PREDIKSI TINGKAT MINAT BELI PELANGGAN PADA APLIKASI SHOPEE MENGGUNAKAN ALGORITMA C4.5

Rangga Wijayadi[1], Nadhif Akmal[2], Naufal Indra Fata[3], Hendri[4], Nur Ali Farabi[5]

[1,2,3,4,5]*Universitas Bina Sarana Informatika*
E-mail : ranggawijayadi50@gmail.com[1], nadhifakmal53@gmail.com[2], naufalindrafata16@gmail.com[3], hendri.hed@bsi.ac.id,nur.naf@bsi.ac.id[4]

ABSTRAK

Perkembangan *e-commerce* di Indonesia, khususnya Shopee, meningkatkan pentingnya analisis perilaku konsumen untuk mengetahui tingkat minat beli pelanggan. Penelitian ini bertujuan membangun model klasifikasi menggunakan algoritma C4.5 untuk memprediksi tingkat minat beli pelanggan berdasarkan variabel frekuensi pembelian, pengalaman pengguna, dan kepercayaan. Penelitian ini juga bertujuan mengetahui performa model serta variabel yang paling dominan memengaruhi minat beli pelanggan. Metode penelitian yang digunakan adalah pendekatan kuantitatif dengan teknik *data mining*. Data diperoleh melalui penyebaran kuesioner kepada 255 responden pengguna Shopee. Tahapan penelitian meliputi pengumpulan data, *preprocessing*, transformasi data, pembangunan model menggunakan algoritma C4.5, dan evaluasi model menggunakan *confusion matrix* dengan bantuan aplikasi RapidMiner. Hasil penelitian menunjukkan bahwa algoritma C4.5 berhasil membentuk model *decision tree* yang mampu mengklasifikasikan tingkat minat beli pelanggan ke dalam kategori rendah, sedang, dan tinggi. Model yang dihasilkan memperoleh tingkat akurasi sebesar 93,35%, sehingga menunjukkan performa yang sangat baik dalam melakukan prediksi. Variabel frekuensi pembelian menjadi faktor paling dominan dalam menentukan tingkat minat beli pelanggan, diikuti oleh kepercayaan dan pengalaman pengguna. Penelitian ini diharapkan dapat menjadi dasar pengambilan keputusan strategis dalam meningkatkan minat beli pelanggan pada platform Shopee.

Kata Kunci : *Data Mining, Algoritma C4.5, Minat Beli, Shopee, RapidMiner*

PENDAHULUAN

Transformasi digital yang berlangsung dalam beberapa tahun terakhir telah membawa perubahan besar pada berbagai sektor, termasuk aktivitas perdagangan elektronik di Indonesia. Peningkatan akses terhadap teknologi informasi dan komunikasi memungkinkan masyarakat melakukan transaksi secara lebih mudah melalui platform daring. Dalam kondisi tersebut, Shopee muncul sebagai salah satu marketplace yang memiliki tingkat penggunaan tinggi dan mencatat pertumbuhan transaksi yang signifikan. Tingginya aktivitas pada platform tersebut mengindikasikan adanya perubahan preferensi konsumen dalam memenuhi kebutuhan, dari yang sebelumnya mengandalkan transaksi tatap muka menjadi lebih memilih proses pembelian yang dilakukan secara *online*. Dalam kondisi persaingan yang semakin ketat antar platform *e-commerce*, pemahaman terhadap perilaku konsumen, khususnya terkait minat beli pelanggan, menjadi faktor krusial dalam menentukan strategi bisnis yang efektif dan berbasis data (Diven & Khoiri) (Khasan et al.)

Dalam kajian teori pemasaran, minat beli diartikan sebagai dorongan atau kecenderungan yang muncul dalam diri individu untuk melakukan pembelian terhadap suatu produk maupun jasa. Kecenderungan tersebut mencerminkan adanya keinginan dan ketertarikan konsumen yang dapat memengaruhi keputusan mereka dalam melakukan transaksi pada masa mendatang. Minat beli diidentifikasi melalui berbagai atribut perilaku konsumen seperti frekuensi pembelian, pengalaman pengguna, dan tingkat kepercayaan terhadap platform. Dalam konteks *e-commerce*, pengalaman pengguna yang baik serta tingkat kepercayaan yang tinggi terbukti meningkatkan kemungkinan terjadinya transaksi. Oleh karena itu, analisis terhadap minat beli pelanggan menjadi penting untuk memahami pola perilaku konsumen secara lebih mendalam dan terukur (Diven & Khoiri), (Khasan et al.).

