
**PENERAPAN ALGORITMA DECISION TREE UNTUK MEMREDIKSI
TINGKAT KEBERHASILAN STRATEGI BISNIS DIGITAL UMKM**

Angelina br Tampubolon¹, Jihan Zulwanda², Efandi H.B Sinuhaji³, Saut Parsaoran
Tamba⁴

Universitas Prima Indonesia, Medan

Pusat Unggulan IPTEK (PUI-ITIK) Inovasi Teknologi Ilmu Komputer

e-mail: ¹angelinatampubolon03@gmail.com, ²jihanzulwanda8@gmail.com,

³efandysinuhaji@gmail.com, ⁴sautparsaorantamba@unprimdn.ac.id

Abstract: *The development of digital marketing requires companies to understand customer behavior to ensure more effective promotional strategies. The problem in this study is the difficulty in identifying customers who are likely to respond to digital marketing campaigns because customer data has not been optimally utilized. To address this issue, this study applies the Decision Tree algorithm as a classification method to analyze and predict customer responses based on characteristics and purchasing behavior. The data used is the Customer Personality Analysis dataset from Kaggle, comprising 2,240 customer data. This data was processed through preprocessing, attribute selection, data transformation, SMOTE application, Decision Tree model development, and evaluation using a confusion matrix in RapidMiner Studio. The results showed that the Decision Tree model achieved an accuracy score of 74.48%, with the AcceptedCmp5 attribute being the most influential attribute due to its highest information gain. Based on these results, the Decision Tree algorithm is quite effective in assisting customer behavior analysis and supporting decision-making in digital marketing strategies*

Keywords: *Decision Tree, Classification, Digital Marketing, Customer Behavior Rapidminer*

Abstrak: Perkembangan pemasaran digital menuntut perusahaan untuk memahami perilaku pelanggan agar strategi promosi dapat berjalan lebih efektif. Permasalahan dalam penelitian ini adalah sulitnya menentukan pelanggan yang berpotensi memberikan respons terhadap kampanye pemasaran digital karena data pelanggan belum dimanfaatkan secara optimal. Untuk mengatasi masalah tersebut, penelitian ini menerapkan algoritma Decision Tree sebagai metode klasifikasi untuk menganalisis dan memprediksi respons pelanggan berdasarkan karakteristik dan perilaku pembelian. Data yang digunakan merupakan dataset Customer Personality Analysis dari Kaggle sebanyak 2.240 data pelanggan yang diolah melalui tahap preprocessing, seleksi atribut, transformasi data, penerapan SMOTE, pembentukan model Decision Tree, dan evaluasi menggunakan confusion matrix pada RapidMiner Studio. Hasil penelitian menunjukkan bahwa model Decision Tree memperoleh nilai accuracy sebesar 74,48% dan atribut AcceptedCmp5 menjadi atribut paling berpengaruh karena memiliki nilai information gain tertinggi. Berdasarkan hasil tersebut, algoritma Decision Tree cukup efektif digunakan untuk membantu analisis perilaku pelanggan dan mendukung pengambilan keputusan dalam strategi pemasaran digital

Kata Kunci: *Pohon Keputusan, Klasifikasi, Pemasaran Digital, Perilaku Pelanggan Rapidminer,*

PENDAHULUAN

(UMKM) Usaha mikro, kecil, dan menengah adalah suatu sektor yang memuat peranan esensial pada sektor ekonomi Indonesia. UMKM Sektor yang menjadi salah satu penopang stabilitas ekonomi nasional, terutama dalam kondisi krisis. Melampaui 60 persen kepada (PDB) produk domestik bruto, disumbangkan oleh UMKM, serta sejumlah besar tenaga kerja mampu diserap olehnya [1]. Keberadaan umkm bukan semata-mata berperan sebagai penggerak ekonomi masyarakat, melainkan juga menjadi salah satu penopang utama dalam meningkatkan kesejahteraan masyarakat [2]. Hal ini menunjukkan bahwa UMKM menyumbang kontribusi strategis dalam pembangunan ekonomi inklusif. Meskipun demikian, umkm saat ini menghadapi berbagai tantangan, terutama dalam menghadapi kemajuan teknologi digital yang semakin cepat berkembang [3].

