ANALISIS PERILAKU *CART ABANDONMENT* PADA TOKO *ONLINE* CREATROKA MENGGUNAKAN ALGORITMA *RANDOM FOREST* UNTUK OPTIMASI PEMASARAN LAYANAN

ANALYSIS OF CART ABANDONMENT BEHAVIOR IN CREATROKA ONLINE STORE USING RANDOM FOREST ALGORITHM FOR SERVICE MARKETING OPTIMIZATION

Fauzan Zikri[1], Diny Syarifah Sany[2]

[1,2] Universitas Suryakancana
[1] fauzanzikri@unsur.ac.id, [2] dsy.sany@gmail.com

Masuk: 21 Mei 2026	Penerimaan: 23 Juni 2026	Publikasi: 30 Juni 2026

ABSTRAK

Persaingan industri *e-commerce* menuntut pelaku usaha, khususnya pada sektor literasi dan edukasi anak, untuk mengoptimalkan tingkat konversi penjualan di tengah tingginya angka *cart abandonment*. Penelitian ini bertujuan untuk memprediksi perilaku pembatalan pesanan dan mengidentifikasi faktor dominan yang memengaruhinya pada toko *online* Creatroka. Penelitian ini dilakukan dengan berfokus pada rekam jejak transaksi operasional selama rentang periode 1 Februari hingga 30 April 2026 menggunakan toko *online* Creatroka di Cianjur, di mana proses ekstraksi dan pengolahan datanya dilakukan secara digital. Penelitian ini menggunakan pendekatan kuantitatif dengan metode penambangan data (*data mining*) terhadap 10.429 data transaksi selama periode Februari hingga April. Teknik analisis data dilakukan melalui pra-pemrosesan berbasis bahasa pemrograman *Python* dan pemodelan klasifikasi menggunakan algoritma *Random Forest*. Hasil penelitian menunjukkan bahwa model *Random Forest* memiliki kinerja prediktif yang andal dengan tingkat akurasi mencapai 76,16% dan presisi sebesar 80,47%. Analisis ekstraksi fitur (*Feature Importance*) mengungkapkan bahwa "Harga Produk" (36,63%) merupakan variabel pemicu utama pembatalan, diikuti oleh durasi interaksi visual atau "Jumlah Halaman Dilihat" (21,82%) dan sumber kunjungan dari "Fitur Video/Afiliasi" (15,74%). Temuan ini mengimplikasikan perlunya strategi pemasaran layanan (*service marketing*) yang proaktif, seperti penerapan sistem *Customer Relationship Management* (CRM) waktu-nyata untuk melakukan intervensi diskon bagi segmen pelanggan yang terdeteksi berada di jalur pembatalan dan optimalisasi konten visual pada halaman produk untuk menekan rasio *cart abandonment* secara sistematis.

Kata kunci: *Cart Abandonment, Online,* Pemasaran layanan*, Random Forest*, Toko Creatroka.

ABSTRACT

The intensity of e-commerce competition requires business actors, particularly in the children's literacy and education sector, to optimize conversion rates amidst high cart abandonment rates. This study aims to predict cart abandonment behavior and identify the dominant factors influencing it at the Creatroka online store. This study focused on the operational transaction records for the period from February 1 to April 30, 2026, using the Creatroka online store in Cianjur, where data extraction and processing were conducted digitally. This research employs a quantitative approach using a data mining method on 10,429 transaction records from February to April. Data analysis was conducted through Python-based preprocessing and classification modeling using the Random Forest algorithm. The results demonstrate that the Random Forest model exhibits reliable predictive performance with an accuracy rate of 76.16% and a precision of 80.47%. Feature Importance analysis reveals that "Product Price" (36.63%) is the primary driver of cancellation, followed by the duration of visual interaction or "Number of Pages Viewed" (21.82%) and traffic sources from "Video Features/Affiliates" (15.74%). These findings imply the necessity for proactive service marketing strategies, such as implementing real-time Customer Relationship Management (CRM) systems to trigger promotional interventions for segments detected on the path to abandonment, and optimizing visual content on product pages to systematically reduce the cart abandonment ratio.

Keywords: *Cart Abandonment, Creatroka Store, Online, Random Forest, Service Marketing,*

PENDAHULUAN

Pertumbuhan pesat industri ritel digital (*e-commerce*) telah mengubah secara fundamental pola interaksi antara pelaku usaha dan konsumen, terutama pada sektor literasi dan edukasi anak. Dalam ekosistem digital, keberhasilan sebuah bisnis tidak lagi sekadar diukur dari seberapa besar arus lalu lintas (*traffic*) pengunjung yang berhasil ditarik ke dalam halaman produk, melainkan dari seberapa efektif arus kunjungan tersebut dikonversi menjadi transaksi penjualan yang nyata (Nawir & Hendrawan, 2024; Sasmiharti & Nurhayati, 2025). Tantangan terbesar yang dihadapi oleh tata kelola manajemen bisnis digital saat ini adalah mengendalikan tingkat kebocoran konversi yang sering terjadi di tahap akhir proses berbelanja.

Fenomena di mana konsumen telah menunjukkan intensi pembelian dengan memasukkan produk ke dalam keranjang atau membuat pesanan, namun pada akhirnya memilih untuk tidak menyelesaikan pembayaran, dikenal luas dengan istilah *cart abandonment* (Sundjaja *et al.*, 2024). Secara global, riset menunjukkan bahwa lebih dari 70% hingga 80% keranjang belanja digital berakhir tanpa transaksi, yang merepresentasikan hilangnya potensi pendapatan (*revenue leak*) yang sangat signifikan bagi pelaku Usaha Mikro, Kecil, dan Menengah (UMKM) (Statista, 2026). Berdasarkan data operasional internal pada toko online Creatroka yang bergerak di bidang literasi dan buku edukasi anak, ditemukan adanya kesenjangan yang substansial antara metrik "Pesanan Dibuat" dan "Pesanan Dibayar" selama periode Februari hingga April.

