



Prediksi Status Pesanan Menggunakan Metode Classification C.45 Pada Toko Stuftech.Id di Shopee

Muhamad Arief Firdaus¹, Fadli Rahman Latarissa², Yanuar Dzaky³, Hidayanti Murtina⁴

¹⁻⁴ Universitas Bina Sarana Informatika

Email author: firdausmuhamadarief@gmail.com¹, fadlilarissa3017@gmail.com², yanuar.dzaky03@gmail.com³, hidayati.hym@bsi.ac.id⁴

Article Info

Article history:

Received Januari 3, 2025

Revised Februari 17, 2025

Accepted June 28, 2025

Keywords:

C4.5 Algorithm

Classification

Shopee

Order Prediction

E-Commerce

ABSTRACT

The increase in transactions on e-commerce platforms such as Shopee necessitates the development of accurate order status prediction systems to optimize services and reduce order cancellations and delays. This study aims to build a classification model for predicting order status (completed or canceled) at StufTech.Id store using the C4.5 algorithm. The dataset consists of transaction attributes such as payment method, shipping location category, and shipping cost. The classification process was conducted using RapidMiner through preprocessing, decision tree generation, and model evaluation stages. The analysis shows that the "Shipping Region Category" attribute has the highest information gain and was selected as the root node. The resulting model achieved an accuracy of 86%, with 100% recall for completed orders but only 6.67% for canceled ones. These findings indicate that the C4.5 algorithm is effective in predicting successful transactions but requires improvement in identifying potential cancellations. Implementing this model can help business owners make proactive operational decisions.

Corresponding Author:

Muhamad Arief Firdaus,

Universitas Bina Sarana Informatika

Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat

Email: firdausmuhamadarief@gmail.com



1. PENDAHULUAN

Perkembangan teknologi digital telah mendorong transformasi besar-besaran dalam berbagai sektor, termasuk perdagangan ritel melalui platform e-commerce. Salah satu pemain besar dalam sektor ini di Asia Tenggara adalah Shopee, yang mencatat peningkatan transaksi luar biasa dalam beberapa tahun terakhir (Santoso et al., 2024). Menurut laporan statistika 2024, Shopee mencatat 228 juta klik bulanan di Indonesia, menjadikannya platform e-commerce dengan trafik tertinggi di kawasan (Nurhayati-Wolff, 2025). Dinamika ini menunjukkan bahwa transaksi digital kini bukan hanya sekadar pelengkap, tetapi telah menjadi tulang punggung perdagangan modern. Namun, seiring dengan meningkatnya volume transaksi, tantangan dalam pengelolaan data pesanan juga kian kompleks. Salah satu tantangan terbesar adalah bagaimana memprediksi status pesanan (seperti "selesai"/"batal") secara akurat untuk meningkatkan efisiensi operasional dan kepuasan pelanggan. Di sinilah pentingnya

penerapan rapid miner dan algoritma klasifikasi seperti C4.5 untuk mengelola dan memproyeksikan data transaksi secara otomatis dan cerdas (Dutalia et al., 2021).

Penelitian ini secara khusus berfokus pada prediksi status pesanan pada toko Stuftech.Id yang beroperasi di Shopee, dengan memanfaatkan metode klasifikasi C4.5 sebagai pendekatan utama. Algoritma C4.5 merupakan pengembangan dari ID3 yang lebih stabil dan efisien, dan sangat sesuai untuk data bertipe kategorikal maupun numerik. Dalam konteks prediksi status pesanan, berbagai variabel seperti metode pembayaran, lokasi pengiriman, dan ongkos kirim dapat dianalisis untuk menghasilkan keputusan prediktif. Hubungan antara variabel-variabel ini sangat erat, karena mampu mencerminkan pola perilaku konsumen serta performa logistik yang kompleks. Melalui pendekatan ini, diharapkan sistem dapat mengklasifikasikan pesanan secara otomatis dan mengantisipasi keterlambatan atau pembatalan, sehingga dapat membantu pelaku usaha untuk mengambil keputusan secara proaktif (Jatmika & Taufi, 2021).