Evolusi teknologi informasi telah memicu pergeseran pola bisnis dari konvensional menuju digital [4]. Perubahan ini dikenal sebagai transformasi digital dalam dunia usaha. Pemasaran yang sebelumnya dilakukan secara tradisional melalui distribusi langsung, penyebaran brosur, dan promosi tatap muka kini mulai beralih ke pemasaran digital melalui media sosial, e-commerce, serta berbagai platform online lainnya [5], [6]. Digitalisasi membuka peluang bagi umkm guna memperlebar cakupan pasar, mengefektifkan kinerja operasional, serta menguatkan daya saing di tengah kompetisi usaha yang semakin sengit [7].

Pada kenyataannya, belum seluruh pihak umkm cakap mendayagunakan teknologi digital dengan maksimal [8]. Ketimpangan kemampuan ini menjadi salah satu faktor yang memengaruhi kesenjangan perkembangan antar pelaku UMKM. Banyak umkm yang masih berhadapan dengan beragam hambatan, misalnya rendahnya literasi digital,

terbatasnya akses terhadap teknologi, serta kurangnya pemanfaatan data dalam pengambilan keputusan bisnis [9], [10]. Akibatnya, evaluasi keberhasilan strategi bisnis digital sering kali masih dilakukan secara subjektif dan berdasarkan pengalaman atau intuisi semata, sehingga sulit untuk mengukur secara objektif tingkat keberhasilan strategi yang telah diterapkan [11].

Kondisi tersebut juga dapat ditemukan pada sebagian umkm yang berada di kota medan sebagaimana dijelaskan dalam penelitian sebelumnya [12]. Hal ini menunjukkan bahwa tantangan dalam pemanfaatan teknologi digital bukan sekedar terjadi pada satu wilayah, melainkan juga bersifat umum di berbagai daerah. Meskipun beberapa pelaku usaha sudah memanfaatkan media sosial serta platform digital guna mempromosikan produk yang mereka hasilkan, namun pemanfaatan data digital dalam menganalisis keberhasilan strategi bisnis masih belum dilakukan secara maksimal [13]. Akibatnya, pengambilan keputusan bisnis masih sering dilakukan tanpa dasar analisis data yang kuat. Kondisi ini memicu banyaknya pelaku umkm kesulitan dalam menentukan strategi bisnis digital yang paling efektif untuk meningkatkan kinerja usaha mereka.

Seiring dengan perkembangan teknologi, penggunaan artificial intelligence khususnya machine learning menjadi salah satu solusi yang bisa dimanfaatkan guna membantu proses analisis data serta pengambilan keputusan secara lebih objektif [14]. Teknologi ini memungkinkan proses analisis dilaksanakan dengan efisien dan tepat ketimbang pendekatan lama. Machine learning mempunyai kapabilitas untuk memproses data dalam jumlah besar serta menemukan pola yang dapat digunakan untuk memprediksi berbagai kondisi bisnis, termasuk dalam memprediksi keberhasilan strategi bisnis digital umkm [15], [3].

Beberapa penelitian sebelumnya telah menerapkan beragam algoritma

machine learning contohnya naïve bayes serta support vector machine ketika menelaah kinerja umkm [16]. Temuan studi tersebut mengungkap bahwa pendekatan berbasis algoritma mampu memberikan hasil prediksi yang cukup akurat. Namun, penelitian yang secara khusus menerapkan algoritma decision tree untuk memprediksi tingkat keberhasilan strategi bisnis digital umkm masih relatif terbatas [17]. Selain itu, penelitian sebelumnya lebih banyak berfokus pada prediksi penjualan atau kepuasan pelanggan, bukan pada keberhasilan strategi bisnis digital secara menyeluruh [18].

Berdasarkan permasalahan diatas, kajian tersebut dilakukan guna mengimplementasikan algoritma decision tree dalam memprediksi tingkat keberhasilan strategi bisnis digital pada umkm. Algoritma decision tree dipilih karena memiliki kemampuan dalam melakukan klasifikasi data serta menghasilkan model prediksi yang mudah dipahami melalui struktur pohon keputusan [19], [20].

Dengan memanfaatkan beberapa variabel yang berkaitan dengan strategi bisnis digital seperti pendapatan, pembelian online, kunjungan website, dan kunjungan respon kampanye pemasaran digital, diharapkan penelitian ini bisa membantu mengenali faktor-faktor yang memengaruhi keberhasilan strategi bisnis digital umkm serta memberikan dasar pengambilan keputusan yang lebih tepat bagi para pelaku usaha [21], [22].