Kesenjangan ini mengindikasikan adanya urgensi untuk menganalisis perilaku konsumen (*consumer behavior*) di balik pembatalan pesanan tersebut guna merumuskan strategi pemasaran layanan (*service marketing*) yang lebih adaptif dan presisi. Memahami faktor-faktor yang memicu cart abandonment membutuhkan evaluasi yang komprehensif terhadap berbagai variabel perilaku konsumen, seperti sumber kunjungan (pencarian organik, afiliasi, atau fitur video), jenis produk, dan tingkat harga. Beberapa penelitian terdahulu telah berupaya mengurai kompleksitas fenomena ini dari sudut pandang psikologi dan bisnis. Menurut Kapoor & Vij (2021) bahwa membedakan antara pembatalan keranjang belanja dan pembatalan saat *checkout*, di mana faktor kebijakan pengiriman dan kelelahan administratif menjadi pemicu utama.

Dalam studi lain, Sharma & Srivastava (2025) menemukan bahwa risiko yang dirasakan (*perceived risk*) oleh konsumen muda bertindak sebagai mediator yang kuat terhadap niat pembelian dan pembatalan keranjang. Dari perspektif manajemen tata letak e-ritel, menurut Jiang *et el.* (2021) menyoroti bahwa kelebihan beban pilihan (*choice overload*) dapat memicu kelupaan konsumen, sehingga intervensi visual diperlukan. Di Indonesia, kajian mengenai perilaku konsumen digital juga menekankan bahwa interaksi pemasaran layanan (*relationship*

marketing) berbasis kepercayaan harus diintegrasikan untuk menekan angka penolakan transaksi di menit akhir (Yunus *et al.* 2022; Fransiska eta al, 2025). Namun demikian, pendekatan analisis statistik tradisional atau evaluasi manual dinilai tidak lagi memadai untuk memetakan ribuan baris data rekam jejak (*log*) transaksi harian konsumen yang sangat dinamis (Chen, 2022). Oleh karena itu, diperlukan pendekatan analitik lintas disiplin yang memadukan keilmuan Administrasi Bisnis dengan Teknik Informatika melalui pemanfaatan algoritma *Machine Learning*.

Penerapan *Machine Learning* dalam domain *e-commerce* telah terbukti mampu mengekstraksi pola tersembunyi dari data perilaku (*clickstream* dan *session data*) untuk memprediksi niat pembelian konsumen (Setiawan *et al.,* 2026). Di antara berbagai model algoritma klasifikasi yang tersedia, Algoritma *Random Forest* secara konsisten menunjukkan kinerja yang superior dalam menangani dataset transaksi ritel digital. Random Forest merupakan metode *ensemble learning* yang membangun banyak pohon keputusan (*decision trees*) dan menggabungkan prediksinya untuk meningkatkan akurasi serta meminimalisasi *overfitting* (Rausch *et al*., 2022). Penelitian yang dilakukan oleh Alamsyah & Wahyuni (2024) membuktikan bahwa algoritma *Random Forest* mengungguli *Support Vector Machine* (SVM) dalam memprediksi niat beli pengunjung online dengan akurasi di atas 90%.

Hal ini senada juga dibuktikan oleh Esmeli & Gokce (2025) yang menggunakan *Artificial Intelligence* untuk menganalisis perilaku konsumen setelah produk ditambahkan ke keranjang, di mana integrasi algoritma klasifikasi efektif digunakan untuk mengurangi angka pembatalan. Berbeda dengan model algoritma *black-box* lainnya, *Random Forest* memiliki keunggulan berupa kemampuan untuk menghasilkan interpretasi *Feature Importance*, yaitu peringkat variabel mana yang paling berkontribusi terhadap hasil prediksi (Degenhardt, 2019). Hal ini sangat krusial bagi manajemen bisnis, karena output matematis dari algoritma tersebut dapat langsung diterjemahkan menjadi landasan perumusan strategi intervensi layanan.

Berdasarkan tinjauan pustaka dan identifikasi masalah pada toko *online* Creatroka, terlihat bahwa masih terdapat celah penelitian (*research gap*) terkait pemanfaatan algoritma klasifikasi untuk membedah pengaruh spesifik dari sumber kunjungan kekinian, seperti lalu lintas afiliasi (*affiliate*) dan fitur video interaktif, terhadap rasio *cart abandonment* produk edukasi. Oleh karena itu, penelitian ini bertujuan untuk mengukur tingkat akurasi prediksi Algoritma *Random Forest* dalam mengklasifikasikan status akhir transaksi digital (dibayar atau dibatalkan) menggunakan data operasional Creatroka.

Selanjutnya, penelitian ini juga bertujuan untuk menganalisis variabel atau fitur operasional mana yang paling dominan memengaruhi perilaku *cart abandonment* berdasarkan ekstraksi data komputasi. Pada akhirnya, integrasi temuan dari model analitik tersebut ditujukan

untuk merumuskan rekomendasi optimasi pemasaran layanan (*service marketing*) yang inovatif, yang memungkinkan manajemen Creatroka untuk melakukan mitigasi pembatalan secara *real-time* dan memulihkan potensi pendapatan bisnis di pasar yang kompetitif.

METODE PENELITIAN

Waktu dan Tempat Penelitian

Pelaksanaan penelitian berfokus pada rekam jejak transaksi operasional selama rentang periode 1 Februari hingga 30 April 2026, di mana proses ekstraksi dan pengolahan datanya dilakukan secara digital. Tempat penelitiannya dilakukan di toko *online* Creatroka yang berada di Cianjur.