Seiring dengan meningkatnya volume data yang dihasilkan dari aktivitas pengguna *e-commerce*, diperlukan pendekatan berbasis data untuk mengolah dan menganalisis informasi tersebut secara optimal. Analisis data dengan pendekatan klasifikasi merupakan salah satu teknik yang umum digunakan dalam bidang *data mining* untuk mengidentifikasi pola dan membuat prediksi berdasarkan data historis. Salah satu algoritma yang sering dimanfaatkan adalah C4.5, yang bekerja dengan membentuk struktur pohon keputusan sebagai dasar penentuan klasifikasi. Keunggulan metode ini terletak pada kemudahan interpretasi hasil serta kemampuannya dalam menangani atribut yang bersifat numerik maupun kategorikal secara optimal. Dengan kemampuan tersebut, algoritma C4.5 dinilai relevan untuk digunakan dalam memprediksi tingkat minat beli pelanggan berdasarkan data historis (Firdaus et al.) (Sunarto et al.)

Hasil dari berbagai penelitian terdahulu menunjukkan bahwa masih terdapat beberapa keterbatasan dalam kajian yang telah dilakukan. Keterbatasan tersebut mengindikasikan adanya *research gap* yang belum terakomodasi secara optimal, sehingga diperlukan penelitian lanjutan untuk melengkapi dan memperkaya temuan-temuan yang sudah ada. Penelitian sebelumnya cenderung berfokus pada kepuasan pelanggan dan status transaksi, sehingga belum mampu memberikan gambaran prediktif terhadap minat beli pelanggan secara langsung. Hal ini menyebabkan hasil penelitian kurang mendukung pengambilan keputusan strategis yang bersifat prediktif (Steffany et al.).. Penelitian terkait prediksi minat beli umumnya menggunakan dataset yang bersifat umum dan tidak spesifik pada platform tertentu, sehingga hasil analisis kurang mencerminkan kondisi nyata perilaku pengguna pada platform Shopee (Sunarto et al.).. Selain itu, variabel perilaku konsumen seperti frekuensi pembelian, pengalaman pengguna, dan kepercayaan

masih sering dianalisis secara terpisah dan belum diintegrasikan secara komprehensif dalam satu model klasifikasi, sehingga mengurangi akurasi dan kedalaman analisis (Diven & Khoiri), (Khasan et al.).

Adanya keterbatasan tersebut menunjukkan bahwa diperlukan suatu pendekatan yang lebih spesifik, terintegrasi, dan berbasis data aktual untuk menghasilkan model prediksi minat beli yang lebih akurat dan kontekstual. Oleh karena itu, penelitian ini mengusulkan penerapan algoritma C4.5 dalam memprediksi tingkat minat beli pelanggan pada platform Shopee dengan menggunakan data yang relevan serta mengintegrasikan variabel perilaku konsumen dalam satu model klasifikasi. Pendekatan ini diharapkan mampu mengatasi keterbatasan penelitian sebelumnya dengan menghasilkan model yang tidak hanya akurat, tetapi juga representatif terhadap kondisi nyata pengguna (Firdaus et al.).. Selain berfokus pada analisis data, penelitian ini juga memanfaatkan RapidMiner sebagai perangkat pendukung dalam proses pengolahan data. Penggunaan tools tersebut mencakup berbagai tahapan, mulai dari persiapan dan pembersihan data (*preprocessing*), pembentukan model, hingga pengujian dan evaluasi kinerja model yang dilakukan secara terstruktur. Penggunaan RapidMiner memungkinkan proses analisis dilakukan secara terstruktur dan efisien sehingga mendukung pengembangan model yang lebih optimal (Hidayatuloh et al.) (Nuryawan et al.). Dengan demikian, penelitian ini memiliki kebaruan (*novelty*) yang terletak pada penerapan algoritma C4.5 sebagai pengguna Shopee dengan mengintegrasikan tiga variabel perilaku yaitu frekuensi pembelian, pengalaman pengguna, dan kepercayaan dalam proses klasifikasi minat beli. Kombinasi variabel tersebut belum ditemukan secara bersamaan pada penelitian yang menjadi rujukan penelitian ini (Firdaus et al.) (Baktiar).

Tabel 1. State of Art

No	Penelitian	Metode	Hasil	Kelemahan/Gap
1	Steffany (2023)	C4.5	Mampu klasifikasi kepuasan baik.	Belum prediksi minat beli.
2	Sunarto (2023)	C4.5+Boost	Meningkatkan akurasi.	Dataset umum, kurang spesifik.
3	Diven (2023)	Perilaku	Frekuensi berpengaruh.	Variabel terpisah, belum ada model.
4	Khasan (2024)	Perilaku	Pengalaman & Kepercayaan ngefek.	Belum integrasi variabel.
5	Firdaus (2025)	C4.5	Efektif status pesanan.	Bukan objek minat beli.
6	Penelitian (2026)	C4.5	Integrasi 3 variabel, akurasi 93,35%.	*Novelty* fokus minat beli.