METODE

Penelitian ini menerapkan pendekatan kuantitatif dengan teknik data mining untuk menganalisis informasi pelanggan serta meramalkan reaksi terhadap strategi pemasaran digital dengan menggunakan algoritma Decision Tree. Data yang digunakan merupakan data sekunder yang diambil dari dataset Analisis Kepribadian Pelanggan yang tersedia di platform Kaggle, dengan total

jumlah data mencapai 2. 240 pelanggan. Dataset tersebut mengandung rincian tentang karakteristik demografi pelanggan, pola belanja, aktivitas digital, serta reaksi pelanggan terhadap kampanye pemasaran yang telah dilakukan sebelumnya. Penelitian ini dilaksanakan pada bulan Februari 2026 dengan menggunakan komputer serta software Microsoft Excel dan RapidMiner Studio sebagai alat utama untuk pengolahan data dan pembangunan model klasifikasi.

Proses penelitian dimulai dengan pengumpulan data dari sumber yang telah ditentukan, dilanjutkan dengan pemahaman data untuk memahami struktur dataset, tipe atribut, dan hubungan antarvariabel yang relevan dalam penelitian. Setelah itu, tahap persiapan data dilakukan yang mencakup pembersihan untuk menangani nilai yang hilang, penghapusan data yang berulang, serta perbaikan data yang tidak konsisten untuk meningkatkan kualitas data yang akan dianalisis. Selanjutnya, dilakukan pemilihan atribut untuk menentukan variabel yang relevan dengan tujuan penelitian, yaitu Income, NumWebPurchases, NumCatalogPurchases, NumStorePurchases, NumWebVisitsMonth, AcceptedCmp1 hingga AcceptedCmp5, serta atribut Response yang berfungsi sebagai label atau target klasifikasi.

Data yang telah dipilih kemudian diubah ke dalam format yang sesuai dengan kebutuhan algoritma melalui proses encoding dan penyederhanaan atribut agar lebih mudah untuk diproses. Selain itu, dilakukan reduksi data pada atribut yang kurang relevan untuk meningkatkan efisiensi saat pemodelan berlangsung. Untuk mengatasi ketidakseimbangan antara jumlah data di setiap kelas, diimplementasikan metode Synthetic Minority Over-sampling Technique (SMOTE) sehingga distribusi data menjadi lebih merata. Dataset yang telah diproses kemudian dibagi menjadi dua bagian: 70% data training untuk

membangun model dan 30% data testing untuk menguji performa model.

Tahapan pemodelan dilakukan dengan menggunakan algoritma Decision Tree yang membuat struktur pohon keputusan berdasarkan nilai information gain tertinggi dari setiap atribut. Algoritma ini dipilih dikarenakan kemampuannya untuk menghasilkan model klasifikasi yang mudah dipahami serta dapat mengidentifikasi atribut yang paling berpengaruh terhadap reaksi pelanggan. Setelah model terbentuk, dilakukan evaluasi melalui confusion matrix untuk mengukur tingkat akurasi, presisi, dan recall dari model yang dihasilkan.

Hasil evaluasi kemudian diinterpretasikan untuk menemukan faktor-faktor paling dominan yang memengaruhi keberhasilan strategi pemasaran digital sekaligus menjadi dasar untuk pengambilan keputusan bisnis yang lebih efektif. Seluruh fase penelitian dilakukan dengan sistematis hingga menghasilkan model klasifikasi yang dapat membantu dalam menganalisis perilaku pelanggan dan meramalkan reaksi terhadap kampanye pemasaran digital dengan lebih tepat.

HASIL DAN PEMBAHASAN

Analisis Data Penelitian

Data yang digunakan dalam kajian adalah data sekunder dari dataset analisis kepribadian pelanggan di platform Kaggle. Terdapat 2.240 data pelanggan dengan 21 atribut tentang karakteristik pelanggan, perilaku pembelian, aktivitas digital, dan respons terhadap kampanye pemasaran. Penelitian ini menggunakan variabel Income, NumWebPurchases, NumCatalogPurchases, NumStorePurchases, NumWebVisitsMonth, AcceptedCmp1 hingga AcceptedCmp5, dengan atribut Response sebagai target klasifikasi. Atribut Response menunjukkan respons pelanggan terhadap kampanye pemasaran terakhir. Dataset ini digunakan untuk membentuk model klasifikasi

menggunakan algoritma Decision Tree untuk menganalisis dan memprediksi perilaku pelanggan terhadap strategi pemasaran digital.

