

Implementasi Algoritma Random Forest untuk Prediksi Permintaan dan Optimasi Stok pada Sistem Manajemen Inventori Layanan Pengiriman Makanan

Achmad Ichwani^{1*}, Egie Irawan², Lintang Wahyu Aji Saputro³, Gibrand Putra Pradana⁴, Joni Maulindar⁵

¹Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta

^{1*}230103256@mhs.udb.ac.id

²Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta

²230103132@mhs.udb.ac.id

³Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta

³230103140@mhs.udb.ac.id

⁴Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta

⁴230103136@mhs.udb.ac.id

⁵Fakultas Ilmu Komputer
Universitas Duta Bangsa Surakarta

⁵joni_maulindar@udb.ac.id

Abstrak— Industri pengiriman makanan menghadapi tantangan besar dalam pengelolaan inventori akibat sifat produk yang mudah rusak dan fluktuasi permintaan yang tinggi. Penelitian ini bertujuan mengembangkan sistem prediksi permintaan dan optimasi stok berbasis algoritma Random Forest, menggunakan metode CRISP-DM dengan data historis dari Kaggle dan implementasi sistem web berbasis Python Flask. Hasil penelitian menunjukkan Random Forest memiliki kinerja terbaik dibandingkan Decision Tree dan Linear Regression, dengan skor MSE 109.570,40, RMSE 331,01, dan R^2 sebesar 0,26. Skor R^2 ini secara spesifik mengindikasikan bahwa model hanya mampu menjelaskan 26% dari variabilitas data permintaan. Performa yang terbatas ini utamanya disebabkan oleh ketiadaan variabel eksternal yang krusial seperti data cuaca atau hari libur nasional dalam dataset yang digunakan, yang diketahui sangat mempengaruhi dinamika pasar. Meskipun demikian, model ini berhasil membuktikan keunggulan Random Forest dalam menangani hubungan non-linier dan dapat mendukung estimasi kebutuhan stok mingguan melalui sistem manajemen inventori yang fungsional. Visualisasi seperti grafik feature importance dan heatmap korelasi juga diimplementasikan untuk membantu pemahaman pola data serta pengambilan keputusan. Penelitian ini menegaskan potensi besar machine learning sebagai fondasi sistem prediksi di rantai pasok makanan, dengan rekomendasi utama untuk pengembangan selanjutnya adalah memperkaya dataset dengan variabel eksternal guna meningkatkan akurasi model secara signifikan.

Kata kunci— Random Forest, prediksi permintaan, manajemen inventori, pengiriman makanan, CRISP-DM.

Abstract— The food delivery industry faces significant challenges in inventory management due to the perishable nature of products and high demand fluctuations. This study aims to develop a demand forecasting and stock optimization system based on the Random Forest algorithm, utilizing the CRISP-DM methodology with historical data from Kaggle and a web-based system implementation using Python Flask. The research results show that Random Forest achieved the best performance compared to Decision Tree and Linear Regression, with an MSE of 109,570.40, RMSE of 331.01, and an R^2 of 0.26. This R^2 score specifically indicates that the model can only explain 26% of the demand data variability. This limited performance is primarily attributed to the absence of crucial external variables such as weather data or national holidays in the dataset used, which are known to significantly affect market dynamics. Nevertheless, this model successfully demonstrates the superiority of Random Forest in handling non-linear relationships and can support weekly stock estimation through a functional inventory management system. Visualizations such as feature importance charts and correlation heatmaps were also implemented to aid in understanding data patterns and support decision-making. This research affirms the significant potential of machine learning as a foundation for forecasting systems in the food supply chain, with the main recommendation for future development being to enrich the dataset with external variables to significantly improve model accuracy.

Keywords— Random Forest, demand forecasting, inventory management, food delivery, CRISP-DM.

I. PENDAHULUAN

Industri pengiriman makanan mengalami pertumbuhan yang sangat pesat seiring dengan perubahan perilaku konsumen, peningkatan

adopsi teknologi digital, serta meningkatnya permintaan terhadap layanan cepat dan fleksibel. Di tengah dinamika ini, pengelolaan inventori menjadi tantangan krusial, terutama karena produk makanan umumnya bersifat mudah rusak

(perishable) dan memiliki umur simpan yang pendek. Ketidaktepatan dalam merencanakan kebutuhan bahan baku dapat menyebabkan pemborosan akibat kelebihan stok atau kerugian layanan karena kekurangan stok. Kedua kondisi ini tidak hanya berdampak pada profitabilitas perusahaan, tetapi juga secara langsung mempengaruhi kepuasan pelanggan [1], [2]. Oleh karena itu, sistem manajemen inventori yang responsif dan presisi sangat dibutuhkan untuk mengantisipasi fluktuasi permintaan dan mendukung proses pengambilan keputusan yang berbasis data [3].