Dari Tabel 1 dapat dilihat bahwa penelitian-penelitian terdahulu mengenai platform Shopee masih berfokus secara parsial pada aspek kepuasan pelanggan atau pelacakan status transaksi tanpa memodelkan kecenderungan prediktif dari niat beli konsumen secara langsung. Melalui pemetaan ilmiah tersebut, perbedaan mendasar sekaligus kebaruan (novelty) pada penelitian ini ditunjukkan dengan keberhasilan mengintegrasikan variabel frekuensi pembelian, tingkat kepercayaan, serta pengalaman pengguna secara simultan ke dalam arsitektur klasifikasi algoritma C4.5 guna memprediksi tingkat minat beli pelanggan secara akurat.

Tinjauan Pustaka

Algoritma C4.5

C4.5 merupakan salah satu algoritma klasifikasi yang banyak diterapkan dalam teknik *data mining* karena kemampuannya dalam membangun model berbentuk pohon keputusan. Algoritma ini dikembangkan sebagai penyempurnaan dari metode ID3 dengan tujuan meningkatkan efektivitas proses klasifikasi data. Dalam pembentukan pohon keputusan, C4.5 memilih atribut yang paling informatif sebagai dasar pemisahan data sehingga setiap cabang yang terbentuk mampu menggambarkan karakteristik data secara lebih jelas. Model yang dihasilkan kemudian dapat dimanfaatkan untuk mendukung proses prediksi dan pengambilan Keputusan (Rahmawati et al.). Pembentukan model pohon keputusan pada algoritma C4.5 diawali dengan proses pemilihan atribut yang memiliki kemampuan terbaik dalam membedakan data, yang ditentukan berdasarkan nilai *information gain or gain ratio* tertinggi. Atribut terpilih kemudian ditempatkan sebagai akar pohon (*root node*) dan digunakan untuk memisahkan data ke dalam beberapa kelompok yang memiliki karakteristik lebih seragam (Kurniawan et al.),

Decision tree

Decision tree atau pohon keputusan merupakan teknik klasifikasi yang banyak digunakan dalam *data mining* untuk menggambarkan proses pengambilan keputusan secara sistematis. Metode ini menyajikan hubungan antar atribut dalam bentuk struktur menyerupai pohon yang terdiri atas beberapa bagian utama. Bagian tersebut meliputi *root node* sebagai titik awal proses pengambilan keputusan, *internal node* yang berfungsi sebagai titik percabangan berdasarkan atribut tertentu, serta *leaf node* yang menunjukkan hasil atau kelas akhir dari proses klasifikasi Baktiar). Pada *decision tree* , setiap *node* menggambarkan atribut atau variabel yang digunakan sebagai dasar pengambilan keputusan, sedangkan cabang menunjukkan nilai dari atribut tersebut. Proses klasifikasi dilakukan dengan menelusuri pohon dari akar hingga ke daun, sehingga menghasilkan suatu keputusan berdasarkan kondisi yang terpenuhi (Darmansah et al.).. Metode *decision tree* melakukan proses klasifikasi dengan membagi kumpulan data ke dalam beberapa kelompok yang lebih spesifik berdasarkan atribut tertentu (Kurniawan et al.).

RapidMiner

RapidMiner merupakan perangkat lunak analisis data yang banyak digunakan dalam penerapan *data mining* karena menyediakan berbagai fitur pengolahan data dalam satu platform yang terintegrasi. Melalui aplikasi ini, pengguna dapat melakukan berbagai tahapan analisis, mulai dari persiapan data hingga pembangunan model, dengan lebih mudah dan efisien (Hidayatuloh et al.). RapidMiner menyediakan berbagai fitur yang mendukung proses *Knowledge Discovery in Database* (KDD), mulai dari tahap data *preprocessing*, transformación data, pemodelan (*modeling*), hingga evaluasi hasil. Setiap tahapan dapat dilakukan dengan menggunakan operator-operator yang tersedia dan disusun dalam bentuk alur kerja (*workflow*) yang sistematis (Nuryawan et al.).. Perangkat lunak ini menyediakan beragam algoritma klasifikasi, seperti C4.5, Naïve Bayes,

dan K-Nearest Neighbor, yang dapat digunakan sesuai dengan kebutuhan analisis (Darmansah et al.)