Pengolahan Data

Tahap pengolahan data diimplementasikan guna menyiapkan dataset menjelang diterapkan pada tahapan klasifikasi menggunakan algoritma Decision Tree. Proses pengolahan data dimulai dengan mengimpor dataset ke dalam RapidMiner Studio, kemudian dilanjutkan dengan pembersihan data untuk menangani missing value menggunakan operator Replace Missing Values. Selanjutnya, dilaksanakan pemilihan atribut guna memilih faktor yang sesuai, transformasi data numerik menjadi bentuk biner true dan false menggunakan metode Numerical to Binomial, serta penentuan atribut Response sebagai label klasifikasi menggunakan operator Set Role. Melalui tahapan tersebut, data menjadi lebih terstruktur dan siap digunakan pada proses pemodelan dan evaluasi Decision Tree.

Penerapan Metode Decision Tree Perhitungan Entropy Dan Information Gain

Perhitungan gain serta informasi entropy dilakukan untuk menentukan atribut terbaik dalam pembentukan model Decision Tree. Proses tersebut dimaksudkan guna mengidentifikasi karakteristik yang paling relevan berpengaruh dalam mengklasifikasikan keberhasilan strategi pemasaran digital berdasarkan variabel Response sebagai label.

Variabel Response memiliki dua kelas, yaitu:

1. **Benar** = pelanggan menanggapi kampanye
2. **Salah** = pelanggan tidak merespons

Perhitungan Dataset Entropi

Sebelum menghitung entropy, dilakukan agregasi data dengan operator Aggregate di RapidMiner Studio untuk mengetahui jumlah data pada setiap kelas

atribut Response. Hasil ini digunakan untuk menghitung probabilitas setiap kelas sebagai dasar perhitungan entropy, yang menilai ketidakcocokan data. Semakin besar skor entropy, maka data semakin kurang teratur atau bercampur, sedangkan semakin rendah nilai entropy menunjukkan bahwa data semakin homogen.

Berlandaskan temuan pengolahan data, didapat jumlah data dibawah ini:

1. Jumlah data dengan kelas 0 (false) sebanyak 1906
2. Jumlah data dengan kelas 1 (true) sebanyak 334
3. Total data sebanyak 2240

Selanjutnya dihitung probabilitas masing-masing kelas:

1. Probabilitas kelas 0: $p(0) = \frac{1906}{2240} \approx 0,851$
2. Probabilitas kelas 1: $p(1) = \frac{334}{2240} \approx 0,149$

Entropy dihitung menggunakan rumus berikut:

$$Entropy(S) = -p(0) \log_2 p(0) - p(1) \log_2 p(1)$$

$$\text{Substitusi nilai probabilitas: } Entropy(S) = - (0,851 \cdot \log_2 0,851) - (0,149 \cdot \log_2 0,149)$$

$$\text{Hasil perhitungan: } Entropy(S) \approx 0,607$$

Nilai entropy sebesar 0,607 menunjukkan bahwa dataset memiliki tingkat ketidakpastian yang sedang. Artinya, data masih mengandung campuran antara kelas 0 dan kelas 1, namun tidak sepenuhnya acak. Dengan kondisi ini, dataset masih cukup baik untuk digunakan dalam proses klasifikasi menggunakan algoritma seperti Decision Tree.

Perhitungan Entropy Setiap Atribut

Pada tahap ini, dilakukan perhitungan entropy untuk setiap atribut guna mengetahui kemampuan atribut dalam memisahkan data ke dalam kelas yang lebih homogen. Atribut yang

dihitung adalah Income, NumWebPurchases, NumStorePurchases, NumCatalogPurchases, Num Web Visits Month, serta AcceptedCmp1 hingga AcceptedCmp5, dengan Response sebagai label klasifikasi. Data ditransformasikan ke dalam bentuk binominal menjadi dua kategori yaitu range1 dan range2. Perhitungan entropy dilakukan dengan menghitung jumlah data berdasarkan kelas Response dan menentukan probabilitas setiap kelas untuk memperoleh nilai entropy. Nilai entropy digunakan untuk mengukur tingkat ketidakpastian data.

Entropy Total AcceptedCmp5

Row No.	Response	AcceptedCmp5	count(Resp...
1	false	range1 [-∞ - 0.500]	1835
2	true	range1 [-∞ - 0.500]	242
3	false	range2 [0.500 - ∞]	71
4	true	range2 [0.500 - ∞]	92

Gambar 1 Hasil agregasi atribut AcceptedCmp5

Tabel 1 Hasil Perhitungan Entropy Atribut

Atribut	Entropy
Income	0,594
NumWebPurchases	0,588
NumCatalogPurchases	0,583
NumStorePurchases	0,604
NumWebVisitsMonth	0,607
AcceptedCmp1	0,564
AcceptedCmp2	0,594
AcceptedCmp3	0,574
AcceptedCmp4	0,588
AcceptedCmp5	0,554

Berdasarkan analisis terhadap nilai entropy Atribut AcceptedCmp5 memiliki nilai entropy terendah, sehingga lebih efektif membedakan kelas data dan digunakan dalam perhitungan Information Gain untuk Decision Tree.