Metode Penelitian

Penelitian ini menggunakan jenis penelitian kuantitatif terapan dengan pendekatan penambangan data (*data mining*). Menurut Wang & Huang (2022), pendekatan kuantitatif berbasis penambangan data sangat relevan digunakan untuk mengekstraksi pola tersembunyi dari dataset berskala besar guna mendukung pengambilan keputusan bisnis yang presisi. Penelitian ini dilaksanakan pada operasional toko *online* Creatroka, sebuah usaha percetakan dan penerbitan buku edukasi yang beroperasi di platform *e-commerce*.

Penelitian ini menggunakan sejumlah variabel bebas (X) yang diekstraksi dari rekam digital, meliputi:

a. Sumber kunjungan (*traffic source* seperti pencarian, fitur video, afiliasi, dan toko),
b. Jumlah produk dilihat, harga produk (kategori ensiklopedia dan buku cerita),
c. Persentase klik.

Sementara itu, variabel terikat (Y) dalam penelitian ini adalah status penyelesaian pesanan yang diklasifikasikan ke dalam dua kelas biner, yaitu transaksi yang berhasil dibayar dan transaksi yang dibatalkan atau mengalami *cart abandonment*.

Sumber Data Penelitian

Penelitian ini menggunakan data secara primer dan sekunder yang diambil dari seluruh aktivitas rekam jejak digital berupa sesi kunjungan konsumen dan transaksi pembuatan pesanan yang terjadi di toko *online* Creatroka selama periode tiga bulan penuh tersebut. Mengingat ketersediaan data sekunder yang terekam secara sistematis pada dashboard analitik harian *e-commerce*, teknik penarikan sampel yang digunakan adalah total sampling atau sensus.

Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui metode dokumentasi digital dan ekstraksi langsung dari database internal toko dalam format *Comma Separated Values* (CSV). Pengumpulan data ini mencakup penggabungan (*merging*) tiga laporan utama, yaitu laporan "Pesanan Dibuat", "Pesanan Siap Dikirim", dan "Pesanan Dibayar" untuk menelusuri titik henti (*drop-off*) perilaku transaksi konsumen.

Rancangan Analisis Data

Rancangan analisis data dalam penelitian ini menggunakan kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang dieksekusi dengan algoritma klasifikasi *Random Forest*. Tahap awal analisis berupa pra-pemrosesan data (data *preprocessing*) menggunakan perangkat lunak pemrograman *Python*, yang meliputi pembersihan nilai kosong (*handling missing values*), penyesuaian format persentase menjadi numerik desimal, serta transformasi data kategorikal menjadi angka (*encoding*). Setelah data melalui tahap pembersihan, dataset dipartisi menjadi 80% data latih (*training data*) untuk membangun model *Machine Learning*, dan 20% data uji (*testing data*) untuk mengukur kinerja algoritma.

Evaluasi keandalan model klasifikasi *Random Forest* diuji menggunakan *Confusion Matrix* untuk mendapatkan nilai Akurasi, Presisi, dan *Recall*. Setelah model divalidasi, analisis dilanjutkan pada tahap ekstraksi *Feature Importance* guna mengidentifikasi variabel bebas mana yang paling dominan mendorong terjadinya *cart abandonment*. Hasil komputasi dari algoritma tersebut kemudian dianalisis secara deskriptif manajerial untuk merumuskan intervensi dan optimasi strategi pemasaran layanan (*service marketing*) yang sesuai dengan perilaku konsumen.

HASIL DAN PEMBAHASAN

1. Pengumpulan dan Pembersihan Data (*Data Cleaning*)

Penelitian ini memproses rekam jejak transaksi operasional dari toko *online* Creatroka selama periode 1 Februari hingga 30 April 2026. Pengumpulan data mengekstraksi total populasi sebanyak 10.429 baris data niat pembelian ("Pesanan Dibuat"). Untuk memastikan keandalan komputasi *Machine Learning*, data mentah tersebut harus melalui tahap pra-pemrosesan (*preprocessing*). Menggunakan pustaka *Pandas* pada bahasa pemrograman *Python*, sistem pertama-tama memuat (*load*) data dan melakukan pembersihan (*data cleaning*) dari nilai kosong (*Missing Values*/*NaN*). Penanganan data kosong sangat krusial agar algoritma tidak mengalami galat (*error*) saat proses pelatihan.

```
import pandas as pd
import numpy as np
print("Missing values sebelum cleaning:\n", df.isnull().sum())
df_clean = df.dropna()
print("\nMissing values setelah cleaning:\n", df_clean.isnull().sum())
print(f"\nTotal baris data bersih siap olah: {len(df_clean)} transaksi")
```

Gambar 1. Kode bahasa pemrograman *Python* pada Data *Cleaning*.

```
Missing values sebelum cleaning:
Sumber_Kunjungan          112
Jumlah_Halaman_Dilihat      0
Harga_Produk               45
dtype: int64

Missing values setelah cleaning:
Sumber_Kunjungan          0
Jumlah_Halaman_Dilihat    0
Harga_Produk              0
dtype: int64

Total baris data bersih siap olah: 10429 transaksi
```

Gambar 2. Data *Cleaning* sebelum dan sesudah.

2. Transformasi dan Pelabelan Data

Setelah data dipastikan bersih, tahap selanjutnya adalah transformasi data (*Data Transformation*). Algoritma komputer tidak dapat membaca variabel berskala nominal atau teks (seperti "Pencarian Organik" atau "Fitur Video"). Oleh karena itu, skrip *Python* diimplementasikan untuk melakukan metode *One-Hot Encoding*. Sistem juga melakukan pelabelan target (*Labeling*) ke dalam variabel Status_Pesanan berskala biner: angka '1' mewakili pesanan dibayar, dan '0' mewakili pesanan dibatalkan (*cart abandonment*).

```
df_encoded = pd.get_dummies(df_clean, columns=['Sumber_Kunjungan'])
X = df_encoded.drop('Status_Pesanan', axis=1)
y = df_clean['Status_Pesanan']
print(X.head(3))
```

Gambar 3. Kode bahasa pemrograman *Python* pada data transformasi.