METODE

Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif yang bertujuan untuk mengolah data numerik yang diperoleh dari responden. Pendekatan yang digunakan adalah *data mining* dengan fokus pada predictive analysis, yaitu membangun model untuk mengklasifikasikan dan memprediksi tingkat minat beli pengguna. Dalam penelitian ini, algoritma yang digunakan adalah *Decision tree* C4.5, yang berfungsi untuk menghasilkan model klasifikasi dalam bentuk pohon keputusan yang mudah diinterpretasikan. Selain itu, model yang dihasilkan akan dievaluasi menggunakan metode pengujian seperti *confusion matrix* untuk mengukur tingkat akurasi, precision, dan recall, sehingga dapat diketahui performa model dalam mengklasifikasikan data secara tepat.

Populasi dan Sampel

Populasi dalam penelitian ini adalah pengguna Shopee yang memenuhi kriteria penelitian dan bersedia menjadi responden yang pernah melakukan transaksi atau memiliki pengalaman dalam menggunakan platform tersebut. Penelitian ini menerapkan teknik *purposive sampling* sebagai metode pengambilan sampel, yaitu proses seleksi responden berdasarkan karakteristik dan pertimbangan khusus yang telah ditentukan oleh peneliti. Kriteria responden yang digunakan dalam penelitian ini meliputi pernah menggunakan aplikasi Shopee dan pernah melakukan pembelian pada aplikasi Shopee. Populasi dalam penelitian ini berjumlah 255 responden. Untuk menentukan jumlah sampel minimal yang diperlukan, digunakan rumus Slovin dengan tingkat kesalahan sebesar 5% sebagai berikut:

$$n = \frac{N}{1+N(e)^2} \quad (1)$$

Perhitungan ukuran sampel menghasilkan kebutuhan minimal sebanyak 156 responden:

$$n = \frac{255}{1+255(0,05)^2} \quad (2)$$

$$n = \frac{255}{1+255(0,0025)} \quad (3)$$

$$n = \frac{255}{1,6375} \quad (4)$$

$$n = 155,7 \approx 156 \quad (5)$$

Pada pelaksanaannya, penelitian ini melibatkan 255 responden yang sesuai dengan kriteria yang telah ditetapkan sehingga dinilai sangat representatif.

Indikator Minat Beli

Pada penelitian ini, pengukuran variabel minat beli dilakukan dengan menggunakan sejumlah indikator yang dikembangkan melalui pernyataan-pernyataan dalam kuesioner. Adapun indikator yang digunakan meliputi: (1) Ketertarikan membeli produk setelah melihat rekomendasi di Shopee, (2) Ketertarikan terhadap produk yang muncul pada beranda Shopee, (3) Preferensi memilih Shopee dibandingkan platform *e-commerce* lain, (4) Keinginan untuk kembali berbelanja di Shopee, dan (5) Ketertarikan mencoba produk baru yang direkomendasikan oleh Shopee.

Uji Validitas dan Reliabilitas

Uji validitas dilakukan menggunakan korelasi Pearson Product Moment dengan rumus berikut:

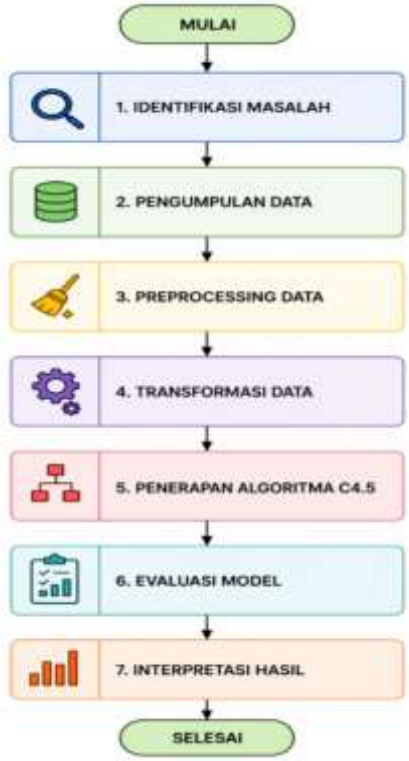
$$r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \quad (6)$$

Uji reliabilitas dilakukan menggunakan metode Cronbach's Alpha dengan rumus sebagai berikut:

$$a = \frac{k}{k-1}\left(1 - \frac{\sum S_i^2}{S_t^2}\right) \quad (7)$$

Tahapan Penelitian

Tahapan penelitian secara terstruktur digambarkan melalui diagram alur pada Gambar 1 di bawah ini.