Perhitungan Information Gain dan Pemilihan Node root

Perhitungan Information Gain dilakukan untuk memilih atribut terbaik sebagai node akar di algoritma Decision

Tree. Nilai gain dihitung dari selisih entropy total dataset dan entropy masing-masing atribut. Entropy total dihitung dari jumlah data Response "true" dan "false". Entropy atribut dihitung setelah proses discretize dan aggregate di RapidMiner.

Rumus Information Gain yang digunakan adalah: $Gain(S, A) =$

$$Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Entropy(S_i)$$

Keterangan:

- **Gain (S, A)** = nilai information gain atribut A
- **Entropy(S)** = entropy total dataset
- **|S_i|** = jumlah data pada partisi ke-i
- **|S|** = jumlah seluruh data

Tabel 2 perhitungan informasi gain

No	Atribut	Entropy atribut	Gain
1	Income	0,594	0,013
2	NumWebPurchases	0,588	0,019
3	NumCatalogPurchases	0,583	0,024
4	NumStorePurchases	0,604	0,003
5	NumWebVisitsMonth	0,607	0,000
6	AcceptedCmp1	0,564	0,043
7	AcceptedCmp2	0,594	0,013
8	AcceptedCmp3	0,574	0,033
9	AcceptedCmp4	0,588	0,019
10	AcceptedCmp5	0,554	0,053

Berlandaskan temuan kalkulasi information gain, atribut AcceptedCmp5 mempunyai skor gain paling tinggi yakni sebesar 0,053.

Tabel 3 Hasil Algoritma Aturan Klasifikasi Pohon Keputusan

Rule	Kondisi (IF)	Prediksi (THEN)	True	False	Interpretasi
1	AcceptedCmp5 = range1, AcceptedCmp3 = range1, AcceptedCmp2 = range1, AcceptedCmp1 = range1, NumWebPurchases = range1, AcceptedCmp4 = range1, NumCatalogPurchases = range1	False	218	651	Pelanggan yang tidak menerima kampanye pemasaran sebelumnya serta memiliki tingkat pembelian melalui website dan katalog yang rendah cenderung tidak memberikan respons terhadap promosi yang dilakukan perusahaan.

Diketahui:

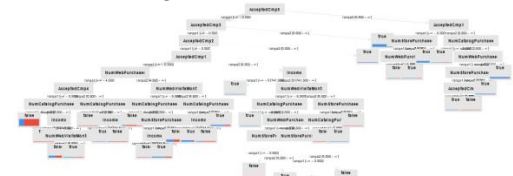
1. Entropy total dataset = 0,607
2. Entropy atribut AcceptedCmp5 = 0,554

Maka: $Gain(AcceptedCmp5) = 0,607 - 0,554 = 0,053$

Nilai gain yang lebih besar menunjukkan bahwa atribut memiliki kemampuan lebih baik dalam memisahkan data berdasarkan kelas Response. Atribut AcceptedCmp5 ditetapkan sebagai root node dalam pohon keputusan, sementara NumWebVisitsMonth memiliki nilai gain terendah sebesar 0,000, menunjukkan pengaruh kecil dalam klasifikasi data. Setelah menetapkan node akar, proses pengembangan cabang dan node dapat dilakukan untuk menghasilkan model Decision Tree yang optimal.

Pembentukan Pohon Keputusan (Decision Tree)

Ilustrasi berikut memperlihatkan hasil pembentukan pohon keputusan menggunakan algoritma Decision Tree berdasarkan atribut menggunakan skor information gain terbesar.



Gambar 2 Tampilan Hasil pembentukan pohon keputusan Decision Tree

2	AcceptedCmp5 = range1, AcceptedCmp3 = range2	True	246	54	Pelanggan yang menerima kampanye pemasaran ketiga memiliki kecenderungan tinggi untuk memberikan respons positif terhadap promosi berikutnya.
3	AcceptedCmp5 = range2, AcceptedCmp1 = range2, NumCatalogPurchases = range2, NumWebPurchases = range2	True	145	2	Pelanggan yang aktif melakukan pembelian melalui katalog dan website serta memiliki riwayat penerimaan kampanye pemasaran sebelumnya memiliki kemungkinan sangat tinggi untuk memberikan respons positif terhadap promosi.