Untuk menjawab kebutuhan tersebut, berbagai pendekatan teknologi telah diterapkan, salah satunya adalah penggunaan algoritma machine learning untuk melakukan prediksi permintaan. Di antara berbagai algoritma yang tersedia, Random Forest (RF) muncul sebagai metode yang unggul dalam menangani kompleksitas data historis, hubungan non-linier antar variabel, serta mampu memberikan hasil prediksi yang akurat dan stabil [4], [5]. Penerapan RF dalam sistem prediksi permintaan telah terbukti efektif di berbagai sektor seperti ritel, farmasi, dan logistik, termasuk dalam konteks layanan pengiriman makanan [6]. Misalnya, penelitian oleh Panda dan Mohanty menunjukkan bahwa RF Regressor mampu memberikan hasil prediksi yang kompetitif dalam lingkungan yang penuh ketidakpastian dan fluktuasi permintaan [6]. Dalam studi lain, Alsanad menemukan bahwa penggunaan RF Classifier dalam logistik dapat meningkatkan akurasi prediksi permintaan harian secara signifikan, yang berdampak langsung pada efisiensi rantai pasok [4].

Namun, tantangan implementasi model prediksi tidak hanya berhenti pada akurasi. Penggabungan hasil prediksi ke dalam sistem manajemen inventori yang operasional dan mudah digunakan oleh pihak manajemen menjadi aspek penting yang tak kalah menentukan keberhasilan sistem secara keseluruhan. Dalam konteks ini, metodologi integrasi seperti CRISP-DM banyak digunakan untuk memandu alur pengembangan

sistem, mulai dari pemahaman kebutuhan bisnis hingga tahap deployment dan evaluasi [7].

Penelitian ini bertujuan untuk mengembangkan model prediksi permintaan berbasis algoritma Random Forest pada layanan pengiriman makanan dan mengintegrasikannya ke dalam sistem manajemen inventori untuk optimasi stok bahan baku. Model prediksi ini akan dibandingkan dengan algoritma lain seperti Decision Tree dan Linear Regression untuk menilai efektivitas relatifnya [8]. Dataset yang digunakan diperoleh dari platform Kaggle dan mencakup data permintaan dari berbagai pusat pemenuhan pesanan selama 145 minggu. Penelitian ini diharapkan mampu memberikan kontribusi baik secara teoritis dalam pengembangan metode prediksi permintaan berbasis machine learning, maupun secara praktis dalam mendukung efisiensi operasional layanan pengiriman makanan melalui pengelolaan stok yang lebih optimal dan adaptif terhadap fluktuasi pasar.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan algoritma Random Forest (RF) dalam konteks manajemen inventori layanan pengiriman makanan. Proses pengembangan sistem prediksi mengikuti metode CRISP-DM (Cross Industry Standard Process for Data Mining) yang terdiri dari enam tahapan utama [9]. Tahap pertama adalah pemahaman bisnis, yaitu mengidentifikasi permasalahan dalam pengelolaan inventori dan merumuskan tujuan prediksi permintaan. Tahap kedua, pemahaman data, dilakukan untuk mengevaluasi struktur dan karakteristik dataset yang digunakan. Selanjutnya, persiapan data mencakup pembersihan data, penggabungan beberapa file, serta transformasi fitur yang relevan. Tahap keempat adalah pemodelan, di mana algoritma Random Forest dibangun dan dibandingkan dengan model Decision Tree dan Linear Regression. Pada tahap evaluasi, model dianalisis menggunakan metrik MSE, RMSE, dan R^2 untuk menentukan performa terbaik. Terakhir, pada tahap deployment, model terbaik diintegrasikan ke dalam sistem manajemen

inventori berbasis web menggunakan Python Flask. Pendekatan CRISP-DM dipilih karena fleksibel dan banyak digunakan dalam proyek data mining berbasis data historis [1].

Data yang digunakan merupakan dataset publik yang diambil dari platform Kaggle dengan judul "Food Demand Forecasting". Dataset ini mencakup informasi permintaan makanan dari berbagai pusat distribusi (fulfillment centers) selama 145 minggu, dengan lebih dari 450.000 baris data. Fitur-fitur penting dalam dataset meliputi: tanggal pemesanan, ID makanan, ID pusat distribusi, kategori makanan, jenis makanan (perishable atau non-perishable), jumlah pemesanan historis, serta indikator promosi. Data ini selanjutnya dibersihkan dari nilai hilang, duplikat, dan outlier menggunakan metode interkuartil (IQR) untuk menjaga kualitas input model.