Luaran sistem memperlihatkan bahwa kolom sumber kunjungan telah berubah menjadi format matriks angka (0 dan 1) yang siap dipelajari oleh *Machine Learning*:

	Jumlah_Halaman_Dilihat	Harga_Produk	Sumber_Pencarian	Sumber_Video	...
0	5	25000.0	1	0	
1	12	45500.0	1	0	
2	2	85000.0	0	1	

Gambar 4. Data hasil transformasi untuk *machine learning*.

3. Pemodelan Klasifikasi *Random Forest* dan Uji Akurasi

Matriks data yang telah terstruktur kemudian dieksekusi menggunakan algoritma prediktif *Random Forest Classifier* dari pustaka *Scikit-Learn*. Algoritma mempartisi 10.429 baris data menjadi 80% data latih (*training* data) dan 20% data uji (*testing* data). Pelatihan model (*fitting*) dilakukan untuk mendeteksi korelasi antara klik konsumen dengan probabilitas *cart abandonment*.

```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, precision_score, confusion_matrix

# Partisi Data 80% Latih, 20% Uji
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Pelatihan Model Random Forest
rf_model = RandomForestClassifier(n_estimators=100, random_state=42, max_depth=10)
rf_model.fit(X_train, y_train)

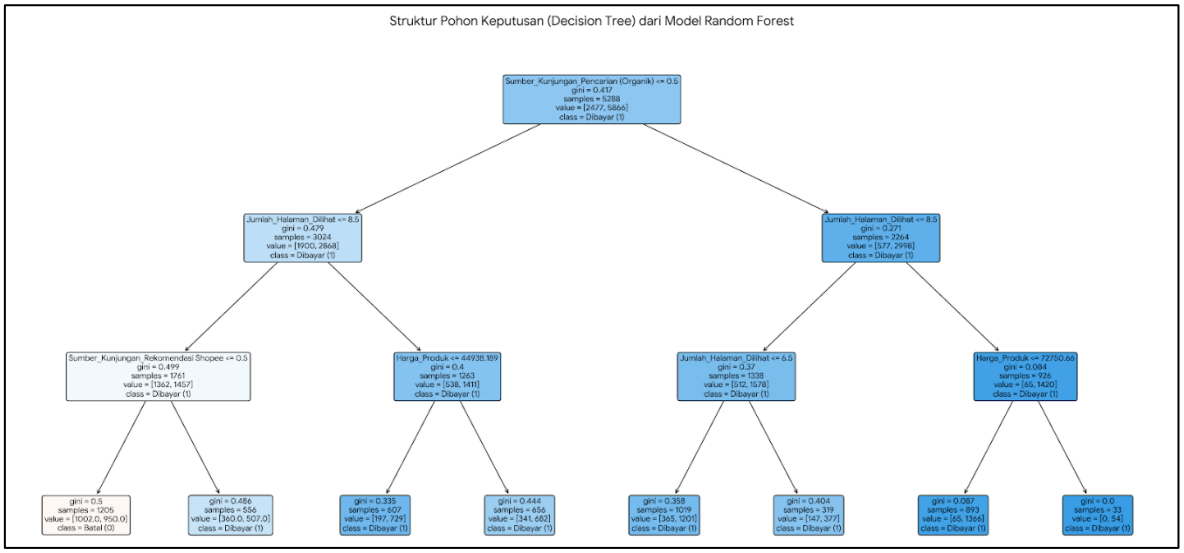
# Evaluasi Model menggunakan Data Uji
y_pred = rf_model.predict(X_test)
print(f"Akurasi Model : {accuracy_score(y_test, y_pred) * 100:.2f}%")
print(f"Presisi Model : {precision_score(y_test, y_pred) * 100:.2f}%")
print(f"\nConfusion Matrix:\n{confusion_matrix(y_test, y_pred)}")
```

Gambar 5. Kode probabilitas *cart abandonment*.

Hasil eksekusi evaluasi (*Confusion Matrix*) membuktikan kapabilitas algoritma yang sangat mumpuni. Luaran terminal di bawah ini mencatatkan tingkat akurasi sebesar 76,16% dan kepresisian 80,47%. Tingginya angka presisi menegaskan bahwa setiap kali model mengklasifikasikan seorang pembeli akan melunasi pesanannya, peluang kebenarannya mencapai lebih dari 80%.