Berdasarkan rule yang dihasilkan, atribut *AcceptedCmp5* menjadi faktor utama dalam menentukan klasifikasi pelanggan. Selain itu, atribut seperti *NumWebPurchases*, *NumCatalogPurchases*, *NumStorePurchases*, dan *AcceptedCmp3* juga berpengaruh dalam menentukan respon pelanggan terhadap promosi perusahaan. Pelanggan dengan aktivitas pembelian tinggi dan riwayat penerimaan kampanye sebelumnya cenderung memberikan respon positif terhadap promosi yang diberikan.

Implementasi Decision Tree

Pada fase pelaksanaan, model klasifikasi dibuat menggunakan algoritma Decision Tree dengan perangkat lunak RapidMiner Studio, berdasarkan data pelanggan yang telah diolah. Proses dimulai dengan mengambil dataset melalui operator Retrieve Data dan menangani nilai yang hilang menggunakan Replace Missing Values dengan metode rata-rata. Atribut yang relevan dipilih melalui Select Attributes. Data numerik kemudian diubah menjadi biner dengan operator Numerical to Binomial. Atribut Response ditetapkan sebagai label klasifikasi menggunakan Set Role. Untuk mengatasi ketidakseimbangan data, digunakan metode SMOTE. Dataset dibagi dengan

rasio 70% untuk data training dan 30% untuk data testing. Model dibangun berdasarkan nilai information gain, lalu diuji dan dievaluasi menggunakan operator Performance (Classification). Hasil menunjukkan bahwa model dapat mengklasifikasikan data pelanggan dengan baik.

Hasil Model Decision Tree

Pada tahap ini dilakukan pengujian model Decision Tree menggunakan data testing yang telah dipisahkan sebelumnya melalui proses Split Data. Model yang telah dibentuk kemudian dievaluasi menggunakan operator Performance (Classification) pada RapidMiner untuk mengetahui tingkat performa klasifikasi. Hasil pengujian model ditunjukkan pada Gambar berikut.

accuracy: 74.48%

	true true	true false	class precision
pred. true	391	111	77.65%
pred. false	181	461	71.81%
class recall	66.36%	80.55%	

Gambar 3 Hasil Evaluasi Model Menggunakan Algoritma Decision Tree

Pengujian Model

Pengujian model dilaksanakan memakai data testing sebesar 30% dari total dataset. Proses pengujian dimaksudkan guna mengetahui kapasitas model Decision Tree pada mengklasifikasikan data pelanggan

berdasarkan atribut yang digunakan. Operator Apply Model dipakai guna menerapkan model Decision Tree ke data testing, sedangkan Performance (Classification) diimplementasikan kedalam perhitungan nilai precision, recall, accuracy, serta confusion matrix.

Evaluasi Model

Berlandaskan temuan pengujian menggunakan RapidMiner didapati hasil yaitu:

Tabel 4 Hasil Evaluasi Model

Metode evaluasi	Nilai
Accuracy	74,48%
Precision kelas true	77,89%
precision kelas false	71,81%
Recall kelas true	68,36%
Recall kelas false	80,59%

Tabel 5 confusion matrix

Prediksi	True	False
true	391	111
false	181	461

Interpretasi Hasil

Berdasarkan penilaian model, tingkat akurasi yang diperoleh adalah 74,48%. Algoritma Decision Tree berhasil mengkategorikan data pelanggan dengan baik. Nilai presisi untuk kelas positif mencapai 77,89%, menandakan bahwa model ini memiliki tingkat kesalahan prediksi positif yang rendah. Recall untuk kelas positif adalah 68,36%, menunjukkan bahwa model mengenali banyak pelanggan yang merespons positif, namun masih ada yang tidak terdeteksi. Recall untuk kelas negatif adalah 80,59%, menunjukkan efisiensi yang lebih tinggi dalam mengidentifikasi pelanggan yang tidak merespons. Model ini efektif dalam klasifikasi perilaku pelanggan, terutama dalam merespons kampanye pemasaran digital perusahaan.

