Jurnal SENOPATI

Sustainability, Ergonomics, Optimization, and Application of Industrial Engineering
Jurnal homepage : ejurnal.itats.ac.id/senopati

Prediksi Preferensi Waktu Pengiriman Pelanggan untuk Optimalisasi Layanan Logistik Menggunakan *Decision Tree*

Femi Yulianti[1*], Aulia Dihas Zahira[2], Hardiyat Hawari [3], Faris Nur Fadhil [4], dan Marisa Dwi S [5]
[1]Jurusan Teknik Logistik, Universitas Telkom Bandung, Jl. Telekomunikasi No. 1, Bandung, Indonesia

INFORMASI ARTIKEL

Halaman:
171 – 176
Tanggal penyerahan:
3 Agustus 2025

Tanggal diterima:
30 April 2026

Tanggal terbit:
30 April 2026

ABSTRACT

This study aims to predict customer delivery time preferences to support logistics service optimization in last-mile delivery operations. Due to the absence of explicit preference data in logistics systems, a data-driven approach is proposed to infer customer preferences from historical delivery data. The dataset is transformed into four categories: morning, afternoon, evening, and night. Three machine learning models – Naive Bayes, Logistic Regression, and Decision Tree – are implemented and evaluated. To address class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) is applied. Model performance is assessed using accuracy, precision, recall, and F1-score metrics.The result show that the Decision Tree model achieves the best performance, with an accuracy of 77% and a macro F1-score of 0.77. Further analysis reveals that traffic conditions, vehicle type, and product category are the most influential factors in determining delivery time preferences. However, the model exhibits limitations in identifying minority classes, particularly morning deliveries, due to data imbalance.This study contributes by providing a data-driven framework for predicting delivery time preferences and integrating customer behavior insights into logistics decision-making. The findings offer practical implications for developing adaptive delivery scheduling, improving service reliability, and enabling personalized delivery time recommendations in e-commerce systems.

Keywords: *predictive analytics, delivery time preference, decision tree, last-mile logistics, customer satisfaction, machine learning.*

EMAIL

[1]femiyulianti@telkomuniversity.ac.id
[2]auliadihasz@gmail.com
[3]hardiyathawari@student.telkomuniversity.ac.id
[4]farisnurfadhil@student.telkomuniversity.ac.id
[5]marisadwis@student.telkomuniversity.ac.id

ABSTRAK

Penelitian ini bertujuan untuk memprediksi preferensi waktu pengiriman pelanggan guna mendukung optimalisasi layanan logistik pada proses last-mile delivery. Keterbatasan sistem logistik yang tidak menyimpan data preferensi pelanggan secara eksplisit mendorong penggunaan pendekatan berbasis data untuk mengekstrak pola preferensi dari data historis pengiriman. Data diklasifikasikan ke dalam empat kategori waktu, yaitu pagi, siang, sore, dan malam. Tiga algoritma *machine learning* digunakan dalam penelitian ini, yaitu *Naive Bayes, Logistic Regression*, dan *Decision Tree*. Untuk mengatasi ketidakseimbangan data, diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE). Evaluasi model dilakukan menggunakan metrik *accuracy, precision, recall,* dan F1-score. Hasil penelitian ini menunjukkan bahwa model *Decision Tree* memberikan performa terbaik dengan akurasi sebesar 77% dan nilai F1-score makro sebesar 0.77. Analisis lebih lanjut menunjukkan bahwa kondisi lalu lintas, jenis kendaraan, dan kategori produk merupakan faktor utama yang memengaruhi preferensi waktu pengiriman. Namun, model masih memiliki keterbatasan dalam mengidentifikasi kelas minoritas, khususnya waktu pagi.