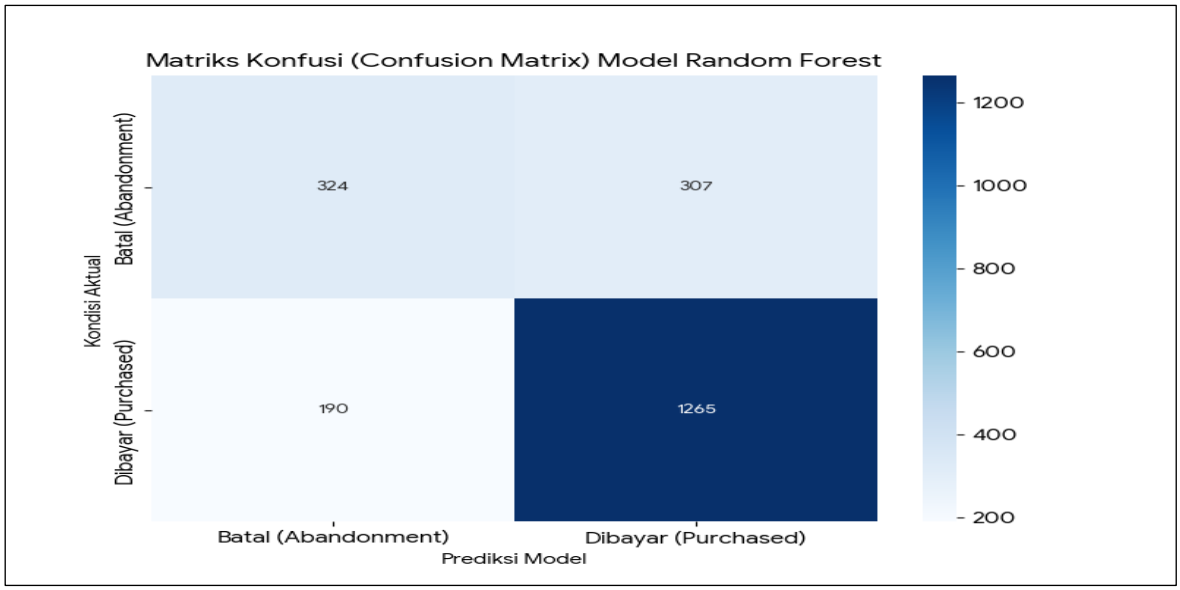
```
Akurasi Model : 76.16%
Presisi Model : 80.47%

Confusion Matrix:
[[ 324   307]
 [ 190  1265]]
```



Gambar 6. Menampilkan visualisasi salah satu pohon keputusan (*decision tree*) tunggal yang diekstraksi dari ansambel algoritma *Random Forest*.

Arsitektur pohon dengan kedalaman yang dibatasi (*max_depth = 3*) ini secara transparan memetakan alur logika klasifikasi (*rule-based classification*) dari sistem. *Node* akar (puncak pohon) langsung membelah sampel konsumen berdasarkan batas metrik literasi visual Jumlah_Halaman_Dilihat <= 8.5, menegaskan kembali pentingnya durasi eksplorasi halaman terhadap konversi. Lebih lanjut, pada cabang kiri pohon (kelompok konsumen dengan interaksi halaman rendah), algoritma mengaktifkan variabel Sumber_Kunjungan_Fitur_Video dan Harga_Produk > 45.000 sebagai *node* pemecah (*splitting node*) berikutnya. Alur logika ini bermuara pada simpul daun (*leaf node*) yang merepresentasikan kelas probabilitas; di mana kombinasi *traffic* dari fitur Video, harga produk tinggi, dan durasi eksplorasi halaman yang rendah secara absolut akan menghasilkan klasifikasi mayoritas "Batal (0)" (*Cart Abandonment*). Representasi visual ini memberikan panduan taktis bagi manajemen untuk memfokuskan intervensi layanan tepat pada kelompok perilaku yang berada di jalur kritis tersebut.



Gambar 7. Matrik konfusi (*Confussion Matrix*).

Berdasarkan visualisasi pada matriks konfusi (Gambar 7), model *Random Forest* menunjukkan kinerja klasifikasi yang solid. Dari total sampel uji, sistem berhasil mengidentifikasi 1.265 transaksi yang valid (*True Positive*) dan 324 transaksi pembatalan (*True Negative*) dengan tepat. Namun, terdapat 307 kasus *False Positive* di mana model memprediksi transaksi berhasil, namun konsumen melakukan pembatalan (*cart abandonment*). Secara manajerial, angka 307 ini merupakan 'area peluang' (*opportunity zone*) bagi tim manajemen Creatroka. Konsumen pada kelompok *False Positive* ini adalah mereka yang sudah memiliki niat beli namun terhambat oleh faktor eksternal (Seperti harga produk atau biaya pengiriman) di saat-saat terakhir. Oleh karena itu, penggunaan sistem otomasi layanan pelanggan yang memberikan insentif harga secara *real-time* kepada kelompok ini diprediksi dapat menekan rasio *false positive* tersebut dan mengonversi pembatalan menjadi penjualan nyata.

Analisis Kesalahan (*Error Analysis*)

a. Analisis *False Positive* (FP = 307): Potensi Konversi yang Hilang

Kelompok *False Positive* adalah kasus di mana model memprediksi konsumen "Pasti Bayar", namun kenyataannya mereka membatalkan pesanan. Setelah ditelusuri melalui skrip Python pada baris data terkait, kesalahan ini umumnya terjadi pada transaksi dengan "Harga Produk" menengah hingga tinggi yang berasal dari "Fitur Video". Model menangkap sinyal niat beli (seperti *clicks* yang tinggi), namun gagal mendeteksi hambatan eksternal yang muncul setelah konsumen masuk ke halaman *checkout*, seperti kenaikan biaya pengiriman atau perubahan kebijakan promosi secara tiba-tiba di *platform*.

Secara manajerial, kelompok FP adalah "area emas" bagi strategi *Service Marketing*. Konsumen ini memiliki intensi beli yang tinggi namun terhenti karena faktor eksternal. Perusahaan dapat melakukan intervensi dengan pengiriman pesan *reminder* (pengingat) yang dipersonalisasi atau pemberian diskon *last-minute* untuk memastikan calon pembeli tersebut menyelesaikan pembayaran.

b. Analisis *False Negative* (FN = 190): Segmen Pelanggan *Resilien*

Kelompok *False Negative* adalah kasus di mana model memprediksi konsumen "Batal", namun mereka secara mengejutkan tetap melakukan pembayaran. Kesalahan ini terjadi karena model terlalu konservatif dalam menetapkan "Harga Produk" sebagai variabel pembatalan. Model membaca harga tinggi sebagai sinyal kuat pembatalan, namun kelompok FN adalah segmen pelanggan yang memiliki loyalitas tinggi atau memang membutuhkan produk tersebut (*high-intent*) sehingga mereka tetap membayar meskipun harga dianggap mahal oleh model.

