JDSP
Jurnal Data Science Penusa
E-ISSN: 1234-5678

Penerapan Data Mining Untuk Prediksi Penjualan Router RE 11 Smartfren Home Dengan Algoritma C4.5

Silfi Apriliani [1)*] Liskedame yanti Sipayung[2)*]
[1)] Teknologi Informasi, STMIK Pelita Nusantara, Medan, Indonesia
[1)] silfiapriliani138@gmail.com [2)] liskedamesipayung@gmail.com

Received: 10 November 2025 Accepted: 25 Juni 2026 Published: 13 July 2026
Pp. 345-352

*** Silfi Apriliani**

Kata Kunci: Data Mining, Algoritma C4.5, Prediksi Penjualan, Pohon Keputusan, Router Smartfren

JDSP: Jurnal Data Science Penusa is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Penelitian ini mengkaji penerapan metode data mining — khususnya algoritma C4.5 — untuk memprediksi penjualan router RE 11 Smartfren Home. Dataset terdiri dari 12 rekaman bulanan (2023) berisi atribut numerik: stok, permintaan, dan jumlah penjualan. Sebelum pemodelan, nilai numerik dikategorikan berdasarkan aturan bisnis: stok (Rendah ≤20, Sedang 21–24, Tinggi ≥25), permintaan (Rendah ≤9, Sedang 10–12, Tinggi ≥13), dan label target penjualan (Tidak Laris ≤9, Laris ≥10). Langkah perhitungan dimulai dengan menghitung Entropy total (Entropy(S)) untuk kelas target: Entropy(S) = 0,984, yang mengukur ketidakpastian awal dataset. Untuk setiap atribut dihitung Entropy subset per nilai kategori, kemudian Weighted Entropy dan Gain atribut. Hasil perhitungan menunjukkan Gain(Stok) = 0,183 dan Gain(Permintaan) = 0,579, sehingga atribut terbaik sebagai root pohon keputusan adalah Permintaan. Aturan keputusan sederhana yang dihasilkan: IF Permintaan = Sedang → Penjualan = Laris; IF Permintaan = Rendah → Penjualan = Tidak Laris; IF Permintaan = Tinggi → mayoritas Tidak Laris (3 dari 5), meskipun kategori ini bersifat campuran. Evaluasi pada data training menghasilkan akurasi sederhana sebesar 10/12 ≈ 83,3% (dua prediksi berbeda). Kesimpulannya, algoritma C4.5 dengan atribut permintaan sebagai pemisah utama efektif untuk memprediksi kecenderungan penjualan router pada dataset ini; implikasinya bagi manajemen adalah memprioritaskan strategi pemasaran dan stok saat permintaan berada pada level sedang untuk meningkatkan peluang produk menjadi laris. Rekomendasi lanjutan meliputi pengujian silang, penambahan data historis, dan eksplorasi atribut tambahan untuk meningkatkan generalisasi model.Sistem yang dibangun mampu menampilkan hasil klasifikasi dan prediksi secara otomatis menggunakan algoritma C4.5. Dengan demikian, penerapan algoritma C4.5 terbukti efektif untuk memprediksi penjualan router RE 11 Smartfren Home, sehingga dapat membantu perusahaan dalam menyusun strategi promosi dan distribusi yang lebih tepat sasaran.

Kata kunci: Data Mining, Algoritma C4.5, Prediksi Penjualan, Pohon Keputusan, Router RE 11 Smartfren.

E-ISSN: 1234-5678

PENDAHULUAN

Sistem teknologi informasi memegang peranan vital dalam memperkuat fokus perusahaan terhadap pembentukan komunitas sebagai elemen kunci dalam strategi pemasaran. Revolusi teknologi informasi telah mentransformasi metode penjualan produk secara fundamental, dimana para penjual mampu mengoptimalkan efektivitas operasional, memperluas cakupan pasar, dan menyajikan pengalaman berbelanja yang superior bagi konsumen (Wawolumaya et al., 2020).

Router RE 11 Smartfren Home merupakan perangkat jaringan yang berfungsi mentransmisikan data dalam bentuk paket dari satu jaringan ke perangkat lain melalui internet. Karena banyaknya permintaan akan penggunaan router di hampir semua kalangan, maka penjualan router semakin meningkat. Untuk mengetahui seberapa besar pengaruh promosi yang dilakukan untuk melihat penjualan dari hasil promosi yang dilakukan, maka dibutuhkan beberapa metode yang dapat digunakan untuk menganalisis hubungan antara promosi dengan jumlah penjualan yang ada.