Penelitian ini memberikan kontribusi berupa kerangka kerja prediktif berbasis data untuk mengidentifikasi preferensi waktu pengiriman serta mengintegrasikan wawasan perilaku pelanggan ke dalam pengambilan keputusan logistik. Hasil penelitian ini dapat dimanfaatkan untuk pengembangan sistem penjadwalan pengiriman yang adaptif, peningkatan keandalan layanan, serta rekomendasi waktu pengiriman yang lebih personal dalam sistem *e-commerce.*

Kata kunci: analitik prediktif, preferensi pengiriman, *decision tree, logistic last-mile,* kepuasan pelanggan, *machine learning.*

PENDAHULUAN

Perkembangan *e-commerce* yang pesat telah meningkatkan kompleksitas sistem logistik, khususnya pada tahap *last-mile delivery* yang menjadi faktor utama dalam menentukan kepuasan pelanggan. Studi menunjukkan bahwa lebih dari 50% total biaya logistik berasal dari aktivitas *last-mile delivery*, sehingga efisiensi pada tahap ini menjadi sangat krusial [1]. Selain itu, pelanggan tidak hanya menuntut kecepatan pengiriman, tetapi juga fleksibilitas waktu pengiriman barang, yang berdampak langsung pada persepsi layanan dan loyalitas pelanggan [2].

Namun, dalam praktiknya masih sering terjadi ketidaksesuaian antara waktu pengiriman aktual dengan preferensi pelanggan, yang berdampak langsung pada kinerja operasional logistik. Ketidaksesuaian ini dapat menyebabkan kegagalan pengiriman, peningkatan biaya operasional akibat proses pengiriman ulang, serta penurunan kepuasan pelanggan [3]. Dampak tersebut dirangkum pada Tabel 1.

Tabel 1. Dampak Ketidaksesuaian Waktu Pengiriman terhadap Kinerja Logistik

Faktor	Dampak
Kegagalan pengiriman	Peningkatan biaya operasional
Pengiriman ulang	Penambahan waktu distribusi
Ketidakpuasan pelanggan	Penurunan kualitas layanan

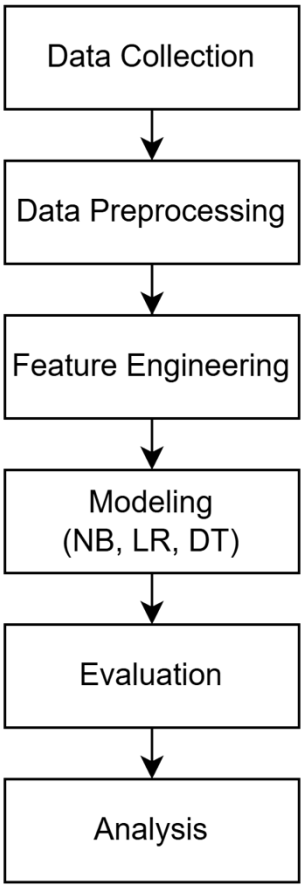
Pendekatan berbasis data (*data-driven*) menjadi solusi untuk mengidentifikasi pola preferensi pelanggan dari data historis. Metode *machine learning* seperti *Naive Bayes, Logistic Regression*, dan *Decision Tree* telah banyak digunakan dalam permasalahan klasifikasi karena kemampuannya dalam menangani data berskala besar dan variabel kategorikal [4] [5]. Diantarai metode tersebut, *Decision Tree* memiliki keunggulan dalam interpretabilitas dan kemampuan menangkap hubungan non-linear antar variabel [6].

Namun, penelitian sebelumnya lebih banyak berfokus pada prediksi performa logistik atau permintaan pengiriman, sementara kajian terkait prediksi preferensi waktu pengiriman pelanggan masih terbatas, khususnya pada kondisi tanpa data preferensi eksplisit [7]. Selain itu, integrasi antara hasil model prediktif dan persepsi pelanggan masih jarang dilakukan, sehingga relevansi praktis model dalam konteks operasional seperti kondisi lalu lintas, jenis kendaraan, dan kategori produk [8][9] [10].