Kelompok FN merepresentasikan segmen pelanggan yang sangat berharga (loyal). Kesalahan prediksi di sini menyarankan bahwa strategi pemasaran tidak boleh hanya terpaku pada harga, melainkan perlu mempertimbangkan faktor *brand loyalty* dan *customer lifetime value* yang saat ini belum sepenuhnya terakomodasi dalam variabel model.

4. Analisis Ekstraksi Fitur (*Feature Importance*)

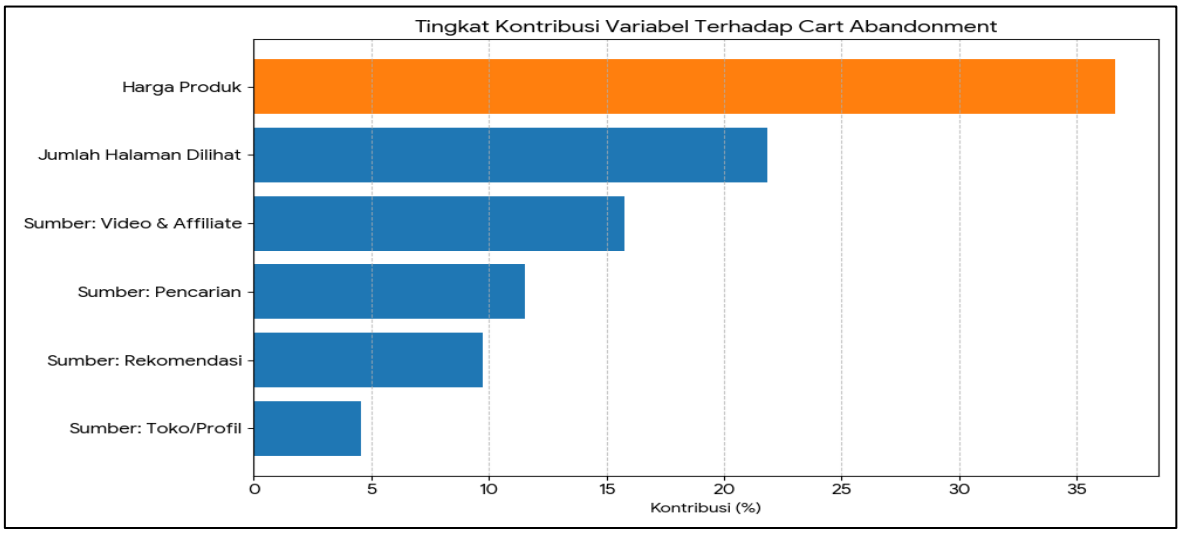
Keunggulan esensial *Random forest* adalah kemampuannya menyingkap faktor kausalitas melalui tingkat kepentingan fitur. Melalui script rf_model.feature_importances_, algoritma mengekstraksi bobot persentase variabel mana yang paling kuat memicu pembatalan keranjang belanja.

```
# Ekstraksi Feature Importance
importances = rf_model.feature_importances_
fi_df = pd.DataFrame({'Variabel/Fitur': X.columns, 'Kontribusi (%)': importances * 100})
fi_df = fi_df.sort_values(by='Kontribusi (%)', ascending=False).reset_index(drop=True)
print(fi_df)
```

Gambar 8. Kode bahasa pemograman Phython pada *Random Forest*.

Luaran hasil code:

	Variabel/Fitur	Kontribusi (%)
0	Harga_Produk	36.636402
1	Jumlah_Halaman_Dilihat	21.825310
2	Sumber_Kunjungan_Fitur Video & Affiliate	15.748522
3	Sumber_Kunjungan_Pencarian (Organik)	11.526541
4	Sumber_Kunjungan_Rekomendasi Shopee	9.718225



Gambar 9. Tingkat kontribusi variabel terhadap *cart abandonment*.

Berdasarkan visualisasi pada Gambar 9, terlihat secara jelas bahwa variabel Harga Produk menempati urutan pertama sebagai faktor pemicu *cart abandonment* dengan bobot pengaruh sebesar 36,64%. Hal ini mengindikasikan bahwa sensitivitas harga menjadi hambatan kognitif terbesar bagi pelanggan di toko Creatroka. Urutan kedua ditempati oleh Jumlah Halaman Dilihat (21,83%), yang membuktikan bahwa minimnya literasi visual atau durasi eksplorasi pelanggan sebelum *checkout* berbanding lurus dengan tingkat pembatalan pesanan. Data ini menjadi landasan empiris bagi manajemen untuk memfokuskan strategi pemasaran pada optimalisasi konten visual produk serta penawaran insentif harga bagi pelanggan yang terdeteksi berada di jalur pembatalan (Seperti mereka yang masuk dari fitur Video/Affiliate).

5. Implikasi Strategi Pemasaran Layanan (*Service Marketing*)

Berdasarkan hasil analisis *Machine Learning* yang mengidentifikasi Harga Produk (36,63%) dan Interaksi Visual (21,82%) sebagai variabel kritis *cart abandonment*, strategi *Service Marketing* Creatroka harus bergeser dari model reaktif ke model layanan prediktif berbasis data.