Gambar 1. Alur Tahapan Penelitian

Dari bagan alur Gambar 1, perhitungan evaluasi akurasi, presisi, dan *recall* diturunkan berdasarkan rumus matematika berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Penelitian

Dataset berisi atribut penelitian seperti jenis kelamin, usia, status responden, frekuensi pembelian, pengalaman pengguna, kepercayaan, dan minat beli yang selanjutnya digunakan sebagai data input dalam proses pengolahan dan pemodelan menggunakan algoritma C4.5 seperti ditunjukkan pada Gambar 2.



Gambar 2. Baris Mentah Dataset Hasil Kuesioner

Dari data mentah Gambar 2, pengujian instrumen dilanjutkan ke tahap uji validitas. Hasil visualisasi *output* korelasi tiap butir instrumen dipetakan pada Gambar 3.

Validitas																
Correlation	0,727	0,721	0,668	0,700	0,698	0,654	0,638	0,710	0,705	0,731	0,774	0,680	0,611	0,695	0,718
R Tabel	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123	0,123
Keputusan	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid

Gambar 3. Hasil uji validitas

Dari grafik Gambar 3, pengujian dilanjutkan dengan uji reliabilitas internal yang hasilnya ditunjukkan pada Gambar 4.

Realibilitas															
Varian	0,923	0,7163	0,7171	0,793	0,716	0,857	0,637	0,625	0,74	0,797	0,781	1,053	0,968	0,627	0,896	85,624
Jumlah Varian	11,847															
Varian Total	85,624															
Keputusan	0,923	Realiabel														

Gambar 4. Hasil uji reabilitas

Data kuesioner kemudian dibersihkan dan dipisahkan menjadi dataset siap modeling yang ditunjukkan pada Gambar 5.

	A	B	C	D
1	Pengalaman Pengguna	Kepercayaan	Frekuensi	Kategori Minat Beli
2	4,0	3,8	47	Tinggi
3	4,0	4,0	32	Sedang
4	4,0	4,0	37	Sedang
5	4,0	4,0	49	Tinggi
6	5,0	4,0	41	Tinggi
7	4,0	4,0	44	Tinggi
8	4,0	4,0	49	Tinggi
9	4,6	4,0	36	Sedang
10	5,6	5,0	48	Tinggi
11	4,2	3,6	31	Sedang
245	4,4	3,8	12	Rendah
246	3,6	2,6	9	Rendah
247	3,4	4,2	23	Sedang
248	3,4	3,6	35	Sedang
249	3,2	3,8	18	Rendah
250	3,8	3,4	6	Rendah
251	4	3,8	41	Tinggi
252	5	3,4	27	Sedang
253	4	4	14	Rendah
254	3,8	3,2	38	Sedang
255	5	5	20	Sedang
256	4	3,8	11	Rendah

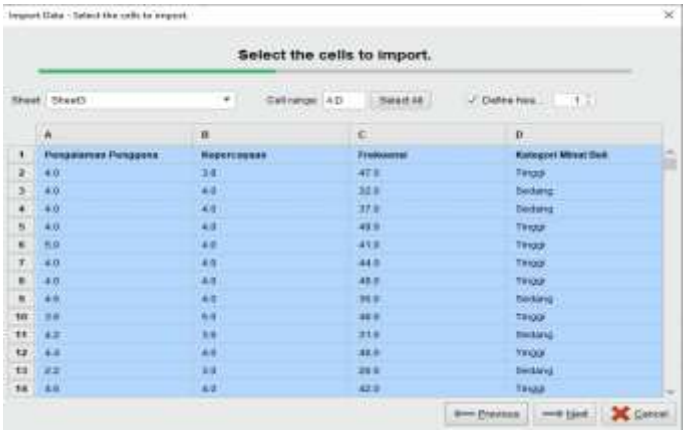
Gambar 5. Dataset siap modeling

Dari dataset Gambar 5, dilakukan import berkas ke aplikasi RapidMiner seperti pada Gambar 6.