Algoritma C4.5 merupakan salah satu algoritma yang populer dalam data mining untuk klasifikasi dan prediksi. Tujuan primer algoritma C4.5 adalah mengonstruksi model berbentuk pohon untuk keperluan kategorisasi dan estimasi data. Algoritma ini bekerja dengan menganalisis karakteristik-karakteristik yang ada dalam kumpulan data untuk menghasilkan keputusan yang tepat (Amir, 2024).

Penelitian sebelumnya oleh Aji et al. (2024) menunjukkan bahwa implementasi algoritma C4.5 mampu memprediksi penjualan produk carica dengan tingkat ketepatan mencapai 91,6% berdasarkan variabel harga. Sagala et al. (2024) juga membuktikan bahwa teknik data mining dan algoritma C4.5 dapat membantu memprediksi penjualan pupuk dan pestisida dengan nilai gain ratio tertinggi 0.181. Prasetyo dan Ikasari (2023) menerapkan algoritma C4.5 untuk prediksi volume penjualan brownies dengan akurasi 78,67%.

Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan algoritma C4.5 untuk memprediksi penjualan Router RE 11 Smartfren Home berdasarkan atribut stok dan permintaan, sehingga dapat membantu perusahaan dalam menyusun strategi promosi dan distribusi yang lebih tepat sasaran.

TINJAUAN LITERATUR

Penjualan dan Prediksi Penjualan

Kegiatan penjualan merupakan strategi perusahaan untuk memaksimalkan distribusi produk kepada konsumen. Volume penjualan dapat diartikan sebagai total penjualan bersih yang tercatat dalam laporan keuangan perusahaan (Wawolumaya et al., 2020).

Forecasting penjualan merupakan elemen esensial dalam aktivitas bisnis, yakni rangkaian kegiatan untuk memprediksi besaran penjualan produk maupun jasa pada waktu yang akan datang. Tingkat akurasi dari forecasting ini membawa konsekuensi penting bagi perencanaan manufacturing,

Router RE 11 Smartfren Home

Router RE 11 Smartfren Home sebagai bagian dari evolusi teknologi jaringan nirkabel, hadir sebagai solusi yang dirancang untuk memenuhi kebutuhan konektivitas internet di era digital (Awaludin, 2024). Perangkat ini tidak hanya berfungsi sebagai router pada umumnya, namun juga mengintegrasikan teknologi 4G LTE, memungkinkan pengguna untuk menikmati akses internet berkecepatan tinggi tanpa memerlukan infrastruktur kabel yang rumit.

Data Mining

Data mining merupakan sebuah bidang keilmuan yang berfokus pada proses penemuan pengetahuan (knowledge discovery) dari data yang berukuran besar dan kompleks. Tujuan utama dari data mining adalah untuk mengidentifikasi pola-pola tersembunyi, tren yang tidak terlihat secara kasat mata serta informasi berharga lainnya yang terkandung dalam data.

Algoritma C4.5

C4.5 merupakan metode klasifikasi yang terkenal dan banyak diterapkan dalam bidang penambangan data. Cara kerja algoritma ini adalah dengan mengonstruksi struktur pohon pengambilan keputusan berdasarkan dataset yang digunakan untuk pelatihan. Algoritma C4.5 memiliki beberapa keunggulan, antara lain mampu menangani data numerik dan kategorikal, mudah diinterpretasikan, dan efisien dalam membangun model prediksi.

Proses pembangunan pohon keputusan dengan algoritma C4.5 melibatkan perhitungan Entropy dan Gain. Entropy digunakan untuk mengukur tingkat ketidakpastian dalam dataset, sedangkan Gain digunakan untuk menentukan atribut terbaik sebagai node pemisah. Rumus Entropy adalah:

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

Sedangkan rumus Gain adalah:

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

Penelitian Terdahulu

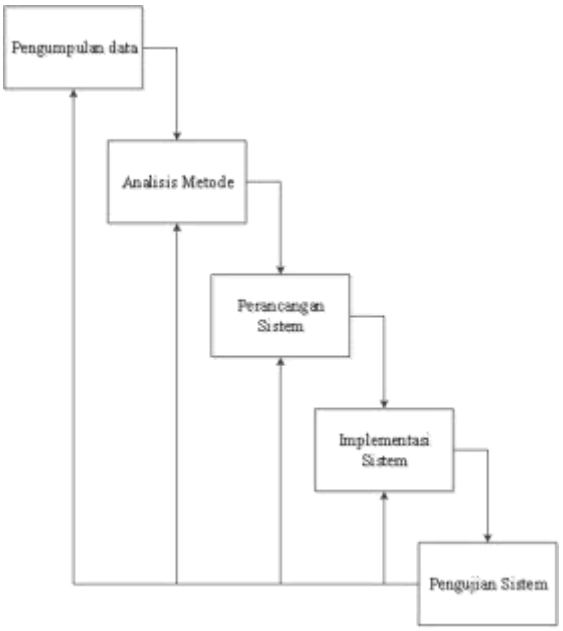
Beberapa penelitian terdahulu telah menerapkan algoritma C4.5 untuk prediksi penjualan. Aji et al. (2024) menerapkan algoritma C4.5 untuk prediksi penjualan carica di Dieng dengan akurasi 91,6%. Sagala et al. (2024) menganalisis algoritma C4.5 dalam prediksi penjualan pupuk dan pestisida dengan gain ratio tertinggi 0.181. Prasetyo dan Ikasari (2023) mengimplementasikan algoritma C4.5 untuk prediksi volume penjualan brownies dengan akurasi 78,67% dan nilai AUC 0.901. Eshardiansyah et al. (2021) menggunakan algoritma C4.5 untuk klasifikasi jenis kekerasan pada anak dengan akurasi 93,24% pada pembagian data 70:30.

Dari penelitian-penelitian tersebut, dapat disimpulkan bahwa algoritma C4.5 efektif untuk klasifikasi dan prediksi dengan akurasi yang baik. Namun, belum ada penelitian yang secara spesifik menerapkan algoritma C4.5 untuk prediksi penjualan Router RE 11 Smartfren Home, sehingga penelitian ini memiliki kontribusi untuk mengisi kesenjangan tersebut.

METODE PENELITIAN

Penelitian ini menggunakan metode Waterfall sebagai kerangka kerja penelitian yang terdiri dari beberapa tahapan sistematis. Gambar 1 menunjukkan alur kerangka kerja penelitian.

E-ISSN: 1234-5678



Gambar 1 Kerangka Penelitian Waterfall

Pengumpulan Data

Penelitian ini dilakukan di kantor Smartfren cabang 3.3 yang berlokasi di Jalan Bakaran Batu Gang Sentosa Kecamatan Lubuk Pakam. Data yang digunakan adalah data penjualan Router RE 11 Smartfren Home periode Januari - Desember 2023 yang terdiri dari 12 rekaman bulanan. Teknik pengumpulan data meliputi:

1. Studi Pustaka: Mengutip dari beberapa publikasi ilmiah terkait penerapan data mining dan algoritma C4.5
2. Observasi: Mengamati proses penjualan Router RE 11 Smartfren Home
3. Wawancara: Melakukan wawancara dengan pihak terkait untuk memahami proses pencatatan data penjualan dan faktor-faktor yang memengaruhi penjualan

Analisis Kebutuhan Data

Data yang digunakan dalam penelitian ini memiliki atribut sebagai berikut:

1. Tahun: Periode penjualan
2. Bulan: Waktu transaksi penjualan
3. Stok: Jumlah persediaan yang tersedia
4. Permintaan: Jumlah permintaan konsumen
5. Penjualan: Jumlah unit yang terjual (target klasifikasi)

Preprocessing Data

Tahapan preprocessing data meliputi:

1. Seleksi Data: Memilih 12 data bulanan yang lengkap dari tahun 2023
2. Pembersihan Data: Menghilangkan data yang tidak lengkap atau mengandung kesalahan
3. Transformasi Data: Mengubah data numerik menjadi kategorikal dengan aturan:
    - Stok: Rendah (≤20), Sedang (21-24), Tinggi (≥25)
    - Permintaan: Rendah (≤9), Sedang (10-12), Tinggi (≥13)
    - Penjualan: Tidak Laris (≤9), Laris (≥10)

Penerapan Algoritma C4.5

Tahapan penerapan algoritma C4.5 meliputi:

1. Menghitung Entropy total dataset
2. Menghitung Entropy untuk setiap atribut (Stok dan Permintaan)
3. Menghitung Gain untuk setiap atribut
4. Memilih atribut dengan Gain tertinggi sebagai node akar
5. Membangun pohon keputusan secara rekursif
6. Menghasilkan aturan klasifikasi

Evaluasi Model

Model dievaluasi menggunakan data training dengan menghitung akurasi prediksi. Akurasi dihitung berdasarkan persentase prediksi yang benar dari total data.