Kontribusi penelitian ini meliputi (1) pengembangan pendekatan prediktif untuk mengidentifikasi preferensi waktu pengiriman tanpa data eksplisit, (2) integrasi hasil model dengan persepsi pelanggan untuk meningkatkan validitas praktis, dan (3) penyediaan dasar pengambilan keputusan dalam pengembangan sistem penjadwalan pengiriman yang adaptif dan berbasis preferensi pelanggan [11] [12] [13].

METODE

Penelitian ini menggunakan pendekatan machine learning untuk mengklasifikasikan preferensi waktu pengiriman pelanggan berdasarkan data historis operasional logistik. Metode disusun secara sistematis untuk memastikan hasil yang valid secara teknis dan relevan secara operasional[14].



Gambar 1. Alur Metodologi Penelitian Berbasis Machine Learning

a. *Data*

Data yang digunakan adalah Amazon Delivery Dataset yang mencerminkan aktivitas pengiriman berbasis *e-commerce*. Dataset mencakup variabel operasional seperti waktu pengiriman, kondisi lalu lintas, kondisi cuaca, jenis kendaraan, dan kategori produk.

b. *Data Preprocessing*

Tahap *preprocessing* meliputi transformasi variabel target dan penanganan ketidakseimbangan data. Transformasi variabel target digunakan untuk mengklasifikasikan waktu pengiriman menjadi empat kategori yaitu pagi, siang, sore, dan malam. Penanganan ketidakseimbangan data menggunakan metode SMOTE untuk menyeimbangkan distribusi kelas [15].

c. *Feature Engineering*

Dilakukan rekayasa fitur untuk meningkatkan representasi data, meliputi (1) *Weather_Traffic_Ratio* dan (2) *Total_Condition_Score.* Fitur yang dikembangkan bertujuan menangkap interaksi antar variabel operasional yang tidak dapat direpresentasikan secara langsung.

d. *Modeling*

Tiga algoritma klasifikasi yang digunakan yaitu *Naive Bayes, Logistic Regression*, dan *Decision Tree*. Model *Decision Tree* di optimasi menggunakan GridSearchCV. Penggunaan tiga algoritma bertujuan untuk membandingkan performa antara model probabilistik, linear, dan non-linear. Model Decision Tree dipilih sebagai kandidat utama karena kemampuannya dalam menangkap pola kompleks.

e. *Evaluation*

Evaluasi model menggunakan metrik yang beragam karena dapat mengevaluasi model secara komprehensif, terutama dalam kondisi data yang tidak seimbang. *Confusion matrix* digunakan untuk mengidentifikasi pola kesalahan klasifikasi yang dapat memberikan *insight* terhadap keterbatasan model.

f. *Pipeline Machine Learning*

Pipeline ini memastikan bahwa proses penelitian berjalan secara sistematis dan terstruktur, sehingga setiap tahapan dapat dievaluasi secara terpisah.

HASIL DAN PEMBAHASAN

a. Deskripsi *Dataset*

Dataset yang digunakan terdiri dari 8.730 data pengiriman, dengan distribusi kelas yang tidak seimbang, yaitu dominasi pada kategori malam (56.8%), diikuti sore (20.1%), siang (12.7%), dan pagi (10.4%). Variabel yang digunakan meliputi waktu pengiriman, kondisi lalu lintas, kondisi cuaca, jenis kendaraan, dan kategori produk.

b. Evaluasi Performa Model

Tabel 2. Evaluasi Performa Model

Model	Akurasi	*Precision*	*Recall*	F-1 *Score*
Naive Bayes	57%	0.59	0.55	0.54
Logistic Regression	68%	0.67	0.63	0.64
Decision Tree	77%	0.83	0.77	0.77

Sumber: Hasil eksperimen pada data uji (test set), klasifikasi preferensi waktu pengiriman

Decision Tree menghasilkan performa terbaik dengan akurasi 77%. Keunggulan ini menunjukkan bahwa hubungan antar variabel logistik bersifat non-linear, sehingga model linear seperti *Logistic Regression* tidak mampu menangkap pola secara optimal.