Pembahasan

Beracuan kepada capaian uji algoritma Decision Tree, diperoleh nilai accuracy sebesar 74,48%, menunjukkan

model bisa mengklasifikasikan perilaku pelanggan dengan baik. Atribut AcceptedCmp5 dipilih sebagai root node karena memiliki information gain tertinggi. Selain itu, atribut AcceptedCmp3, NumWebPurchases, NumCatalogPurchases, dan NumStorePurchases juga berpengaruh. Model lebih baik dalam mengklasifikasikan kelas false dengan nilai recall 80,59%. dibandingkan kelas true sebesar 68,36%. Secara keseluruhan, algoritma Decision Tree cukup efektif digunakan untuk mengidentifikasi perilaku pelanggan berdasarkan riwayat pembelian dan respon terhadap kampanye pemasaran digital.

SIMPULAN

Berdasarkan penelitian, algoritma Decision Tree efektif untuk mengklasifikasikan reaksi pelanggan terhadap kampanye pemasaran digital dengan 2.240 data konsumen. Penelitian dilakukan melalui tahapan seperti pengolahan data, pemilihan fitur, penyeimbangan data menggunakan SMOTE, perhitungan entropy dan information gain, pembuatan Decision Tree, dan evaluasi model menggunakan RapidMiner Studio. Model yang dihasilkan mencapai akurasi 74,48%, menunjukkan keberhasilan algoritma ini dalam mengklasifikasikan perilaku pelanggan berdasarkan karakteristik dan respons kampanye. Atribut AcceptedCmp5 memiliki nilai information gain tertinggi dan dipilih sebagai node akar. Untuk penelitian selanjutnya, disarankan menggunakan data lebih banyak, atribut relevan, serta membandingkan dengan algoritma lain seperti Random Forest, Naive Bayes, dan XGBoost. Pengolahan data dan pengoptimalan parameter model juga perlu ditingkatkan.

DAFTAR PUSTAKA

- T. Budiman and B. Rahman, “Pemanfaatan Machine Learning Dalam Mengoptimalkan Strategi Penjualan Berbasis E-Commerce Untuk Meningkatkan Pendapatan Umk,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 3, pp. 3917–3923, 2025, doi:10.36040/jati.v9i3.13463.
- M. Cerdas, A. Telukdarie, S. Philbin, B. G. Mwanza, and M. Munsamy, “Platform Digital untuk Pemberdayaan UMKM,” vol. 00, pp. 1–9, 2022.
- A. Rahardi, R. Aprilia Andini, A. S. Tri Agil, G. Ramadhan, and D. Salim, “Pemanfaatan Machine Learning untuk Prediksi Kepuasan Pelanggan pada UMKM Digital,” *J. Data Sci. Methods Appl.*, vol. 01, no. 02, pp. 66–74, 2025, doi:10.30873/jodmapps.v1i2.pp66-74.
- N. A. Khotamin, W. Damayanti, and M. Aula, “Pendampingan Konversi Bisnis Manual Ke Digital Bisnis Pada Usaha Kecil Menengah di Kecamatan Metro Utara,” *Educommunity J. Pengabd. Masy.*, vol. 2, no. 2, pp. 50–57, 2024, doi: 10.71365/ejpm.v2i2.59.
- A. L. Kilay, B. H. Simamora, and P. Putra, “Pengaruh Layanan E-Payment dan E-Commerce terhadap Kinerja Rantai Pasokan : Implikasi Inovasi dan Solusi Terbuka terhadap Digitalisasi Usaha Mikro , Kecil , dan Menengah (UMKM) di Indonesia,” 2022.
- L. Hakim, A. Luthfi, and A. Hasyim, “Konsep Pemasaran Digital dan Konvensional UMKM (Literatur Studi di Kabupaten Mojokerto),” *ALTAMKINJurnal Ekon. Berbas. Pemberdaya. Masy.*, vol. 1, no. 1, pp. 16–22, 2025.
- B. S. Nora, A. Mukoffi, and R. Risnaningsih, “Pengaruh Perkembangan Ekonomi Digital terhadap Kelangsungan Usaha Kecil Menengah Kota Malang,” vol. 3, no. 2, pp. 82–97, 2025, [Online]. Available: <https://rinjani.unitri.ac.id/handle/071061/4483%0Ahttps://rinjani.unitri.ac.id/bitstream/handle/071061/4483/artikel.pdf?sequence=1&isAllowed=y>
- G. Santoso, M. Rizal, H. Wiyana, and S. N. Subagja, “Digitalisasi UMKM : Strategi Dan Model Bisnis Berbasis Teknologi Untuk Keberlanjutan,” *JUBISDIGI J. Bisnis Digit.*, vol. 01, no. 01, pp. 21–30, 2025.
- X. Zhang and Y. Xu, “Penelitian Faktor Keberhasilan dan Mekanisme yang Mempengaruhi Transformasi Digital pada UKM,” 2022.
- S. Ramadhan, Z. Alamin, Miftahul Jannah, Muhammad Akbar, and Rizki Fikriyansah, “Data-driven MSME Success Prediction Using Decision Tree-Based Machine Learning Techniques,” *Journix J. Informatics Comput.*, vol. 1, no. 1, pp. 1–9, 2025, doi: 10.63866/journix.v1i1.3.
- Sutisna and M. Nurur Raudhan, “Strategi Pengambilan Keputusan Berbasis Data Analytics pada UMKM untuk Meningkatkan Penjualan,” *Data Sci. Indones.*, vol. 5, no. 2, pp. 38–49, 2025, doi: 10.47709/dsi.v5i2.6768.
- A. Hutagaol *et al.*, “Analisis Pengaruh Digitalisasi Terhadap Umkm Di Kota Medan,” *J. Ekuilnomi*, vol. 6, no. 3, pp. 729–738, 2024, doi: 10.36985/kgvmq881.
- I. Kamil, A. A. Bakri, S. Salingkat, A. Ardenny, J. P. Tahirs, and A. Alfiana, “Pendampingan UMKM melalui Pemanfaatan Digital Marketing pada Platform E-Commerce,” *Amalee Indones. J.*