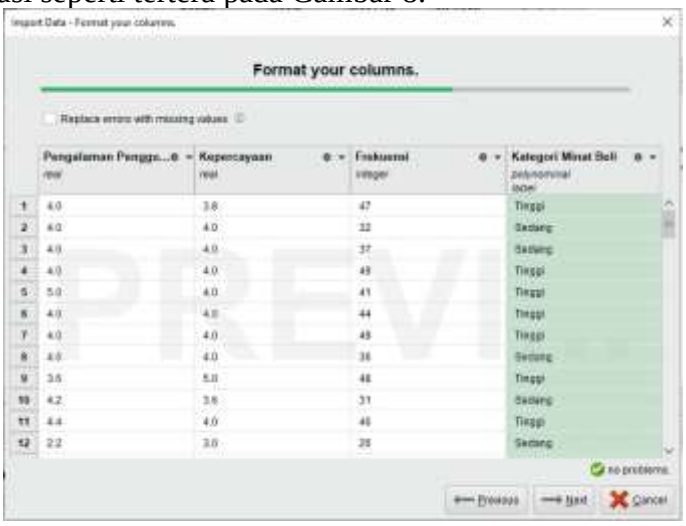
Gambar 6. Impor file dari aplikasi rapidminer

Tahapan dilanjutkan dengan memilih sheet aktif pengolahan data seperti tertera pada Gambar 7.



Gambar 7. Impor sheet yang ingin di gunakan

Perlakuan khusus diterapkan pada variabel target dengan mengeset peran kolom atribut target menjadi label klasifikasi seperti tertera pada Gambar 8.



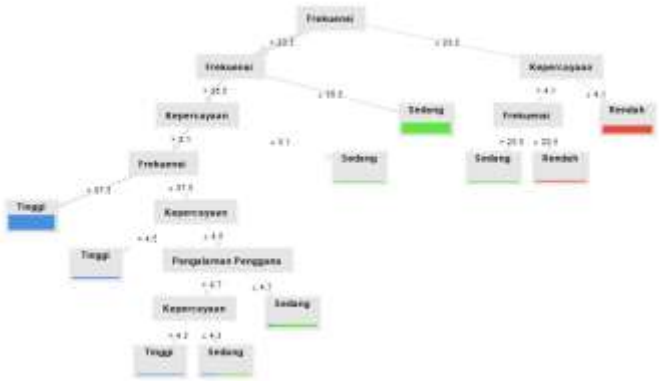
Gambar 8. Menentukan role atribut dari minat beli

Rangkaian operator pemodelan pohon keputusan disusun pada kanvas proses utama seperti ditunjukkan pada alur Gambar 9.



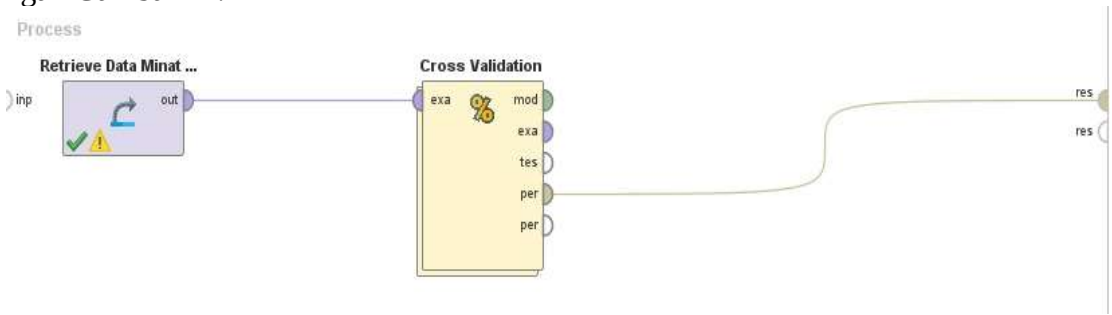
Gambar 9. Implementasi Menggunakan *Decision Tree*

Eksekusi dari alur operator menghasilkan bagan visualisasi struktur percabangan *decision tree* utuh yang digambarkan pada Gambar 10.



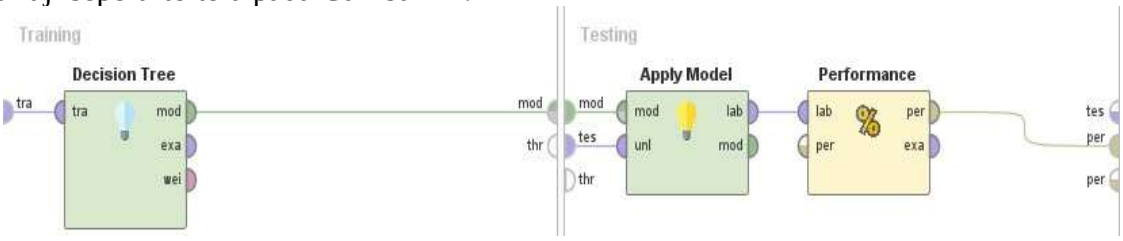
Gambar 10. Hasil dari decesion tree

Evaluasi keandalan model diuji menggunakan kerangka validitas silang seperti pada rancangan Gambar 11.