Pada tahap pemodelan, algoritma Random Forest dipilih karena kemampuannya menangani data besar dan kompleks dengan fitur heterogen, serta mengurangi risiko overfitting melalui ensemble learning [4]. Sebagai pembandingan, dua algoritma lainnya digunakan yaitu Decision Tree Regressor dan Linear Regression, untuk mengevaluasi kinerja relatif RF. Model diuji menggunakan teknik split data dengan proporsi 80:20 untuk data latih dan data uji. Proses pelatihan model dilakukan menggunakan scikit-learn di Python dengan parameter default dan eksperimen grid search terbatas untuk tuning.

Evaluasi model dilakukan dengan tiga metrik utama: Mean Squared Error (MSE), Root Mean Square Error (RMSE), dan R-squared (R^2).

Ketiga metrik ini memberikan gambaran komprehensif mengenai seberapa dekat prediksi terhadap nilai aktual serta kemampuan model menjelaskan variasi data [10]. Model dengan nilai MSE dan RMSE paling kecil serta R^2 mendekati 1 dipertimbangkan sebagai model terbaik untuk digunakan dalam integrasi sistem manajemen inventori.

Setelah diperoleh model prediksi terbaik, tahap selanjutnya adalah integrasi ke dalam sistem manajemen inventori berbasis web sederhana. Sistem ini dibangun menggunakan framework Python Flask untuk back-end dan Bootstrap untuk antarmuka pengguna. Fungsi utama sistem mencakup: prediksi permintaan berdasarkan input pengguna (tanggal dan ID makanan), perhitungan kebutuhan stok optimal berdasarkan hasil prediksi, dan pelaporan stok harian. Uji coba sistem dilakukan secara simulatif untuk menilai akurasi, kemudahan penggunaan, dan efektivitas operasional.

III. HASIL DAN PEMBAHASAN

Pada bagian ini, akan dipaparkan hasil analisis data, evaluasi model prediksi yang dikembangkan, serta pembahasan mengenai temuan-temuan kunci dari penelitian. Penelitian ini menggunakan beberapa dataset utama, yaitu `train.csv`, `meal_info.csv`, dan `fulfillment_center_info.csv`, yang digabungkan untuk membentuk satu dataset final. Dataset ini mencakup berbagai variabel yang memiliki peran penting dalam proses prediksi jumlah pesanan makanan (`num_orders`). Di Tabel 1 berikut merangkum deskripsi setiap variabel dalam dataset.

Tabel 1. Deskripsi Variabel Dataset

Nama Variabel	Jenis Data	Satuan	Peran	Keterangan
<code>id</code>	<code>int64</code>	-	Identifier	ID unik untuk setiap baris (dihapus saat pemodelan).
<code>week_center_id</code>	<code>int64</code>	Minggu ke-	Fitur Waktu	Minggu dalam dataset.
<code>meal_id</code>	<code>int64</code>	-	Fitur Kategori	ID pusat distribusi. Dikelompokkan ke dalam beberapa rentang kategori.
<code>checkout_price</code>	<code>float64</code>	Dolar US	Fitur Numerik	ID jenis makanan. Dikelompokkan ke dalam beberapa rentang kategori.
<code>base_price</code>	<code>float64</code>	Dolar US	Fitur Numerik	Harga akhir setelah diskon. Nilai outlier telah ditangani.
<code>emailer_for_promotion</code>	<code>int64</code>	Boolean (0/1)	Fitur Biner	Harga dasar sebelum diskon.
<code>homepage_featured</code>	<code>int64</code>	Boolean (0/1)	Fitur Biner	1 jika ada promosi via email, 0 jika tidak.
<code>num_orders</code>	<code>int64</code>	Unit Pesanan	Target	1 jika ditampilkan di homepage, 0 jika tidak.
				Variabel target; Jumlah total pesanan.

Setelah memahami struktur data, proses pemodelan dilakukan dengan menggunakan algoritma *Random Forest Regressor*, yang kemudian dibandingkan dengan dua model lainnya yaitu *Decision Tree* dan *Linear Regression*. Evaluasi model menggunakan tiga metrik yaitu Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan R-squared (R^2). Hasilnya menunjukkan bahwa Random Forest unggul dengan MSE sebesar 109,570.40, RMSE sebesar 331.01, dan R^2 sebesar 0.26. Decision Tree mencatat MSE 110,773.22, RMSE 332.82, dan R^2 0.26, sementara Linear Regression mencatat MSE 111,857.88, RMSE 334.45, dan R^2 0.22.