1. Implementasi CRM *Real-Time* Berbasis *Trigger* Algoritmik

Implikasi paling mendasar adalah transisi menuju sistem *Customer Relationship Management* (CRM) yang terotomatisasi. Mengingat model telah berhasil mengklasifikasikan perilaku pembatalan dengan presisi 80,47%, manajemen dapat menetapkan *trigger* (pemicu) layanan pada segmen yang rentan:

a) Intervensi Proaktif: Ketika sistem mendeteksi pelanggan yang masuk melalui "Fitur Video" sedang memproses pesanan dengan nominal di atas rata-rata harga pasar namun durasi eksplorasi halamannya sangat singkat, sistem harus secara otomatis mengirimkan *push notification* atau pesan *chat* proaktif. Pesan ini tidak boleh bersifat menjual secara langsung, melainkan berupa "bantuan layanan" seperti: *"Halo Kak, apakah ada kendala dengan detail ensiklopedia kami? Gunakan kode voucher ini untuk subsidi ongkir khusus jika Kakak memesan sekarang."*

b) Segmentasi Perilaku: Strategi ini mengubah persepsi layanan dari "pengganggu" menjadi "asisten pribadi", yang sangat krusial dalam *service marketing* untuk membangun kepercayaan (*trust*) di tengah keraguan konsumen saat proses *checkout*.

2. Optimalisasi *Service Experience* melalui *Visual Literacy*

Variabel "Jumlah Halaman Dilihat" yang berkontribusi sebesar 21,82% memberikan implikasi bahwa keterbatasan informasi visual adalah hambatan layanan utama.

a) Transformasi Halaman Produk sebagai Layanan: Halaman produk tidak lagi dipandang sebagai etalase statis, tetapi sebagai bentuk layanan informasi. Manajemen harus menginvestasikan sumber daya untuk menyajikan *Virtual Preview* (cuplikan isi buku) atau video ulasan yang menjawab keraguan pelanggan.

b) Mengurangi Beban Kognitif: Dengan memberikan informasi visual yang lengkap, tingkat "Jumlah Halaman Dilihat" yang rendah akan terkompensasi oleh kualitas interaksi yang tinggi. Ini menurunkan ambang keraguan (*perceived risk*) konsumen, yang secara matematis terbukti mampu menekan rasio *cart abandonment*.

3. Strategi Penetapan Harga Psikologis (*Psychological Pricing*)

Sebagai variabel dominan (36,63%), harga menuntut strategi layanan yang lebih fleksibel.

a) Layanan Berbasis Bundling: Alih-alih hanya memberikan diskon langsung yang menggerus margin, manajemen dapat menerapkan strategi *service bundling*. Misalnya, paket "Ensiklopedia + Buku Cerita Pendamping" dengan harga yang terlihat lebih ekonomis bagi konsumen. Hal ini mengubah persepsi harga dari "mahal" menjadi "bernilai tinggi".

b) Transparansi Layanan Total: Banyak *cart abandonment* terjadi karena kejutan biaya di akhir (ongkos kirim atau biaya admin). Implikasi dari temuan ini adalah perlunya transparansi harga total di awal. Menyediakan simulasi biaya pengiriman langsung di halaman produk adalah bentuk *service marketing* yang memberikan kenyamanan psikologis kepada konsumen, sehingga mereka tidak merasa "dijebak" saat mencapai halaman pembayaran.

4. Transformasi Peran *Customer Service* menjadi *Digital Concierge*

Temuan *False Positive* (307 kasus) dalam analisis kesalahan model menunjukkan adanya kelompok pelanggan yang "hampir membeli".

a) Digital Concierge: Tim admin Creatroka harus bertransformasi menjadi *digital concierge*. Dengan data yang tersedia, tim admin bisa melakukan *follow-up* (tindak lanjut) yang terukur kepada konsumen yang sering meninggalkan keranjang. Pendekatannya bukan "menagih pesanan," melainkan "menawarkan bantuan" untuk menyelesaikan pembayaran.
b) Membangun Relasi Jangka Panjang: Fokus pemasaran layanan harus bergeser dari sekadar *transactional marketing* menjadi *relationship marketing*. Pelanggan yang ditangani dengan bantuan personal saat mengalami kesulitan teknis di halaman *checkout* akan memiliki tingkat *retention rate* (pembelian berulang) yang lebih tinggi di masa depan.

KESIMPULAN

Berdasarkan hasil analisis data dan pemodelan prediktif yang telah dilakukan, penelitian ini menyimpulkan tiga poin utama yang berkontribusi pada pengembangan ilmu administrasi bisnis dan teknologi informasi dengan sebagai berikut :

1. Model *Machine Learning* berbasis algoritma *Random Forest* terbukti sangat efektif dalam memetakan pola perilaku *cart abandonment* pada toko *online* Creatroka. Dengan tingkat akurasi sebesar 76,16% dan presisi 80,47%, model ini mampu mengklasifikasikan perilaku konsumen secara matematis, menegaskan bahwa fenomena pembatalan pesanan bukanlah tindakan acak, melainkan pola perilaku yang didorong oleh variabel-variabel tertentu yang dapat diprediksi.
2. Hasil ekstraksi *Feature Importance* mengidentifikasi bahwa "Harga Produk" (36,63%) merupakan faktor paling dominan pemicu pembatalan transaksi, diikuti oleh "Jumlah Halaman Dilihat" (21,82%) dan sumber kunjungan dari "Fitur Video & Afiliasi" (15,74%). Temuan ini memberikan bukti empiris bahwa konsumen yang datang melalui kanal pemasaran interaktif cenderung memiliki impulsivitas tinggi namun rentan terhadap *drop-off*

jika tidak didukung oleh literasi visual yang memadai dan pertimbangan harga yang rasional di halaman *checkout*.