- Community Res. Engagem.*, vol. 3, no. 2, pp. 517–526, 2022, doi: 10.37680/amalee.v3i2.2782.
- R. Suarantalla, “Peran Artificial Intelligence (AI) dalam Optimalisasi Pemasaran Digital pada Usaha Mikro, Kecil, dan Menengah (UMKM),” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 8480–8491, 2025, doi: 10.31004/riggs.v4i3.3277.
- M. Qasim, M. Khan, W. Mehmood, F. Sobieczky, M. Pichler, and B. Moser, “A Comparative Analysis of Anomaly Detection Methods for Predictive Maintenance in SME,” *Commun. Comput. Inf. Sci.*, vol. 1633 CCIS, pp. 22–31, 2022, doi: 10.1007/978-3-031-14343-4_3.
- M. Harahap, B. P. A. Sihombing, O. A. F. Laia, B. T. Saragih, F. Teknologi, and U. P. Indonesia, “ANALISIS SENTIMEN REVIEW PENJUALAN PRODUK UMKM PADA MACHINE LEARNING Kuandi Dharma,” *METHOMIKA J. Manaj. Inform. Komputerisasi Akunt.*, vol. 5, no. 2, pp. 147–154, 2021.
- A. Saepuloh, P. A. Sanusi, and E. Rilvani, “PT. Media Akademik Publisher PREDIKSI KEBERHASILAN USAHA KECIL MENENGAH MENGGUNAKAN CREDAL C4.5 DAN CREDAL DECISION TREE: ANALISIS KOMPARATIF,” *Jma*, vol. 3, no. 7, pp. 3031–5220, 2025, [Online]. Available: <https://doi.org/10.12345/jeb.2021.15.2.112>
- S. Rizal, P. Studi, T. Informatika, and U. Yudharta, “Penerapan Algoritma Naïve Bayes Untuk Prediksi Penerimaan,” vol. 10, no. 1, pp. 14–21, 2018.
- F. Satria, “Analisis Data Mining Strategi Digital Marketing terhadap Keputusan Pembelian Mahasiswa,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 9, no. 2, p. 427, 2025, doi: 10.26798/jiko.v9i2.1910.
- M. S. Rahayu and B. -Barati, “Prediksi Perilaku Pembelian Konsumen di Platform E-Commerce Menggunakan Algoritma Decision Tree,” no. 19, pp. 1–5, 2025, [Online]. Available: <https://www.researchgate.net/publication/393631892>
- L. Nur Khasanah, I. Nafiyah, N. Khomsah, and U. K. Abdurrahman Wahid Pekalongan, “Peran Digital Marketing Bagi Umkm Di Masa Pandemi,” *J. Sahmiyya*, vol. 1, no. 2, 196–202, 2022, [Online]. Available: <https://e-journal.uingusdur.ac.id/index.php/sahmiyya/article/view/58>
- S. A. Qalati, D. Ostic, M. A. B. A. Sulaiman, A. A. Gopang, and A. Khan, “Social Media and SMEs’ Performance in Developing Countries: Effects of Technological-Organizational-Environmental Factors on the Adoption of Social Media,” *SAGE Open*, vol. 12, no. 2, 2022, doi: 10.1177/215824402210945