

---

## PERBANDINGAN METODE *DECISION TREE* DAN *NAIVE BAYES* PADA TINGKAT PENJUALAN MINUMAN KOPI DI KOPI PAWON NUSANTARA

Muhamad Iqbal<sup>1</sup>, Siti Miskiyah<sup>2</sup>, Said Luthfil Sham<sup>3</sup>, Syaiful Anwar<sup>4</sup>, Mohammad Haddiel Fuad<sup>5</sup>

<sup>1,2,3,4,5</sup>Universitas Bina Sarana Informatika

Jl. Kramat Raya No.98, Senen, Jakarta Pusat, Indonesia

e-mail: <sup>1</sup>19231732@bsi.ac.id, <sup>2</sup>19232393@bsi.ac.id, <sup>3</sup>19231479@bsi.ac.id, <sup>4</sup>Syaiful.sfa@bsi.ac.id,

<sup>5</sup>mhfud@gmail.com

---

Artikel Info : Diterima : 22-06-2024 | Direvisi : 19-07-2024 | Disetujui : 12-08-2024

---

**Abstrak** - Tujuan dari penelitian ini adalah untuk membandingkan dua model klasifikasi, yaitu *Decision Tree* dan *Naive Bayes*, dalam konteks meningkatkan penjualan di Kopi Pawon Nusantara. Penelitian ini menggunakan data penjualan minuman kopi untuk memahami pola pembelian pelanggan. Pemilik Kopi Pawon Nusantara menghadapi tantangan dalam menentukan minuman kopi yang paling diminati oleh pelanggan mereka. Data yang digunakan berasal dari penjualan sebelumnya di Kopi Pawon Nusantara. Dua algoritma, *Decision Tree* dan *Naive Bayes*, dievaluasi untuk mengukur performanya menggunakan metrik seperti akurasi, presisi, *recall*, dan *area under curve* (AUC) dari kurva ROC. Hasil *Decision Tree* mencapai akurasi sebesar 97,83% dengan AUC 0,957. Sementara itu, *Naive Bayes* memiliki akurasi sebesar 94,35% dengan AUC 0,969. Tujuan dari penelitian ini adalah untuk mendapatkan pemahaman tentang metode klasifikasi mana yang paling cocok untuk menangani data penjualan Kopi Pawon Nusantara. Hasilnya diharapkan dapat digunakan sebagai dasar untuk penelitian lanjutan yang akan memilih metode klasifikasi yang tepat, baik dengan dataset yang sama maupun dengan metode klasifikasi yang berbeda.

Kata Kunci : *Decision Tree*, Minuman kopi, *Naive Bayes*, Penjualan

**Abstracts** - The aim of this study is to compare two classification models, *Decision Tree* and *Naive Bayes*, to enhance sales at Kopi Pawon Nusantara. Using coffee sales data, the research seeks to understand customer purchasing behavior. Owners of Kopi Pawon Nusantara are challenged in identifying the most favored coffee drinks by their customers. Data is sourced from prior sales at Kopi Pawon Nusantara. Both *Decision Tree* and *Naive Bayes* algorithms are assessed for their performance using metrics like accuracy, precision, recall, and area under the ROC curve (AUC). The *Decision Tree* achieved 97.83% accuracy with an AUC of 0.957, while *Naive Bayes* showed 94.35% accuracy with an AUC of 0.969. The objective is to determine the most appropriate classification method for analyzing Kopi Pawon Nusantara's sales data, providing a foundation for future studies exploring different methods or datasets.

Keywords : *Coffee Beverages*, *Decision Tree*, *Naive Bayes*, *Sales*

## PENDAHULUAN

Dalam era digital ini, data telah menjadi aset penting untuk pengambilan keputusan bisnis. Kopi Pawon Nusantara, sebagai salah satu pelaku industri kopi, menghadapi tantangan untuk meningkatkan penjualan di pasar yang sangat kompetitif. Dengan adanya beragam faktor yang mempengaruhi keputusan pembelian konsumen, perusahaan harus menganalisis data penjualan secara efektif guna mengidentifikasi pola dan tren yang dapat mendukung strategi pemasaran dan penjualan mereka.

Dihini mengemukakan “ Industri kopi Indonesia berkembang pesat , menurut data badan pusat statistik , nilai ekspor kopi indonesia pada tahun 2020 sebesar US\$ 809,2 (Bei & Jamu, 2022).