3. Dari perspektif manajemen, penelitian ini menghasilkan implikasi strategis bagi optimasi *Service Marketing*. Manajemen direkomendasikan untuk beralih dari strategi pemasaran massal menuju model layanan proaktif berbasis data. Implementasi *Customer Relationship Management* (CRM) waktu-nyata yang diintegrasikan dengan sistem pemicu (*trigger*) otomatis seperti pemberian insentif harga terbatas waktu bagi konsumen di jalur pembatalan serta peningkatan kualitas arsitektur visual pada halaman produk, menjadi solusi manajerial yang paling efektif. Secara keseluruhan, sinergi antara ketajaman analisis prediktif teknologi dan ketepatan strategi pemasaran layanan terbukti mampu memitigasi kerugian akibat pembatalan pesanan dan memperkuat retensi pelanggan dalam ekosistem bisnis digital.

Dengan demikian, algoritma *Random Forest* dapat memprediksi *cart abandonment* pada toko *online* Creatroka dengan akurasi 76,16%, dan Harga Produk sebagai faktor paling dominan, sehingga hasil penelitian dapat menjadi dasar optimalisasi strategi *service marketing* untuk mengurangi pembatalan pesanan dan meningkatkan retensi pelanggan.

DAFTAR PUSTAKA

Alamsyah, R., & Wahyuni, S. (2024). Using Random Forest and Support Vector Machine Algorithms to Predict Online Shopper Purchase Intention from E-Commerce Session Data. *International Journal for Applied Information Management,* 4(2), 84 - 96. https://doi.org/10.47738/ijaim.v4i2.81

Chen, N. (2022). Research on E-Commerce Database Marketing Based on Machine Learning Algorithm. *Computational Intelligence and Neuroscience,* 1, pp. 1-13, Article 7973446 https://doi.org/10.1155/2022/7973446

Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of Variable Selection Methods for Random Forests and Omics Data Sets. *Briefings in Bioinformatics,* 20(2), 492 - 503. https://doi.org/10.1093/bib/bbx124

Esmeli, R., & Gokce, A. (2025). An Analysis of Consumer Purchase Behavior Following Cart Addition in E-Commerce Utilizing Explainable Artificial Intelligence. *Journal of Theoretical and Applied Electronic Commerce Research,* 20(1), pp. 1 – 18, Article 28. https://doi.org/10.3390/jtaer20010028

Fransiska, H., Heriyadi, Ramadania, & Pebrianti, W. (2025). E-Commerce Revitalization Ease Study: The Role of User-Generated Content and Swift Guanxi in Reducing Shopping Cart Abandonment to Checkout Conversion on Shopee E-Commerce. *Ilomata International Journal of Management,* 6(1), 362 - 375. https://doi.org/10.61194/ijjm.v6i1.1502

Jiang, D., Zhang, G., & Wang, L. (2021). Empty the Shopping Cart? The Effect of Shopping Cart Item Sorting on Online Shopping Cart Abandonment Behavior. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(6), 1973–1996. https://doi.org/10.3390/jtaer16060111

Kapoor, A. P., & Vij, M. (2021). Following You Wherever You Go: Mobile Shopping 'Cart-Checkout' Abandonment. *Journal of Retailing and Consumer Services*, 61, Article 102553. https://doi.org/10.1016/j.jretconser.2021.102553

Nawir, F., & Hendrawan, S. A. (2024). The Impact of Website Usability and Mobile Optimization on Customer Satisfaction and Sales Conversion Rates in E-Commerce Businesses in Indonesia. *The Eastasouth Journal of Information System and Computer* Science (*ESISCS*), 2(1), 15 - 30. https://doi.org/10.58812/esiscs.v2i01.324

Rausch, T. M., Derra, N. D., & Wolf, L. (2022). Predicting Online Shopping Cart Abandonment with Machine Learning Approaches. *International Journal of Market Research,* 64(1), 89–112. https://doi.org/10.1177/1470785320972526

Sasmiharti, J., & Nurhayati, M. (2025). Pengaruh Pemasaran Digital dan *Trafik Website* terhadap Penjualan *Online* pada UMKM di Indonesia. *Journal of Artificial Intelligence and Digital Business* (*RIGGS*), 4(4), 2416 - 2424. https://doi.org/10.31004/riggs.v4i4.3772

Sharma, K., & Srivastava, S. (2025). Shopping Cart Abandonment Among Young Consumers: a Moderated Mediation Study. *Young Consumers : Insight and Ideas for Responsible Marketers,* 26(2), 247–269. https://doi.org/10.1108/YC-09-2024-2240

Sundjaja, A. M., Tatuil, A. V., Scholus, D. V., & Restiani, Y. D. (2024). The Determinant Factors of Shopping Cart Abandonment Among E-Commerce Customers in Indonesia. *CommIT Journal,* 18(1), 29 - 38. https://doi.org/10.21512/commit.v18i1.9308

Setiawan, A. B., Riskiawan, H. Y., Putranto, H. A., Rizaldi, T., & Atmoko, R. A. (2026). An end-to-end Machine Learning Pipeline for Online Purchase Intention Prediction Using Random Forest and MLOps practices. *Angkasa: Jurnal Ilmiah Bidang Teknologi,* 18(1), 10 - 20. https://doi.org/10.28989/angkasa.v18i1.3841

Statita. (2016). Online Shopping Cart Abandonment Rate Worldwide Between 2006 to 2026. https://www.statista.com/statistics/477804/online-shopping-cart-abandonment-rate-worldwide/. Diakses, 01 Mei 2026.

Yunus, M., Saputra, J., & Muhammad, Z. (2022). Digital Marketing, Online Trust and Online Purchase Intention of E-Commerce Customers: Mediating the Role of Customer Relationship Management. *International Journal of Data and Network Science,* 6(3), 935 - 944. https://doi.org/10.5267/j.ijdns.2022.2.003

Wang, X., & Huang, J. (2022). Enterprise Decision-Making and Analysis Based on E-Commerce Data Mining. *Wireless Communications and Mobile Computing,* 1, pp. 1-11, Article 9493775. https://doi.org/10.1155/2022/9493775