Gambar 11. Implementasi menggunakan *cross validation*

Detail sub-proses pengerjaan di dalam operator validation dipisah menjadi modul latih dan modul uji seperti tertera pada Gambar 12.



Gambar 12. Proses *cross validation*

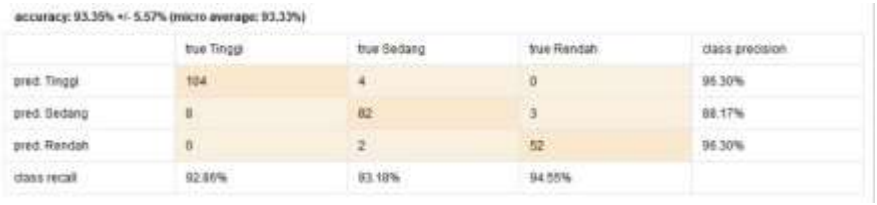
Hasil perhitungan matriks kesalahan terekam pada *PerformanceVector* ringkas seperti ditunjukkan pada Gambar 13.

PerformanceVector

```
PerformanceVector:
accuracy: 93.35% +/- 5.57% (micro average: 93.33%)
ConfusionMatrix:
True:   Tinggi  Sedang  Rendah
Tinggi: 104     4       0
Sedang: 8       82      3
Rendah: 0       2       52
kappa: 0.9 +/- 0.1 (micro average: 0.9)
ConfusionMatrix:
True:   Tinggi  Sedang  Rendah
Tinggi: 104     4       0
Sedang: 8       82      3
Rendah: 0       2       52
```

Gambar 13. Hasil performance vector

Rincian perbandingan sebaran prediksi divalidasi silang pada lembar rekapitulasi akhir Gambar 14.



accuracy: 93.35% +/- 5.57% (micro average: 93.33%)

	true Tinggi	true Sedang	true Rendah	class precision
pred. Tinggi	104	4	0	96.30%
pred. Sedang	8	82	3	88.17%
pred. Rendah	0	2	52	96.30%
class recall	92.86%	93.18%	94.55%	

Gambar 14. Hasil akurasi cross validation

Dari data Gambar 14, parameter prapemrosesan rentang kelas skor minat beli dipetakan pada Tabel 2 di bawah ini.

Tabel 2. Kategori Minat Beli

No	Nilai Batas	Kategori Kelas
1	>= 15	Tinggi
2	10.1 - 15	Sedang
3	<= 10	Rendah

Dari deskripsi Tabel 2, hasil performa akhir pengujian model terekam secara komprehensif pada Tabel 3.

Tabel 3. Capaian Kinerja Model C4.5

Metrik Evaluasi	Nilai Eksperimen
Accuracy	93.35%
Precision	93.59%
Recall	93.53%
F1-Score	93.56%

Berdasarkan data operasional pada Tabel 3, model terbukti kokoh dan stabil untuk kebutuhan klasifikasi data perilaku konsumen secara lokal dalam kolom.

Pembahasan

Berdasarkan pengujian menggunakan algoritma C4.5, model yang dihasilkan mampu mengklasifikasikan tingkat minat beli pelanggan dengan tingkat akurasi sebesar 93,35%. Nilai ini

menunjukkan bahwa model memiliki performa yang baik dalam mengklasifikasikan data berdasarkan atribut yang digunakan. Tingginya nilai akurasi menunjukkan bahwa algoritma C4.5 mampu membangun model klasifikasi dari data perilaku konsumen yang terdiri dari atribut frekuensi pembelian, kepercayaan, dan pengalaman pengguna. Proses pembentukan model dilakukan dengan memilih atribut terbaik menggunakan pendekatan *gain ratio* untuk menghasilkan struktur *decision tree* . Struktur pohon keputusan yang dihasilkan relatif sederhana dan mudah diinterpretasikan, sehingga memudahkan dalam memahami aturan (*rules*) yang terbentuk dari proses klasifikasi. Selain itu, struktur model yang tidak terlalu kompleks menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik terhadap data baru. Meskipun demikian, evaluasi model tidak hanya bergantung pada akurasi, tetapi juga perlu mempertimbangkan metrik lain seperti precision dan recall untuk melihat performa model pada masing-masing kelas. Perbedaan utama penelitian ini dibandingkan penelitian terdahulu adalah menggunakan model klasifikasi berbasis algoritma C4.5, mengintegrasikan beberapa atribut dalam satu model klasifikasi, and menghasilkan output berupa *decision tree* yang dapat digunakan untuk membaca pola klasifikasi secara langsung.