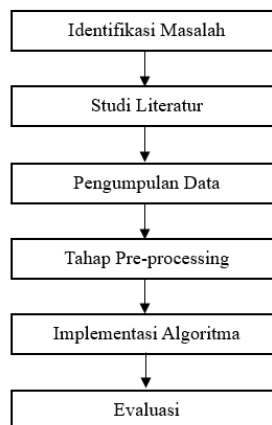
Kopi merupakan salah satu produk pertanian yang banyak ditanam di Indonesia dan sangat populer sebagai minuman. Provinsi Sulawesi Selatan merupakan salah satu wilayah di Indonesia yang terkenal sebagai penghasil kopi. Di provinsi ini, ada beberapa kabupaten yang terkenal sebagai penghasil kopi, seperti Enrekang, Toraja Utara, Sinjai, dan Bantaeng. Saat ini, kehadiran kafe atau kedai kopi sudah menjadi hal umum di berbagai tempat dan wilayah, terutama di era milenial ini. Minat yang tinggi terhadap kopi menjadi dorongan bagi para pemilik warung kopi atau kedai kopi untuk berinovasi dalam penyajian dan pengelolaan minuman kopi mereka. Setiap kedai kopi berusaha menciptakan varian minuman kopi yang unik sebagai ciri khas mereka. Sebagai contoh, Konijawa, sebuah kedai kopi di Kabupaten Bantaeng, menawarkan beragam varian minuman kopi. Menu mereka terbagi menjadi dua kategori, yaitu minuman berbasis kopi dan minuman kopi unggulan. Minuman berbasis kopi mencakup berbagai macam seperti *affogato*, *americano*, hingga *espresso*, sedangkan minuman kopi unggulan mencakup varian seperti *avocado coffee*, kopi susu klasik, dan kopi susu pandan (Suci Amaliah et al., 2022).

Kopi Pawon Nusantara adalah tempat mengopi yang berada dipinggir jalan namun dengan rasa tidak kalah enak dengan kopi yang ada di kafe-kafe. Kopi Pawon Nusantara sendiri adalah merk dagang yang terbentuk karena animo para pelanggan setia penikmat kopi. Kopi Pawon Nusantara mempunyai visi & misi melestarikan kopi rasa nusantara yang dapat di nikmati seluruh kalangan dengan harga terjangkau rasa bintang lima dengan jargon "Kopi Nikmat Harga Hemat". Menurut manajernya Kopi Pawon Hadir ditengah-tengah masyarakat Indonesia dengan membawa misi originalitas kopi Nusantara. Terbukti bahwa Kopi Pawon Nusantara baru merintis selama 5 tahun sejak 10 Juni 2018 dan telah memiliki 52 outlet di berbagai tempat di pulau jawa.

Kopi Pawon Nusantara berada di Jl. Sarang Bangau Komplek PASMAR I No.17 ,Cilincing, Jakarta Utara. Kopi Pawon Nusantara sendiri mulai buka pada jam 19.30 sampai dengan kurang lebih jam 01.00 dini hari. Untuk harganya sendiri tidak terlalu merogoh kantong karena kopi di Kopi Pawon harganya mulai dari Rp.10.000. Sebagai bisnis di industri kuliner, Kopi Pawon Nusantara harus terus berinovasi agar tetap kompetitif di pasar yang semakin padat. Salah satu strategi untuk meningkatkan penjualan adalah dengan memanfaatkan data yang ada untuk memahami pola pembelian pelanggan. Selain itu, Pemilik Kopi Pawon Nusantara mengalami kesulitan dalam menentukan jenis minuman kopi yang populer dan yang kurang diminati. Penelitian ini bertujuan untuk membandingkan dua metode klasifikasi, yaitu *decision tree* dan *naive bayes*, guna menganalisis tingkat keakuratan dalam menentukan minuman kopi yang populer dan yang tidak.

## METODE PENELITIAN

Pada bagian ini, penulis akan memaparkan proses penelitian langkah demi langkah, yang dimulai dengan mengidentifikasi masalah dan diakhiri dengan evaluasi.



Sumber : Penelitian Mandiri (2024)

Gambar 1. Kerangka Penelitian

### 1. Identifikasi Masalah

Menentukan masalah penelitian yang akan dipecahkan adalah membandingkan seberapa efektif algoritma *decision tree* dan *naive bayes* dalam memprediksi jumlah penjualan minuman di Kopi Pawon Nusantara.