Fenomena utama yang menjadi latar belakang penelitian ini adalah tingginya angka pembatalan dan keterlambatan pengiriman pesanan yang masih terjadi secara signifikan dalam e-commerce Indonesia. Menurut data dari Databoks, sekitar 23% dari pengguna e-commerce di Indonesia mengalami keterlambatan pengiriman dan 14% mengalami pembatalan pesanan karena kendala sistem atau logistik (Ahdiat, 2023). Oleh karena itu, penting adanya sistem cerdas yang mampu memprediksi status pesanan sejak dini berdasarkan pola transaksi sebelumnya. Khususnya bagi toko-toko skala menengah seperti Stuftech.Id, yang menghadapi tekanan kompetitif tinggi, sistem prediksi semacam ini dapat menjadi kunci diferensiasi layanan.

Tujuan utama dari penelitian ini adalah untuk membangun model prediksi status pesanan pada toko Stuftech.Id di Shopee dengan memanfaatkan metode klasifikasi C4.5 secara komprehensif. Penelitian ini memiliki urgensi yang tinggi, baik dari sisi akademik maupun praktis. Secara akademik, penelitian ini memberikan kontribusi terhadap pengembangan pemodelan prediktif dalam sistem e-commerce Indonesia dengan pendekatan yang lebih aplikatif dan kontekstual. Sementara secara praktis, hasil dari penelitian ini dapat diterapkan langsung oleh pelaku UMKM digital untuk meningkatkan efisiensi pengelolaan transaksi dan kepercayaan pelanggan. Ruang lingkup penelitian meliputi pengumpulan data transaksi pesanan dari toko Stuftech.Id, pengolahan data dengan metode preprocessing, pembangunan model klasifikasi dengan algoritma C4.5, serta evaluasi akurasi dan interpretasi hasil prediksi.

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi berbasis algoritma C4.5. Rancangan penelitian yang diterapkan adalah penelitian eksperimental kuantitatif, di mana data transaksi pesanan dianalisis menggunakan perangkat lunak RapidMiner untuk membangun model prediksi status pesanan.

2.1. Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data primer berupa riwayat transaksi pesanan dari toko Stuftech.Id yang beroperasi di platform e-commerce Shopee. Dataset mencakup atribut seperti metode pembayaran, ongkos kirim, status pesanan, dan kategori pulau.

Table 1. Data Transaksi Stuftech,id

No	Status Pesanan	Metode Pembayaran	Ongkos Kirim	Kategori Pulau
1	Selesai	Online Payment	9.000	Pulau Jawa
2	Selesai	Saldo ShopeePay	22.000	Pulau Jawa
3	Selesai	COD (Bayar di Tempat)	17.000	Pulau Jawa
4	Selesai	Online Payment	20.000	Pulau Jawa
5	Selesai	COD (Bayar di Tempat)	17.000	Pulau Jawa
6	Selesai	Online Payment	9.000	Pulau Jawa

.....
1550	Batal	COD (Bayar di Tempat)	37.000	Luar Pulau Jawa
1551	Selesai	Spaylater	9.000	Luar Pulau Jawa

2.2. Teknik Pengumpulan Data

Data diperoleh langsung dari pemilik toko Stuftech.Id dalam format CSV. Data kemudian dibersihkan dan disesuaikan untuk keperluan proses klasifikasi.

2.3. Teknik Analisis Data

Tahapan analisis data dan alur proses analisis dalam RapidMiner ditunjukkan pada Gambar 1 berikut:

1. Preprocessing Data

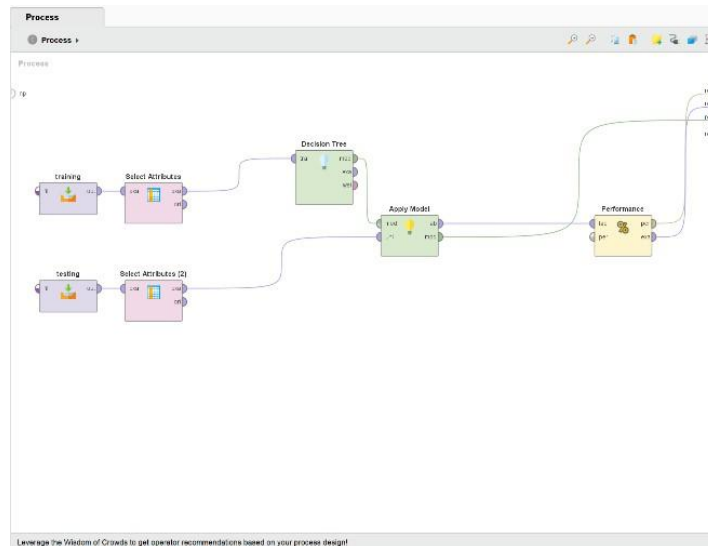
Meliputi pembersihan data dari nilai kosong (missing value), penyesuaian format data, transformasi variabel kategorikal menjadi numerik (jika diperlukan), dan seleksi atribut. Proses ini dilakukan untuk memastikan data siap digunakan dalam proses klasifikasi.

2. Pengembangan Model Klasifikasi

Algoritma C4.5 digunakan untuk membentuk decision tree berdasarkan atribut-atribut yang tersedia. Algoritma ini dipilih karena kemampuannya dalam menangani data kontinu, mengelola nilai yang hilang, serta menyediakan proses pruning untuk menghindari overfitting.

3. Evaluasi Model

Evaluasi dilakukan dengan teknik *performance classification* untuk mengukur akurasi, precision, dan recall dari model yang dibangun. Hasil evaluasi ini menjadi dasar dalam mengukur performa dan keandalan prediksi status pesanan.



Gambar 1. Diagram Proses Eksperimen di RapidMiner

Alat Bantu Penelitian

Perangkat lunak yang digunakan dalam penelitian ini adalah RapidMiner Studio, yang mendukung proses data mining secara visual dan intuitif. RapidMiner dipilih karena fleksibilitasnya dalam membangun model klasifikasi, melakukan evaluasi model, serta mendukung berbagai format input data.

3. HASIL DAN ANALISIS

Hasil Klasifikasi Menggunakan Algoritma C4.5

Model klasifikasi dibangun menggunakan algoritma C4.5 melalui perangkat lunak RapidMiner. Atribut utama yang digunakan dalam model prediksi status pesanan adalah Kategori Pulau, Metode Pembayaran, dan Ongkos Kirim. Hasil training menunjukkan struktur pohon keputusan yang cukup mendalam, dan mampu menangkap pola-pola penting dari data transaksi.

- a) Menghitung *Entropy* Total :
- $$= \left(-\frac{1332}{1551} \times \log_2 \left(\frac{1332}{1551}\right)\right) + \left(-\frac{219}{1551} \times \log_2 \left(\frac{219}{1551}\right)\right)$$
- $$= 0,5873$$
- b) Menghitung *Gain* jumlah :
- Entropy* [Jumlah - Pulau Jawa]

$$\begin{aligned}
&= \left(-\frac{1073}{1551} \times \log_2\left(\frac{1073}{1551}\right)\right) + \left(-\frac{146}{1551} \times \log_2\left(\frac{146}{1551}\right)\right) \\
&= 0,5287 \\
&\text{Entropy [Jumlah - Luar Pulau Jawa]} \\
&= \left(-\frac{259}{1551} \times \log_2\left(\frac{259}{1551}\right)\right) + \left(-\frac{73}{1551} \times \log_2\left(\frac{73}{1551}\right)\right) \\
&= 0,7599 \\
&= \text{Gain} = \text{Entropy (s)} - \left(\frac{1219}{1551} \times 0,5287 + \frac{332}{1551} \times 0,7599\right) \\
&= 0,0091
\end{aligned}$$

Tabel berikut merangkum hasil perhitungan manual nilai entropy dan gain untuk masing-masing atribut:

Table 2. Perhitungan *Node 1*

Node 1		Selesai	Batal	Jumlah Data	Entropy Total (S)	Gain
		1332	219	1551	0,5873708814	
Metode Pembayaran	COD	331	77	408	0,6987976574	0,008892108331
	Online Payment	390	68	458	0,6060022694	
	Saldo ShopeePay	276	29	305	0,4532090425	
	SPayLater	248	30	278	0,4935907028	
	SeaBank Bayar Instan	46	8	54	0,6051865766	
	Kartu Kredit/Debit	31	4	35	0,512709142	
	Indomaret/i.Saku	6	2	8	0,8112781245	
	Alfamart/Alfamidi/Dan+Dan	1	1	2	1	
	BCA OneKlik	3	0	3	0	
Kategori Pulau	Luar Pulau Jawa	259	73	332	0,7599474455	0,009170909948
	Pulau Jawa	1073	146	1219	0,5287002493	
Ongkos Kirim	<=20.440	957	139	1096	0,5486654829	0,002859648922
	>20.440	375	80	455	0,670856159	
	<=17.000	818	117	935	0,5439396402	
	>17.000	514	102	616	0,6475034187	

Berdasarkan Tabel 2, atribut Kategori Pulau memiliki nilai gain tertinggi yaitu 0,0091, sehingga atribut ini dipilih sebagai node akar dalam pembangunan pohon keputusan.

Table 3. Perhitungan *Node 1.1*

Kategori Pulau - Luar Jawa	Selesai	Batal	Jumlah Data	Entropy Total (S)	Information Gain

	TOTAL DATA	259	73	332	0,7599474455	
Luar Jawa - Metode Pembayaran	COD	108	45	153	0,8739810481	0,07791705733
	Online Payment	56	17	73	0,7829924502	
	Saldo ShopeePay	40	0	40	0	
	SPayLater	47	7	54	0,5564215673	
	SeaBank Bayar Instan	4	2	6	0,9182958341	
	Kartu Kredit/Debit	2	1	3	0,9182958341	
	Indomaret/i.Saku	2	1	3	0,9182958341	
	Alfamart/Alfamidi /Dan+Dan	0	0	0	0	
	BCA OneKlik	0	0	0	0	
Luar Jawa - Ongkos Kirim	<=20.440	34	13	47	0,8507707041	0,00213063701
	>20.440	225	60	285	0,7424875695	2

Dari hasil perhitungan pada Tabel 3, atribut Metode Pembayaran kembali menjadi atribut dengan nilai gain tertinggi, yaitu 0,0779, sehingga dipilih sebagai node berikutnya.

Table 4. Perhitungan *Node* 1.1.1

Metode Pembayaran - Seabank Bayar Instan	Selesai	Batal	Jumlah Data	Entropy Total (S)	Gain
TOTAL DATA	2	1	3	0,9182958341	
Metode Pembayaran - Seabank Bayar Instan <=20.440	0	1	1	0	0,9182958341
>20.440	2	0	2	0	

Hasil analisis pada Tabel 4 menunjukkan bahwa jika metode pembayaran menggunakan SeaBank Bayar Instan, maka:

Jika ongkos kirim > 20.440, maka transaksi diprediksi Selesai.

Jika ongkos kirim ≤ 20.440, maka transaksi diprediksi Batal.

Hasil Visualisasi Pohon Keputusan

Berdasarkan hasil visualisasi pohon keputusan (Gambar 2), atribut pertama yang menjadi pembeda utama dalam proses klasifikasi adalah Kategori Pulau. Kemudian diikuti oleh atribut Metode Pembayaran dan Ongkos Kirim, yang memiliki pengaruh signifikan terhadap status pesanan (Selesai atau Batal). Hal ini menunjukkan bahwa faktor geografis dan jenis pembayaran memiliki peran penting dalam keberhasilan transaksi.

