100\% = 83.33\%$$

d. AUC

Dari hasil pengujian juga mendapatkan nilai *AUC* seperti pada Gambar 9.



Gambar 9. Nilai AUC dan Diagram ROC

Nilai *AUC* di atas yang didapatkan dari proses pengujian model klasifikasi data penjualan produk di Koperasi Al-Baedlowi menggunakan algoritma *Naïve Bayes* mendapat nilai sebesar 0.917 untuk kelas positif Kurang Laris yang berarti model yang telah dibuat berdasarkan kriteria nilai *AUC* dapat dikategorikan sebagai klasifikasi yang sangat baik (*Excellent Classification*).

e. PerformanceVector

Serta dari proses evaluasi model diberikan keterangan *PerformanceVector* yang dimana mendapatkan *accuracy* sebesar 80.00%, *precision* sebesar 83.33%, *recall* sebesar 83.33%, *AUC* (*optimistic*) sebesar 0.917, *AUC* sebesar 0.917, dan *AUC* (*pessimistic*) sebesar 0.917 dengan semuanya memiliki *positive class*: Kurang Laris seperti pada Gambar 10.

PerformanceVector

```
PerformanceVector:
accuracy: 80.00%
ConfusionMatrix:
True:  Laris  Kurang Laris
Laris: 3      1
Kurang Laris: 1      5
precision: 83.33% (positive class: Kurang Laris)
ConfusionMatrix:
True:  Laris  Kurang Laris
Laris: 3      1
Kurang Laris: 1      5
recall: 83.33% (positive class: Kurang Laris)
ConfusionMatrix:
True:  Laris  Kurang Laris
Laris: 3      1
Kurang Laris: 1      5
AUC (optimistic): 0.917 (positive class: Kurang Laris)
AUC: 0.917 (positive class: Kurang Laris)
AUC (pessimistic): 0.917 (positive class: Kurang Laris)
```

Gambar 10. PerformanceVector

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian dalam klasifikasi data penjualan produk di Koperasi Al-Baedlowi menggunakan algoritma *Naïve Bayes* dengan metode penelitian *Knowledge Discovery in Database (KDD)* didapatkan bahwa hasil pengujian model mendapat *accuracy* sebesar 80.00%, *precision* kelas positif Kurang Laris sebesar 83.33%, dan *recall* kelas positif Kurang Laris sebesar 83.33%, di mana dari sebanyak 10 data pengujian terbagi menjadi sejumlah 4 untuk data berlabel Laris dengan prediksi dari model yaitu 3 benar dan 1 salah, dan sejumlah 6 untuk data berlabel Kurang Laris dengan prediksi yaitu 5 benar dan 1 salah. Selain itu didapatkan nilai *AUC* sebesar 0.917 untuk kelas positif Kurang Laris yang berarti model yang telah dibuat berdasarkan kriteria nilai *AUC* dapat dikategorikan sebagai klasifikasi yang sangat baik (*Excelent Classification*). Saran untuk penelitian selanjutnya yaitu dapat memperbanyak jumlah data yang digunakan untuk penelitian, mengubah parameter algoritma, atau menambah kompleksitas pada proses *data mining* untuk mendapatkan hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] M. F. Rozi, "Penerapan Data Mining Menggunakan Metode Naive Bayes Untuk Klasifikasi Data Penentuan Hasil Penjualan Dalam Strategi Pemasaran," *J. Komput. Teknol. Inf. dan Sist. Inf.*, vol. 2, no. 2, pp. 444–454, 2023, doi: 10.62712/juktisi.v2i2.137.
- [2] Alvina Felicia Watratan, Arwini Puspita. B, and Dikwan Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 7–14, Jul. 2020, doi: 10.52158/jacost.v1i1.9.
- [3] Y. S. T. Allo, V. Sofica, N. Hasan, and M. Septiani, "Penggunaan Metode Naïve Bayes Dalam Mengklasifikasi Pengangguran Pada Desa Bojong Kulur," *Bianglala Inform.*, vol. 10, no. 1, pp. 30–35, 2022, doi: 10.31294/bi.v10i1.12333.
- [4] S. Lestari, A. Akmaludin, and M. Badrul, "Implementasi Klasifikasi Naive Bayes Untuk Prediksi Kelayakan Pemberian Pinjaman Pada Koperasi Anugerah Bintang Cemerlang," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 7, no. 1, pp. 8–16, 2020, doi: 10.30656/prosisko.v7i1.2129.
- [5] A. Julianto and S. Andayani, "Penerapan Data Mining Untuk Klasifikasi Produk Terlaris Menggunakan Algoritma Naive Bayes Pada Bengkel Motor," *JuSiTik J. Sist. dan Teknol. Inf. Komun.*, vol. 7, no. 2, pp. 50–58, 2024, doi: 10.32524/jusitik.v7i2.1148.
- [6] I. M. B. Adnyana, "Implementasi Naïve Bayes Untuk Memprediksi Waktu Tunggu Alumni Dalam Memperoleh Pekerjaan," *Semin. Nas. Teknol. Komput. Sains*, vol. 1, no. 1, pp. 131–134, 2020, [Online]. Available: <http://prosiding.seminar-id.com/index.php/sainteks/article/view/418>
- [7] L. Febriyanti and H. Zakaria, "Implementasi Data Mining Untuk Memprediksi Produktivitas Pada Tanaman Kacang Tanah Menggunakan Metode Naive Bayes (Studi Kasus : Perkebunan Kacang Tanah Di Kota Bogor)," *Log. J. Ilmu Komput. dan Pendidik.*, vol. 1, no. 2, pp. 105–118, 2023, [Online]. Available: <https://journal.mediapublikasi.id/index.php/logic/article/view/2353>
- [8] R. P. V and P. M, "Intelligent deep learning based bidirectional long short term memory model for automated reply of e-mail client prototype," *Pattern Recognit. Lett.*, vol. 152, pp. 340–347, Dec. 2021, doi: 10.1016/j.patrec.2021.10.021.
- [9] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.
- [10] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [11] N. S. Fauziah and R. D. Dana, "Implementasi Algoritma Naive bayes dalam Klasifikasi Status Kesejahteraan Masyarakat Desa Gunungsari," *Blend Sains J. Tek.*, vol. 1, no. 4, pp. 295–305, 2023, doi: 10.56211/blendsains.v1i4.234.
- [12] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [13] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 20–33, 2023, doi: 10.56211/helloworld.v2i1.230.
- [14] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.