### 2. Studi Literatur

Penelitian ini melakukan tinjauan literatur dengan menghimpun informasi dari berbagai sumber referensi, termasuk literatur akademis dan jurnal penelitian terkait. Data tersebut meliputi landasan teoritis mengenai metodologi yang akan digunakan dalam meramal tingkat penjualan minuman kopi, serta hasil temuan dari penelitian-penelitian terdahulu.

### 3. Pengumpulan Data

Kumpulkan informasi mengenai penjualan berbagai jenis minuman kopi di Kopi Pawon Nusantara, termasuk rincian seperti nama menu, jumlah penjualan, harga, tanggal penjualan, dan faktor-faktor lain yang relevan. Selain itu, cari juga data tentang atribut yang mungkin memengaruhi tingkat penjualan minuman kopi, seperti promosi yang dilakukan dan lokasi dari Kopi Pawon Nusantara.

### 4. Tahap Pre-processing

*Pre-processing* adalah tahap penting dalam pengolahan data yang membantu meningkatkan dan membersihkan data sehingga hasil evaluasinya lebih optimal. Beberapa proses selesai pada saat ini, termasuk memilih dan mengkonfigurasi atribut dan normalisasi data., serta penanganan nilai yang hilang. Langkah awal yang penting adalah memeriksa nilai yang hilang dalam dataset yang digunakan. Selanjutnya, operator split data digunakan untuk memastikan validitas data dengan membagi data menjadi data pelatihan dan pengujian. Ini dilakukan menggunakan operator *split* data dalam *RapidMiner*. Setelah tahap ini, data siap untuk diproses dalam proses pemodelan (Chandra Wirawan, 2020).

### 5. Implementasi Algoritma

Mengimplementasikan algoritma *decision tree* dan *naive bayes* menggunakan *software* analisis data yang menyediakan kedua algoritma tersebut.

#### *Decision Tree*

*Decision tree* merupakan struktur diagram alur yang menyerupai bentuk pohon, yang bertujuan untuk memecah kumpulan data besar menjadi kelompok data yang lebih kecil dengan menerapkan aturan keputusan. Setiap kelompok anggota kemudian dikelompokkan berdasarkan kesamaannya. Dalam pohon keputusan, atribut-atribut dibagi menjadi node untuk klasifikasi sesuai dengan label yang telah ditentukan. Algoritma pembentukan pohon keputusan C4.5 terdiri dari dua tahapan utama, yaitu pencarian root dan proses pembuatan cabang. Pencarian root dilakukan dengan mempertimbangkan setiap atribut dari data latih yang telah diberi label (Sulastri, 2022). *Decision Tree* adalah teknik klasifikasi yang membangun model dengan menggunakan informasi dari entropy. Model ini dibangun dengan menghitung nilai entropy dan gain dari sampel data yang digunakan. Entropy adalah sebuah indikator statistik yang mengukur tingkat ketidakpastian dalam data. Sementara gain merupakan ukuran yang menilai efektivitas suatu atribut dalam membagi contoh pelatihan sesuai dengan kelas target (Fihir et al., 2023). Langkah-langkah untuk menghitung nilai entropy dan gain juga disertakan :

1. *Entropy*

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$

Keterangan

S = Himpunan kasus

n = Jumlah partisi A

pi = proporsi Si terhadap S

2. *Gain*

$$Gain(S, A) = Entropy(S) - \sum_i^n = 1 \frac{|S_i|}{|S|} x Entropy(S_i) \quad (2)$$

Keterangan

S = Himpunan kasus

A = Jumlah Partisi A

|Si| = Proporsi Si terhadap S

|S| = Jumlah kasus dalam S

#### *Naive Bayes*

Algoritma klasifikasi *Naive Bayes* adalah suatu metode sederhana dalam pengklasifikasi yang mengandalkan *teorema Bayes*, menekankan independensi yang kuat antara fitur-fitur yang digunakan. Dengan