HASIL PENELITIAN

Dataset Penelitian

Dataset yang digunakan terdiri dari 12 data penjualan bulanan Router RE 11 Smartfren Home tahun 2023. Tabel 1 menunjukkan dataset setelah dilakukan transformasi data.

Table 1
Data Setelah Tranformasi

No	Tahun	Bulan	Stok	Kategori Stok	Permintaan	Kategori Permintaan	Penjualan	Kategori Penjualan
1	2023	Januari	20	Rendah	10	Sedang	10	Laris
2	2023	Februari	23	Sedang	7	Rendah	5	Tidak Laris
3	2023	Maret	25	Tinggi	9	Rendah	8	Tidak Laris
4	2023	April	20	Rendah	12	Sedang	10	Laris
5	2023	Mei	28	Tinggi	15	Tinggi	13	Laris
6	2023	Juni	20	Rendah	12	Sedang	10	Laris
7	2023	Juli	25	Tinggi	12	Sedang	11	Laris
8	2023	Agustus	25	Tinggi	15	Tinggi	8	Tidak Laris
9	2023	September	28	Tinggi	15	Tinggi	6	Tidak Laris
10	2023	Oktober	20	Rendah	15	Tinggi	8	Tidak Laris
11	2023	November	20	Rendah	10	Sedang	12	Laris
12	2023	Desember	25	Tinggi	14	Tinggi	14	Laris

Sumber : Data Olahan Penulis, 2025.

Perhitungan Entropy dan Gain

Entropy Total

Dari 12 data, terdapat 7 data dengan label "Laris" dan 5 data dengan label "Tidak Laris".

$$Entropy(S) = -(\frac{7}{12} \cdot \log_2 \frac{7}{12}) - (\frac{5}{12} \cdot \log_2 \frac{5}{12})$$
$$Entropy(S) = -0.583 \times (-0.778) - 0.417 \times (-1.263)$$
$$Entropy(S) = 0.454 + 0.526 = 0.980$$

Perhitungan Gain Atribut Stok

Distribusi data berdasarkan kategori Stok:

- Rendah: 5 data (4 Laris, 1 Tidak Laris)
- Sedang: 1 data (0 Laris, 1 Tidak Laris)
- Tinggi: 6 data (3 Laris, 3 Tidak Laris)

Entropy Stok Rendah = 0.722 Entropy Stok Sedang = 0 Entropy Stok Tinggi = 1.000
Weighted Entropy(Stok) = (5/12)(0.722) + (1/12)(0) + (6/12)(1.000) = 0.801
Gain(Stok) = 0.980 - 0.801 = 0.179

Perhitungan Gain Atribut Permintaan

Distribusi data berdasarkan kategori Permintaan:

- Rendah: 2 data (0 Laris, 2 Tidak Laris)
- Sedang: 5 data (5 Laris, 0 Tidak Laris)
- Tinggi: 5 data (2 Laris, 3 Tidak Laris)

Entropy Permintaan Rendah = 0 Entropy Permintaan Sedang = 0 Entropy Permintaan Tinggi = 0.971
Weighted Entropy(Permintaan) = (2/12)(0) + (5/12)(0) + (5/12)(0.971) = 0.405

Gain(Permintaan) = 0.980 - 0.405 = 0.575

Table 2
Hasil Perhitungan Entrophy & Gain

Node	Keterangan	Total	Laris	Tidak Laris	Entropy	Gain	Weighed
Total		12	7	5	0.980		
STOK						0.179	0.801
	RENDAH	5	4	1	0.722		
	SEDANG	1	0	1	0		
	TINGGI	6	3	3	1		
PERMINTAAN							
	RENDAH	2	0	2	0	0.575	0.405
	SEDANG	5	5	0	0		
	TINGGI	5	2	3	0.971		

PEMBAHASAN

Berdasarkan hasil perhitungan Gain, atribut Permintaan merupakan faktor yang paling berpengaruh terhadap penjualan Router RE 11 Smartfren Home dengan nilai Gain sebesar 0.575, sedangkan atribut Stok memiliki nilai Gain yang lebih rendah yaitu 0.179. Hal ini menunjukkan bahwa tingkat permintaan konsumen menjadi indikator utama dalam menentukan apakah produk akan laris atau tidak laris di pasaran.