c. Analisis Distribusi Kelas

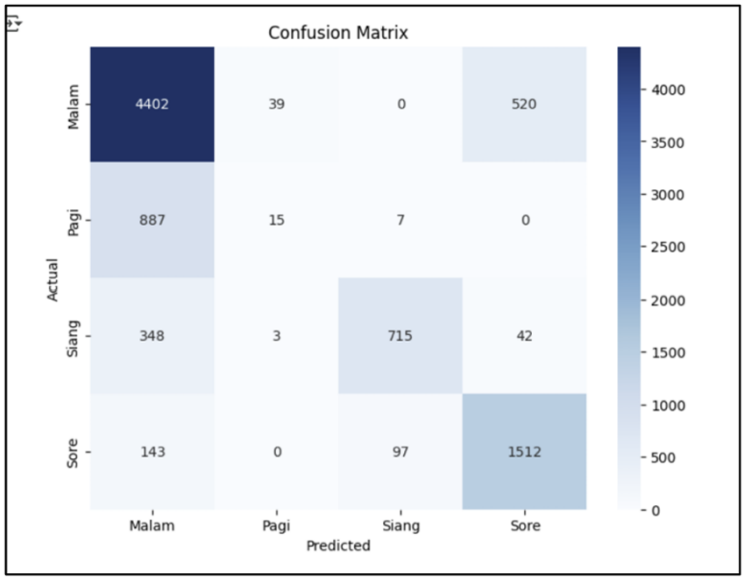
Tabel 3. Perbandingan Data Aktual dan Prediksi

Kategori Waktu	Jumlah Data Aktual	Jumlah Data Diprediksi	Benar Diklasifikasi	*Recall* per Kelas
Pagi	909	57	15	0.02
Siang	1.108	819	715	0.65
Sore	1.752	2.074	1.512	0.86
Malam	4.961	5.780	4.402	0.89

Sumber: Hasil eksperimen pada data uji (test set), klasifikasi preferensi waktu pengiriman

Model memiliki performa tinggi pada sore dan malam, namun sangat rendah pada pagi. Hal ini menunjukkan adanya bias terhadap kelas dominan, dimana pola pengiriman lebih banyak terpusat pada sore dan malam. Rendahnya *recall* pada pagi juga mengindikasikan keterbatasan fitur dalam membedakan karakteristik waktu tersebut.

d. Analisis Distribusi dan Kesalahan Klasifikasi



Gambar 2. *Confusion Matrix Decision Tree*

Model menunjukkan performa tinggi pada sore dan malam, namun sangat rendah pada pagi (*recall* 0.02). Hal ini menunjukkan adanya bias model terhadap kelas dominan. Analisis *confusion matrix* menunjukkan bahwa sebagian besar data pagi diklasifikasikan sebagai malam. Hal ini mengindikasikan bahwa model cenderung mengikuti pola operasional sistem logistik yang lebih aktif pada malam hari.

e. Implikasi Operasional dan Sistem *E-Commerce*

Hasil menunjukkan bahwa distribusi pengiriman masih terpusat pada sore dan malam, sehingga mengindikasikan ketidakseimbangan beban operasional. Model yang dikembangkan dapat digunakan untuk mengoptimalkan penjadwalan pengiriman berbasis preferensi, meningkatkan alokasi sumber daya secara dinamis, serta mengurangi risiko pengiriman ulang.

Secara operasional, model ini mendukung peralihan menuju *predictive scheduling*, yang berpotensi meningkatkan efisiensi melalui pengurangan biaya dan peningkatan utilisasi kapasitas. Dalam konteks *e-commerce*, model dapat diintegrasikan ke dalam sistem *checkout* untuk rekomendasi waktu pengiriman serta pengaturan *delivery slot* yang lebih adaptif.

KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi preferensi waktu pengiriman pelanggan menggunakan pendekatan *machine learning* berbasis data historis. Hasil menunjukkan bahwa model *Decision Tree* memberikan performa terbaik dengan akurasi sebesar 77% dan F1-score makro sebesar 0.77, yang mengindikasikan kemampuannya dalam menangkap pola non-linear pada data operasional logistik.

Analisis menunjukkan bahwa model ini memiliki performa tinggi pada kategori sore dan malam, namun terbatas pada kategori pagi akibat ketidakseimbangan data. Kondisi ini sejalan dengan hasil pembahasan yang menunjukkan dominasi pengiriman pada waktu tertentu, yang berpotensi menyebabkan ketidakefisienan distribusi dan peningkatan risiko pengiriman ulang.

Secara praktis, model ini dapat digunakan untuk mendukung penjadwalan pengiriman berbasis preferensi, meningkatkan pemanfaatan kapasitas distribusi, dan mengurangi ketidaksesuaian waktu pengiriman dalam sistem logistik.

Penelitian selanjutnya disarankan untuk meningkatkan performa pada kelas minoritas (pagi), memperbaiki keseimbangan dataset, serta mengembangkan model dengan variabel tambahan dan integrasi data *real-time* guna meningkatkan akurasi dan relevansi implementasi.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan kontribusi dalam pelaksanaan penelitian ini, khususnya dalam proses pengumpulan data, pengembangan model, dan analisis hasil. Apresiasi juga disampaikan atas kolaborasi yang telah mendukung terselesaikannya penelitian ini dengan baik.

DAFTAR PUSTAKA

[1] S. Chopra and P. Meindl, Supply Chain Management: Strategy, Planning, and Operation, 6th ed. Pearson, 2016.
[2] D. Boysen, S. Fedtke, and M. Schwerdfeger, “Last-mile delivery concepts: A survey,” Omega, vol. 100, 2021.
[3] S. Agatz, A. Campbell, M. Fleischmann, and M. Savelsbergh, “Time slot management in attended home delivery,” Transp. Sci., vol. 45, no. 3, pp. 435–449, 2011.
[4] C. Hübner, H. Kuhn, and A. Sternbeck, “Demand and supply chain planning in retail: A review,” Eur. J. Oper. Res., vol. 289, no. 1, pp. 1–16, 2021.
[5] D. Zhang, H. Zhao, Y. Wang, and F. Xu, “Application of machine learning in logistics and supply chain management,” IEEE Access, vol. 7, pp. 144535–144551, 2019.
[6] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Elsevier, 2022.

[7] J. D. Kelleher, Fundamentals of Machine Learning for Predictive Data Analytics. MIT Press, 2020.
[8] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, Classification and Regression Trees. Wadsworth, 2000.
[9] Y. Liu, H. He, and X. Liu, "Machine learning for supply chain management: A systematic review," Comput. Ind. Eng., vol. 137, 2019.
[10] A. Rai, A. Constantinides, and S. Sarker, "Next-generation digital platforms: Toward human–AI hybrids," MIS Q., vol. 43, no. 1, pp. iii–ix, 2019.
[11] M. Winkenbach, D. Janjevic, and J. Hübner, "Urban last-mile logistics: Challenges and opportunities," Transp. Res. Part E, vol. 129, pp. 1–16, 2019.
[12] X. Chen, S. Zhang, and Y. Wang, "Integrated inventory planning with price-dependent demand under promotions," Int. J. Prod. Econ., vol. 240, 2021.
[13] M. Ghanbari and R. Zanjirani Farahani, "Robust inventory and shelf space allocation under demand uncertainty," Comput. Ind. Eng., vol. 179, 2023.
[14] S. Zajac and H. Kuhn, "Joint shelf design and space planning problem with placement options," Prod. Oper. Manag., 2025.
[15] A. Ostermeier, T. Düsterhöft, and A. Hübner, "Shelf space allocation and replenishment planning in retail," Omega, vol. 102, 2021.