kata lain, dalam *Naive Bayesian*, asumsi bahwa setiap fitur berdiri sendiri tanpa ketergantungan pada fitur lain dalam data. Keunggulan utamanya adalah kemampuannya untuk melakukan klasifikasi dengan membutuhkan hanya sejumlah kecil data pelatihan untuk mengestimasi parameter yang dibutuhkan, seperti mean dan variance variabel, untuk melakukan klasifikasi (Harahap et al., 2021). Algoritma *Naive Bayes* adalah suatu prosedur yang menggunakan probabilitas dan statistik untuk menyelesaikan masalah klasifikasi. Cara kerjanya melibatkan perhitungan nilai probabilitas  $P(x/y)$  dengan mengenali probabilitas kelas  $X$  untuk klasifikasi. Penentuan kelas dalam klasifikasi dilakukan secara probabilistik dengan memilih nilai maksimum  $P(x/y)$ . Salah satu keunggulannya adalah bahwa klasifikasi ini memerlukan jumlah informasi pelatihan yang relatif kecil untuk mengestimasi parameter yang diperlukan untuk klasifikasi (Indriyani Indriyani & Agus Bahtiar, 2023). Menurut algoritma *Naive Bayes*, formula yang digunakan untuk menentukan nilai  $P(x/y)$  adalah seperti yang berikut ini :

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \quad (3)$$

## 6. Evaluasi

Evaluasi kinerja setiap algoritma pada data uji dapat dilakukan dengan menggunakan metrik seperti akurasi, presisi, *recall*, atau *area under curve* (AUC) .

Selanjutnya, *ROC (Receiver Operating Characteristics)* digunakan untuk menampilkan performa dua dimensi dari setiap klasifikasi yang diuji. Sumbu horizontal menggambarkan hasil positif yang salah, sedangkan sumbu vertikal menggambarkan hasil positif yang benar. *AUC* (Area Under Curve) merupakan area yang terletak di bawah kurva *ROC*. Kualitas klasifikasi dapat dinilai berdasarkan nilai *AUC*-nya, dan tabel berikut menunjukkan kualitas klasifikasi yang sesuai dengan nilai *AUC* tersebut.

Tabel 1. Standar AUC

Nilai AUC	Penjelasan
0,90-1,00	<i>Excellent</i>
0,80-0,90	<i>Good</i>
0,70-0,80	<i>Fair</i>
0,60-0,70	<i>Poor</i>
0,50-0,60	<i>Failure</i>

Sumber : Penelitian Mandiri (2024)

## HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan Data

Data ini diambil dari penjualan minuman kopi di Kopi Pawon Nusantara periode bulan januari samapai dengan bulan desember 2023. Memiliki data sebanyak 228 entri.

BULAN	NO	ITEM	JUMLAH	HARGA	TOTAL
JANUARI	1	AMERICANO	640	15,000	9,600,000
	2	ARABICA	515	8,000	4,120,000
	3	BLACK WHITE COFFEE	94	15,000	1,260,000
	4	CAPPUCINO	137	15,000	2,055,000
	5	CHOCO LATTE	156	15,000	2,340,000
	6	COFFE LATTE	243	15,000	3,645,000
	7	COFFE MILK	117	15,000	1,755,000
	8	COFFE SUJREN	53	15,000	795,000
	9	ESPRESSO SINGLE	115	15,000	1,725,000
	10	ESPRESSO DOUBLE	80	15,000	1,200,000
	11	JAPANESE COFFEE	113	15,000	1,695,000
	12	KOPI SUSU TARIK	71	15,000	1,065,000
	13	LANANG	305	7,000	2,135,000
	14	MACHIATO	152	15,000	2,280,000
	15	MOCCACINO	199	15,000	2,985,000
	16	ROBUSTA	132	6,000	792,000
	17	V60	85	15,000	1,275,000
	18	VANILLA COFFEE	117	15,000	1,755,000
	19	VIETNAM DRIP	109	15,000	1,635,000
...					
DESEMBER	1	AMERICANO	25	15,000	375,000
	2	ARABICA	100	8,000	800,000
	3	BLACK WHITE COFFEE	94	15,000	1,260,000
	4	CAPPUCINO	137	15,000	2,055,000
	5	CHOCO LATTE	156	15,000	2,340,000
	6	COFFE LATTE	250	15,000	3,750,000
	7	COFFE MILK	117	15,000	1,755,000
	8	COFFE SUJREN	53	15,000	795,000
	9	ESPRESSO SINGLE	115	15,000	1,725,000
	10	ESPRESSO DOUBLE	80	15,000	1,200,000
	11	JAPANESE COFFEE	113	15,000	1,695,000
	12	KOPI SUSU TARIK	450	15,000	6,750,000
	13	LANANG	200	7,000	1,400,000
	14	MACHIATO	152	15,000	2,280,000
	15	MOCCACINO	199	15,000	2,985,000
	16	ROBUSTA	132	6,000	792,000
	17	V60	85	15,000	1,275,000
	18	VANILLA COFFEE	45	15,000	675,000
	19	VIETNAM DRIP	99	15,000	1,485,000