KESIMPULAN

Berdasarkan hasil eksperimen, algoritma C4.5 berhasil diimplementasikan untuk memprediksi tingkat minat beli pelanggan pada platform Shopee dengan capaian akurasi sebesar 93,35%, presisi 93,59%, dan recall 93,53%, di mana variabel frekuensi pembelian terbukti menjadi faktor paling dominan (*root node*) yang memengaruhi minat beli konsumen diikuti oleh aspek kepercayaan dan pengalaman pengguna. Untuk pengembangan ke depan, disarankan bagi penelitian selanjutnya untuk memperluas volume dataset responden, mengintegrasikan variabel perilaku baru seperti efektivitas promosi dan kualitas layanan platform, serta membandingkan performa model dengan algoritma alternatif seperti Random Forest atau Naïve Bayes guna memperoleh model prediktif yang lebih komprehensif bagi pengambilan keputusan bisnis *e-commerce*.

DAFTAR PUSTAKA

A. Baktiar, "Decission Tree Sebagai Metode Penentuan Penjurusan Perguruan Tinggi Berdasarkan Minat Dan Bakat Melalui Data Raport Dengan Uji Algoritma C4.5 (Studi Kasus di SMKN 1 Donorojo Pacitan)," Jurnal Pilar Teknologi Jurnal Ilmiah Ilmu Ilmu Teknik, vol. 7, no. 1, pp. 40–45, 2022.

D. Darmansah, W. Novrian, T. N. Putra, and B. Satria, "Penerapan Algoritma C.45 Dalam Pengklasifikasian Data Penjualan Gerai Makanan Cepat Saji Berbasis Website," Journal of Information System Research (JOSH), vol. 5, no. 1, pp. 331–337, 2023.

D. Diven and M. Khoiri, "Pengaruh Kualitas Produk, Iklan, dan Persepsi Konsumen Terhadap keputusan Pembelian di Shopee," Jesya, vol. 6, no. 2, pp. 1509–1516, 2023.

M. A. Hidayatuloh, K. P. Kartika, and D. F. H. Permadi, "Implementasi Algoritma C4.5 Untuk Memprediksi Capaian Pembelajaran Daring (Studi Kasus Siswa MAN 3 Blitar)," Jurnal Algoritme, vol. 3, no. 1, pp. 33–47, 2022.

U. Khasan, I. A. Muttaqin, M. A. Sihabudin, and al., "Analisis Perilaku Konsumen dalam Memutuskan Pembelian secara *Online* pada *E-commerce* Shopee," Agrimics, vol. 1, pp. 15–24, 2024.

Y. I. Kurniawan, A. Fatikasari, M. L. Hidayat, and M. Waluyo, "Prediction for Cooperative Credit Eligibility Using *Data mining* Classification With C4.5 Algorithm," Jurnal Teknik Informatika (Jutif), vol. 2, no. 2, pp. 67–74, 2021.

M. A. Firdaus, F. R. Latarissa, Y. Dzaky, and H. Murtina, "Prediksi Status Pesanan Menggunakan Metode Classification C.45 Pada Toko Stuftech.Id di Shopee," Elkom: Jurnal Elektronika Dan Komputer, vol. 18, no. 1, pp. 336–344, 2025.

A. T. Octa Nuryawan, M. Hasbullah, M. Rizal, M. F. Rajab, and N. Agustina, "Algoritma *Decision tree* Untuk Analisis Sentimen Public Terhadap Marketplace Diindonesia," Naratif: Jurnal Nasional Riset, Aplikasi Dan Teknik Informatika, vol. 5, no. 1, pp. 18–25, 2023.

A. Rahmawati, S. N. Rakhmah, and L. I. Prahartiwi, "Prediksi Target Sistem Productivity Services menggunakan Penerapan Algoritma C4.5 pada PT. Sanggar Sarana Baja Jakarta," Applied Industrial Engineering Journal, vol. 5, no. 1, pp. 1–13, 2021.

A. Steffany, K. A. Aldzikri, M. A. R. Tricahya, and M. I. Arfiandono, "Penerapan Algoritma C4.5 Untuk Menganalisis Kepuasan *E-commerce* Shopee," Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI), vol. 6, no. 6, pp. 700–706, 2023.

A. Sunarto, P. N. Kencana, B. Munadjat, I. K. Dewi, A. Z. Abidin, and R. Rahim, "Application of Boosting Technique with C4.5 Algorithm to Reduce the Classification Error Rate in *Online* Shoppers Purchasing Intention," Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, vol. 14, no. 2, pp. 01–11, 2023.