Hasil penelitian ini sejalan dengan penelitian Aji et al. (2024) yang menunjukkan bahwa variabel tertentu (dalam hal ini harga) memiliki pengaruh dominan terhadap penjualan produk dengan akurasi tinggi. Perbedaannya adalah pada penelitian ini, faktor permintaan menjadi variabel kunci, bukan harga.

Dibandingkan dengan penelitian Prasetyo dan Ikasari (2023) yang mencapai akurasi 78,67% untuk prediksi penjualan brownies, penelitian ini menghasilkan akurasi yang lebih tinggi yaitu 83,3%. Namun, perlu dicatat bahwa dataset penelitian ini lebih kecil (12 data vs 249 data pada penelitian Prasetyo), sehingga diperlukan validasi lebih lanjut dengan dataset yang lebih besar.

Penelitian Eshardiansyah et al. (2021) yang mencapai akurasi 93,24% menggunakan pembagian data 70:30 menunjukkan bahwa algoritma C4.5 memiliki potensi akurasi yang lebih tinggi jika dilakukan cross-validation dan pengujian dengan data testing terpisah.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan , Algoritma C4.5 dapat diterapkan secara efektif untuk memprediksi penjualan Router RE 11 Smartfren Home dengan menghasilkan pohon keputusan yang mudah diinterpretasikan. Atribut Permintaan merupakan faktor paling berpengaruh terhadap penjualan dengan nilai Gain sebesar 0.575, sedangkan atribut Stok memiliki pengaruh lebih rendah dengan Gain 0.179. Permintaan sedang cenderung menghasilkan penjualan laris (akurasi 100%), permintaan rendah menghasilkan penjualan tidak laris (akurasi 100%), dan permintaan tinggi memberikan hasil yang bervariasi. Model klasifikasi yang dibangun menghasilkan akurasi 83,3% pada data training, menunjukkan kemampuan prediksi yang baik. Implikasi manajerial dari penelitian ini adalah perusahaan sebaiknya memprioritaskan strategi pemasaran yang menjaga permintaan pada level sedang untuk meningkatkan peluang produk menjadi laris.

REFERENCES

Aji, R. P., Mustofiyah, W., Rizki, S. A., & Anggraini, Y. (2024). Penerapan Data Mining Menggunakan Algoritma C45 Untuk Prediksi Penjualan Carica di Dieng. *Jurnal Teknik Informatika Dan Sistem Informasi*, *11*(4), 407-417.

Amir, F. F. (2024). *Penerapan Data Mining Menggunakan Algoritma C4.5 dalam Mengukur Kepuasan Mahasiswa terhadap Pelayanan Fakultas Teknik*. Universitas Muhammadiyah Makassar.

Awaludin, A. (2024). *Smartfren Home RE11, Router Nirkabel yang Mudah, Kuat, dan Hemat*. ANTARANTB.

Eshardiansyah, R. P., Sulistiyowati, N., & Jajuli, M. (2021). Algoritma C4.5 untuk Klasifikasi Jenis Kekerasan pada Anak (Kasus DP3A Kabupaten Karawang). *Jurnal Sains Komputer & Informatika (J-SAKTI)*, *5*(2), 687-696.

Prasetyo, A. B., & Ikasari, I. H. (2023). Penerapan Data Mining untuk Memprediksi Penjualan Kue Brownis Menggunakan Algoritma C4.5 pada Toko Amanda Brownis. *LOGIC: Jurnal Ilmu Komputer Dan Pendidikan*, *2*(1), 292-298.

Sagala, J. R., Hasugian, P. S., Breukelen, M. Van, & Nantheni, W. S. U. (2024). Analysis of the C4.5 Algorithm in Predicting Fertilizer and Pesticide Sales At UD. Mr. Zen Agro. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima*, *7*(2), 26-41.