Sumber : Kopi Pawon Nusantara

Gambar 2. Data penjualan minuman kopi

### 3.2 Tahap Pre-processing

#### Seleksi Data

Proses seleksi data dilakukan untuk memilih variabel yang akan digunakan dalam penelitian. Variabel yang digunakan pada penelitian yaitu item, jumlah, harga, dengan menambahkan atribut kategori, proses, dan keterangan. Atribut yang akan dihapus yaitu bulan, total, dan nomor. Berikut pada gambar dibawah ini dataset yang sudah di seleksi.

No.	ITEM	KATEGORI	PROSES	HARGA	JUMLAH	KETERANGAN
1	AMERICANO	ESPRESSO BASED	TIDAK RACIKAN	15000	640	LARIS
2	ARABICA	KOPI TUBRUK	TIDAK RACIKAN	8000	515	LARIS
3	BLACK WHITE COFFEE	ESPRESSO BASED	TIDAK RACIKAN	15000	84	TIDAK LARIS
4	CAPPUCINO	ESPRESSO BASED	TIDAK RACIKAN	15000	137	TIDAK LARIS
5	CHOCO LATTE	ESPRESSO BASED	TIDAK RACIKAN	15000	156	LARIS
6	COFFE LATTE	ESPRESSO BASED	TIDAK RACIKAN	15000	243	LARIS
7	COFFE MILK	ESPRESSO BASED	TIDAK RACIKAN	15000	117	TIDAK LARIS
8	COFFEE SUREN	ESPRESSO BASED	RACIKAN	15000	53	TIDAK LARIS
9	ESPRESSO SINGLE	ESPRESSO BASED	TIDAK RACIKAN	15000	115	TIDAK LARIS
10	ESPRESSO DOUBLE	ESPRESSO BASED	TIDAK RACIKAN	15000	80	TIDAK LARIS
11	JAPANESE COFFEE	MANUAL BREW	TIDAK RACIKAN	15000	113	TIDAK LARIS
12	KOPI SUSU TARIK	ESPRESSO BASED	TIDAK RACIKAN	15000	72	TIDAK LARIS
13	LANANG	ESPRESSO BASED	RACIKAN	7000	305	LARIS
14	MACHIATO	ESPRESSO BASED	RACIKAN	15000	152	LARIS
....						
211	ARABICA	KOPI TUBRUK	RACIKAN	8000	100	TIDAK LARIS
212	BLACK WHITE COFFEE	ESPRESSO BASED	TIDAK RACIKAN	15000	84	TIDAK LARIS
213	CAPPUCINO	ESPRESSO BASED	RACIKAN	15000	137	TIDAK LARIS
214	CHOCO LATTE	ESPRESSO BASED	RACIKAN	15000	156	LARIS
215	COFFE LATTE	ESPRESSO BASED	TIDAK RACIKAN	15000	250	LARIS
216	COFFE MILK	ESPRESSO BASED	RACIKAN	15000	117	TIDAK LARIS
217	COFFEE SUREN	ESPRESSO BASED	RACIKAN	15000	53	TIDAK LARIS
218	ESPRESSO SINGLE	ESPRESSO BASED	RACIKAN	15000	115	TIDAK LARIS
219	ESPRESSO DOUBLE	ESPRESSO BASED	TIDAK RACIKAN	15000	80	TIDAK LARIS
220	JAPANESE COFFEE	MANUAL BREW	RACIKAN	15000	113	TIDAK LARIS
221	KOPI SUSU TARIK	ESPRESSO BASED	RACIKAN	15000	450	LARIS
222	LANANG	KOPI TUBRUK	TIDAK RACIKAN	7000	200	LARIS
223	MACHIATO	ESPRESSO BASED	TIDAK RACIKAN	15000	152	LARIS
224	MOCCACINO	ESPRESSO BASED	TIDAK RACIKAN	15000	199	LARIS
225	ROBUSTA	KOPI TUBRUK	RACIKAN	6000	132	TIDAK LARIS
226	V60	MANUAL BREW	RACIKAN	15000	85	TIDAK LARIS
227	VANILLA COFFEE	ESPRESSO BASED	RACIKAN	15000	45	TIDAK LARIS
228	VIETNAM DRIP	MANUAL BREW	RACIKAN	15000	99	TIDAK LARIS