Tree

```
Kategori Pulau = Luar Pulau Jawa
| Metode Pembayaran = COD (Bayar di Tempat)
| | Ongkos Kirim = <=20.440: Selesai {Selesai=19, Batal=4}
| | Ongkos Kirim = >20.440: Selesai {Selesai=89, Batal=41}
| Metode Pembayaran = Indomaret/i.Saku
| | Ongkos Kirim = <=20.440: Batal {Selesai=0, Batal=1}
| | Ongkos Kirim = >20.440: Selesai {Selesai=2, Batal=0}
| Metode Pembayaran = Kartu Kredit/Debit: Selesai {Selesai=2, Batal=1}
| Metode Pembayaran = Online Payment
| | Ongkos Kirim = <=20.440: Batal {Selesai=4, Batal=5}
| | Ongkos Kirim = >20.440: Selesai {Selesai=52, Batal=12}
| Metode Pembayaran = SPayLater
| | Ongkos Kirim = <=20.440: Selesai {Selesai=5, Batal=2}
| | Ongkos Kirim = >20.440: Selesai {Selesai=42, Batal=5}
| Metode Pembayaran = Saldo ShopeePay: Selesai {Selesai=40, Batal=0}
| Metode Pembayaran = SeaBank Bayar Instan
| | Ongkos Kirim = <=20.440: Batal {Selesai=0, Batal=1}
| | Ongkos Kirim = >20.440: Selesai {Selesai=4, Batal=1}
Kategori Pulau = Pulau Jawa
| Metode Pembayaran = Alfamart/Alfamidi/Dan+Dan: Selesai {Selesai=1, Batal=1}
| Metode Pembayaran = BCA OneKlik: Selesai {Selesai=3, Batal=0}
| Metode Pembayaran = COD (Bayar di Tempat): Selesai {Selesai=223, Batal=32}
| Metode Pembayaran = Indomaret/i.Saku
| | Ongkos Kirim = <=20.440: Selesai {Selesai=4, Batal=0}
| | Ongkos Kirim = >20.440: Batal {Selesai=0, Batal=1}
| Metode Pembayaran = Kartu Kredit/Debit
| | Ongkos Kirim = <=20.440: Selesai {Selesai=10, Batal=3}
| | Ongkos Kirim = >20.440: Selesai {Selesai=19, Batal=0}
| Metode Pembayaran = Online Payment
| | Ongkos Kirim = <=20.440: Selesai {Selesai=272, Batal=39}
| | Ongkos Kirim = >20.440: Selesai {Selesai=62, Batal=12}
| Metode Pembayaran = SPayLater
| | Ongkos Kirim = <=20.440: Selesai {Selesai=169, Batal=18}
| | Ongkos Kirim = >20.440: Selesai {Selesai=32, Batal=5}
| Metode Pembayaran = Saldo ShopeePay
| | Ongkos Kirim = <=20.440: Selesai {Selesai=206, Batal=28}
| | Ongkos Kirim = >20.440: Selesai {Selesai=30, Batal=1}
| Metode Pembayaran = SeaBank Bayar Instan
| | Ongkos Kirim = <=20.440: Selesai {Selesai=37, Batal=5}
| | Ongkos Kirim = >20.440: Selesai {Selesai=5, Batal=1}
```

Gambar 2. Hasil Output Text Pohon Keputusan (Simplifikasi Logika Klasifikasi)

Sebagai contoh, untuk pelanggan di Luar Pulau Jawa yang memilih metode pembayaran COD dengan ongkos kirim di atas Rp20.440, maka probabilitas pesanan Selesai cenderung lebih tinggi ({Selesai=89, Batal=41}). Sebaliknya, untuk metode pembayaran Indomaret/i.Saku dengan ongkos kirim kurang dari Rp20.440, justru berakhir Batal ({Selesai=0, Batal=1}). Pola serupa juga terjadi pada pelanggan dari Pulau Jawa, di mana metode pembayaran seperti Saldo ShopeePay dan Online Payment memiliki jumlah keberhasilan transaksi yang tinggi.

Visualisasi Pohon Keputusan

Visualisasi decision tree (Gambar 3) memperkuat hasil klasifikasi dengan menyajikan cabang-cabang keputusan secara visual. Node utama dimulai dari atribut Kategori Pulau, kemudian bercabang ke Metode Pembayaran, dan selanjutnya ke nilai Ongkos Kirim. Daun (leaf) pada pohon menampilkan keputusan akhir (Selesai/Batal) beserta distribusi jumlah data. Tampak bahwa sebagian besar node terminal menghasilkan keputusan “Selesai”, menunjukkan dominasi kelas ini dalam dataset.

Pola yang terbentuk dari visualisasi menunjukkan bahwa keberhasilan pesanan sangat berkaitan dengan nilai ongkos kirim yang tinggi (> Rp20.440), terutama bila dikombinasikan dengan metode pembayaran digital seperti ShopeePay, SPayLater, dan Online Payment.