Sumber : Penelitian Mandiri (2024)

Gambar 3. Dataset yang telah di seleksi

#### Tranformasi Data

Berikut ini adalah hasil transformasi data yang siap untuk diolah :

No.	ITEM	KATEGORI	PROSES	N JUAL	KETERANGAN
1	AMERICANO	ESPRESSO BASED	TIDAK RACIKAN	Rp 9,600.000	LARIS
2	ARABICA	KOPI TUBRUK	TIDAK RACIKAN	Rp 4,120.000	LARIS
3	BLACK WHITE COFFEE	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,260.000	TIDAK LARIS
4	CAPPUCINO	ESPRESSO BASED	TIDAK RACIKAN	Rp 2,055.000	TIDAK LARIS
5	CHOCO LATTE	ESPRESSO BASED	TIDAK RACIKAN	Rp 2,340.000	LARIS
6	COFFE LATTE	ESPRESSO BASED	TIDAK RACIKAN	Rp 3,645.000	LARIS
7	COFFE MILK	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,755.000	TIDAK LARIS
8	COFFEE SUREN	ESPRESSO BASED	RACIKAN	Rp 795.000	TIDAK LARIS
9	ESPRESSO SINGLE	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,725.000	TIDAK LARIS
10	ESPRESSO DOUBLE	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,200.000	TIDAK LARIS
11	JAPANESE COFFEE	MANUAL BREW	TIDAK RACIKAN	Rp 1,695.000	TIDAK LARIS
12	KOPI SUSU TARIK	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,080.000	TIDAK LARIS
13	LANANG	ESPRESSO BASED	RACIKAN	Rp 2,135.000	LARIS
14	MACHIATO	ESPRESSO BASED	RACIKAN	Rp 2,280.000	LARIS
....					
211	ARABICA	KOPI TUBRUK	RACIKAN	Rp 800.000	TIDAK LARIS
212	BLACK WHITE COFFEE	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,260.000	TIDAK LARIS
213	CAPPUCINO	ESPRESSO BASED	RACIKAN	Rp 2,055.000	TIDAK LARIS
214	CHOCO LATTE	ESPRESSO BASED	RACIKAN	Rp 2,340.000	LARIS
215	COFFE LATTE	ESPRESSO BASED	TIDAK RACIKAN	Rp 3,750.000	LARIS
216	COFFE MILK	ESPRESSO BASED	RACIKAN	Rp 1,755.000	TIDAK LARIS
217	COFFEE SUREN	ESPRESSO BASED	RACIKAN	Rp 795.000	TIDAK LARIS
218	ESPRESSO SINGLE	ESPRESSO BASED	RACIKAN	Rp 1,725.000	TIDAK LARIS
219	ESPRESSO DOUBLE	ESPRESSO BASED	TIDAK RACIKAN	Rp 1,200.000	TIDAK LARIS
220	JAPANESE COFFEE	MANUAL BREW	RACIKAN	Rp 1,695.000	TIDAK LARIS
221	KOPI SUSU TARIK	ESPRESSO BASED	RACIKAN	Rp 6,750.000	LARIS
222	LANANG	KOPI TUBRUK	TIDAK RACIKAN	Rp 1,400.000	LARIS
223	MACHIATO	ESPRESSO BASED	TIDAK RACIKAN	Rp 2,280.000	LARIS
224	MOCCACINO	ESPRESSO BASED	TIDAK RACIKAN	Rp 2,985.000	LARIS
225	ROBUSTA	KOPI TUBRUK	RACIKAN	Rp 792.000	TIDAK LARIS
226	V60	MANUAL BREW	RACIKAN	Rp 1,275.000	TIDAK LARIS
227	VANILLA COFFEE	ESPRESSO BASED	RACIKAN	Rp 675.000	TIDAK LARIS
228	VIETNAM DRIP	MANUAL BREW	RACIKAN	Rp 1,485.000	TIDAK LARIS

Sumber : Penelitian Mandiri (2024)