Shopee dengan tingkat akurasi sebesar 86%. Atribut yang paling berpengaruh dalam proses klasifikasi adalah Kategori Pulau, diikuti oleh Metode Pembayaran dan Ongkos Kirim. Model ini menunjukkan performa yang sangat baik dalam mengklasifikasikan pesanan yang selesai, dengan recall mencapai 100%. Namun demikian, masih terdapat kelemahan dalam mendeteksi pesanan yang batal, ditunjukkan dengan recall yang rendah (6,67%). Oleh karena itu, untuk penerapan di masa mendatang, disarankan pengembangan model lanjutan dengan algoritma yang lebih sensitif terhadap kelas minoritas, atau dengan pendekatan data balancing agar model lebih seimbang dalam mengklasifikasikan kedua kategori status pesanan.

DAFTAR PUSTAKA

- [1] Aditya Nugroho, C. R., & Kristiana, T. (2022). Penerapan Algoritma C4.5 Untuk Kepuasan Pelanggan Toko Online Parfume Chantik. *Jurnal Algoritme*, 3(1), 10–21. <https://doi.org/10.35957/algoritme.v3i1.3169>
- [2] Ahdiat, A. (2023). Ini Kendala Usaha E-Commerce di Indonesia menurut Survei BPS. *Databoks*.
- [3] Apriyadi, A., Lubis, M. R., & Damanik, B. E. (2022). Penerapan Algoritma C5.0 Dalam Menentukan Tingkat Pemahaman Mahasiswa Terhadap Pembelajaran Daring. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, 11(1), 11–20. <https://doi.org/10.34010/komputa.v11i1.7386>
- [4] Arhami, M., & Nasir, M. (2020). *Data Mining - Algoritma dan Implementasi*. CV Andi Offset.
- [5] Dutalia, S., Lalo, A. K., Batarius, P., Carmeneja, Y., & Siki, H. (2021). Implementasi Algoritma C4 . 5 Untuk Klasifikasi Penjualan. 06, 1–12.
- [6] Jatmika, D., & Taufi, G. (2021). Penerapan Algoritma C4.5 Untuk Klasifikasi Keberhasilan Pengiriman Barang. 8(1), 36–43.
- [7] Muktar, R., Syam, M., Kunda, A., & Natsir, M. (2023). Analisis Penerapan Data Mining dalam Klasifikasi Penjualan Pakaian Pada Toko Online Shopee Menggunakan Algoritma C4.5. *Jurnal Dipanegara Komputer Sistem Informasi*, XVII(1), 1–8.
- [8] Novrianto, & Elisa, E. (2024). ANALISIS KLASIFIKASI ALGORITMA C.45 DALAM MEMPREDIKSI TINGKAT PENJUALAN PERLENGKAPAN IBADAH. *Comasie*, 6(2), 107–118.
- [9] Nurhayati-Wolff, H. (2025). Situs e-commerce terbesar di Indonesia pada tahun 2024, berdasarkan lalu lintas bulanan. *Statista*.
- [10] Santoso, N. A., Redaputri, A. P., & Rizkyna, S. (2024). JURNAL CEMERLANG : Pengabdian pada Masyarakat PEMANFAATAN E - COMMERCE SHOPEE SEBAGAI STRATEGI JURNAL CEMERLANG : Pengabdian pada Masyarakat. 7(1), 98–110.
- [11] Steffany, A., Aldzikri, K. A., Tricahya, M. A. R., & Arfiandono, M. I. (2023). Penerapan Algoritma C4.5 Untuk Menganalisis Kepuasan E-Commerce Shopee. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 6(6), 700–706. <https://doi.org/10.32672/jnkti.v6i6.7112>
- [12] Suripto, Nurul Rahmanita, R., Kirana, A. S., & 2022. (2022). Teknik pre-processing dan classification dalam data science. *Student Corner*.
- [13] Wibawa, A. P., Guntur, M., Purnama, A., Fathony Akbar, M., & Dwiyanto, F. A. (2018). Metode-metode Klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi*, 3(1), 134–138.