Gambar 4. Hasil dari transformasi data

Jenis Kelas	Jumlah
LARIS	73
TIDAK LARIS	155
Jumlah	228

**Tabel 2.** Pengelompokan data berdasarkan kategori labelnya dan jumlahnya

Sumber : Penelitian Mandiri (2024)

### 3.3 Implementasi Algoritma

Berikut adalah tabel yang menampilkan hasil klasifikasi yang telah diperoleh :

**Tabel 3.** Hasil dari *precision* dan *recall* pada *Decision Tree*

Label	Precision	Recall
LARIS	98,55%	94,44%
TIDAK LARIS	97,48%	99,36%

Sumber : Penelitian Mandiri (2024)

Berdasarkan tabel 3, hasil evaluasi *decision tree* menunjukkan bahwa label LARIS memiliki nilai *precision* sebesar 98,55% dan *recall* sebesar 94,44%, sedangkan label TIDAK LARIS menunjukkan nilai *precision* sebesar 97,48% dan *recall* sebesar 99,36%.

**Tabel 4.** Hasil dari *precision* dan *recall* pada *Naive Bayes*

Label	Precision	Recall
LARIS	91,55%	90,28%
TIDAK LARIS	95,54%	96,15%

Sumber : Penelitian Mandiri (2024)

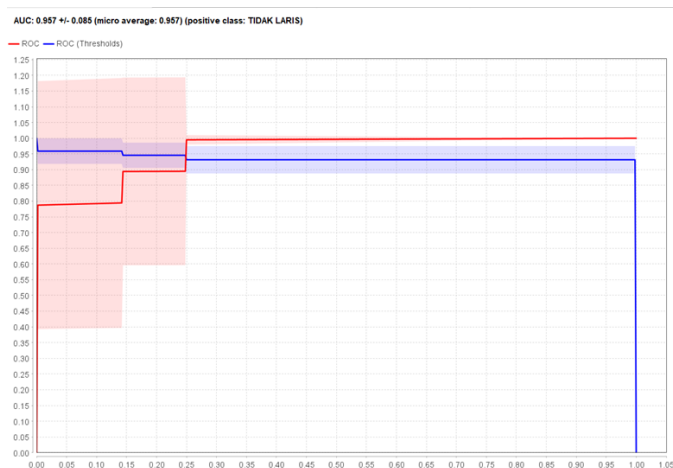
Berdasarkan tabel 4, hasil evaluasi *naive bayes* menunjukkan bahwa label LARIS memiliki nilai *precision* sebesar 91,55% dan *recall* sebesar 90,28%, sedangkan label TIDAK LARIS menunjukkan nilai *precision* sebesar 95,54% dan *recall* sebesar 96,15%.

**Tabel 5.** Perbandingan tingkat keakuratan antara metode klasifikasi *Decision tree* dan *Naive Bayes*.

Metode klasifikasi	Nilai akurasi
Decision Tree	97,83%
Naive Bayes	94,35%

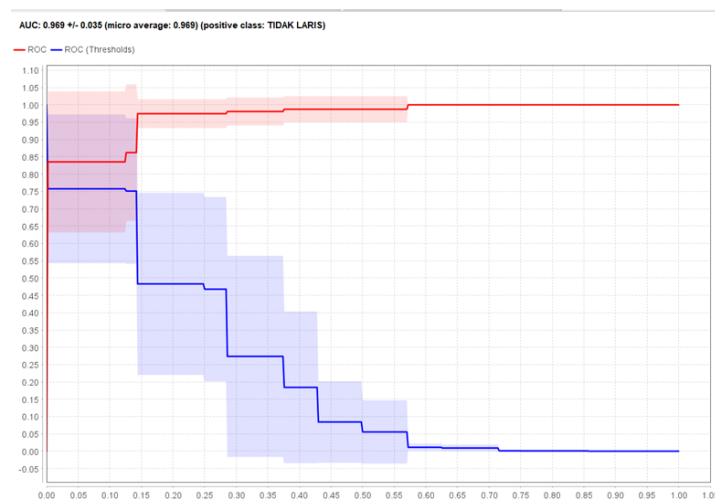
Sumber : Penelitian Mandiri (2024)

Setelah membandingkan presisi dan recall, penggunaan akurasi digunakan untuk menilai performa klasifikasi yang dilakukan. Hasil evaluasi akurasi dari kedua metode klasifikasi menunjukkan bahwa *Decision Tree* memiliki tingkat akurasi tertinggi, mencapai 97,83%. Sebaliknya, *Naive Bayes* memiliki tingkat akurasi yang sedikit lebih rendah, yakni sebesar 94,35%.



Sumber : Penelitian Mandiri (2024)

**Gambar 5.** Nilai Area Under Curve (AUC) pada *Decision Tree*



Sumber : Penelitian Mandiri (2024)

**Gambar 6.** Nilai Area Under Curve (AUC) pada *Naive Bayes*

Grafik AUC menampilkan bahwa performa *Decision Tree* dan *Naive Bayes* diklasifikasikan dalam kategori Excellent berdasarkan nilai AUC mereka. *Decision Tree* memiliki AUC sebesar 0,957 sedangkan *Naive Bayes* memiliki AUC sebesar 0,969.

## KESIMPULAN

Tujuan dari penelitian ini adalah untuk membandingkan dua model klasifikasi, yaitu *Decision Tree* dan *Naive Bayes*, dalam upaya meningkatkan penjualan di Kopi Pawon Nusantara dengan menggunakan data penjualan minuman kopi untuk memahami pola pembelian pelanggan. Pemilik menghadapi tantangan dalam menentukan minuman kopi yang paling diminati oleh pelanggan. Dua algoritma, *Decision Tree* dan *Naive Bayes*, dievaluasi berdasarkan akurasi, presisi, *recall*, dan AUC dari kurva ROC. Hasilnya menunjukkan bahwa *Decision Tree* mencapai akurasi 97,83% dengan AUC 0,957, sementara *Naive Bayes* memiliki akurasi 94,35% dengan AUC 0,969. Kesimpulannya, *Decision Tree* merupakan algoritma terbaik untuk mengolah data penjualan minuman kopi di Kopi Pawon Nusantara dengan akurasi tertinggi sebesar 97,83%.

## REFERENSI

- Adam, F., Agustina, R., & Fadhil, R. (2022). Pengujian Cita Rasa Kopi Arabika dengan Metode Cupping Test. *Jurnal Ilmiah Mahasiswa Pertanian*, 7(1), 517–521. <https://doi.org/10.17969/jimfp.v7i1.19021>
- Bei, F., & Jamu, M. E. (2022). Strategi Pemasaran Pada Usaha Kopi Kita Dalam Mempertahankan Penjualan Di Tengah Pandemi Covid-19. *Analisis*, 12(2), 193–207. <https://doi.org/10.37478/als.v12i2.2081>
- Chandra Wirawan. (2020). *Teknik Data Mining Menggunakan Algoritma Decision Tree C4*. 3(1), 47–52.
- Fihir, M., Martanto, & Hayati, U. (2023). *KLASIFIKASI TINGKAT KEPUASAN PELANGGAN KOPI KENANGANMENGUNAKAN METODE DECISION TREE PADAAPLIKASI KOPI KENANGAN*. 7(6), 3830–3833.
- Harahap, F., Saragih, N. E., Siregar, E. T., & Sariangshah, H. (2021). Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat. *Jurnal Ilmiah Informatika*, 9(01), 19–23. <https://doi.org/10.33884/jif.v9i01.3702>
- Indriyani Indriyani, & Agus Bahtiar. (2023). Implementasi Data Mining Untuk Mengklasifikasikan Data Penjualan Pada Supermarket Menggunakan Algoritma Naive Bayes. *Jurnal Manajemen Dan Bisnis Ekonomi*, 1(1), 207–220. <https://doi.org/10.54066/jmbe-itb.v1i1.70>
- Nosiel, N., Sriyanto, S., & Maylani, F. (2021). Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Pada UMKM Gerabah. *Prosiding Seminar Nasional Darmajaya*, 1, 72–86.
- Suci Amaliah, Nusrang, M., & Aswi, A. (2022). Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(3), 121–127. <https://doi.org/10.35580/variasiunm31>
- Sulastri, R. (2022). Identifikasi Tingkat Penjualan Produk Herbal HWI Menggunakan Algoritma C4.5. *Jurnal Informatika Ekonomi Bisnis*, 4(3), 84–89. <https://doi.org/10.37034/infep.v4i3.141>
- Wahyudi, M. D. (2023). Penerapan data mining dengan algoritma c4. 5 dalam prediksi penjualan buku. *Jurnal Teknorama (Informatika dan Teknologi El Rahma)*, 1(1).