Model Random Forest terbukti lebih akurat karena mampu menangkap hubungan kompleks dan non-linear antar fitur dalam data, termasuk kombinasi variabel waktu, numerik, biner, dan kategori. Teknik ensemble pada Random Forest juga meningkatkan generalisasi dan ketahanan terhadap overfitting dibandingkan model tunggal seperti Decision Tree.

Dalam implementasinya, model ini dapat digunakan untuk mendukung keputusan operasional, seperti memperkirakan kebutuhan stok makanan pada masing-masing pusat distribusi setiap minggunya. Visualisasi pendukung seperti grafik garis antara prediksi dan aktual, grafik batang *feature importance*, serta heatmap korelasi antar variabel numerik akan sangat membantu dalam menganalisis dan menyampaikan hasil model kepada pengambil keputusan.

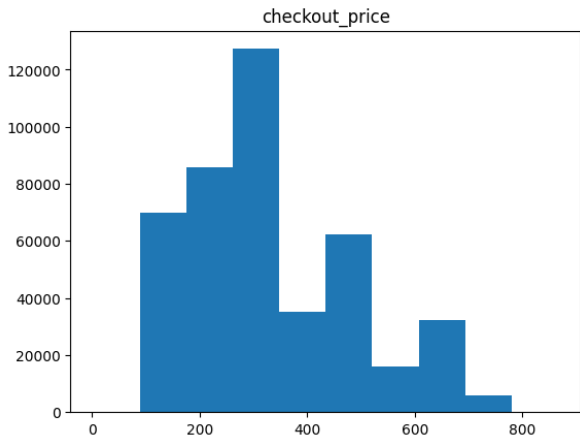
Hasil prediksi yang dihasilkan oleh model Random Forest selanjutnya digunakan untuk mendukung proses optimasi stok pada pusat distribusi makanan. Prediksi jumlah pesanan mingguan memungkinkan sistem untuk memperkirakan kebutuhan stok secara lebih presisi, sehingga pengadaan bahan baku dapat disesuaikan dengan permintaan aktual yang diproyeksikan. Hal ini penting dalam konteks makanan perishable, karena dapat mengurangi risiko pemborosan akibat kelebihan stok dan

menghindari kekurangan yang berdampak pada hilangnya potensi pendapatan. Dengan memanfaatkan hasil prediksi ini, sistem memberikan rekomendasi jumlah stok optimal untuk setiap ID makanan dan periode waktu tertentu, yang ditampilkan dalam antarmuka pengguna berbasis web. Meskipun sistem belum menggunakan metode optimasi matematis eksplisit, penggunaan prediksi permintaan sebagai dasar pengambilan keputusan stok sudah merepresentasikan bentuk optimasi operasional berbasis data historis.

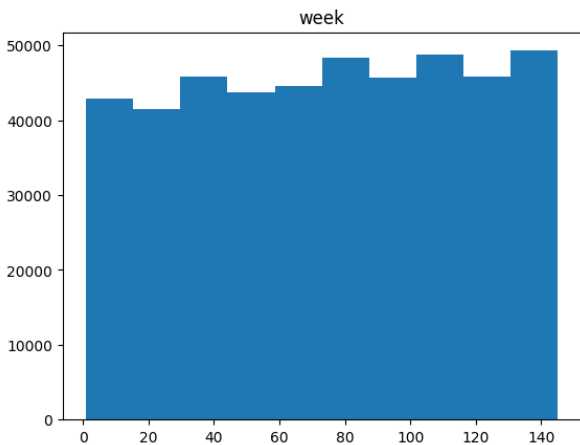
Jika dibandingkan dengan penelitian oleh Panda dan Mohanty [6], yang menggunakan model regresi dalam konteks rantai pasok makanan dengan hasil akurasi tinggi, nilai R^2 dalam penelitian ini (0.26) memang lebih rendah. Hal ini kemungkinan disebabkan oleh keterbatasan fitur dalam dataset yang digunakan, terutama tidak adanya variabel eksternal seperti cuaca atau hari libur yang diketahui dapat meningkatkan akurasi prediksi permintaan. Selain itu, berbeda dengan pendekatan klasifikasi yang digunakan oleh Alsanad [4] dalam memprediksi permintaan harian secara kategorikal, penelitian ini fokus pada prediksi nilai numerik (regression-based) yang lebih relevan untuk estimasi volume stok. Sementara Alsanad menunjukkan akurasi tinggi dalam klasifikasi permintaan tinggi/rendah, pendekatan regresi di penelitian ini memungkinkan estimasi yang lebih terukur dan terintegrasi langsung ke sistem inventori berbasis web.

Dengan demikian, hasil penelitian ini menegaskan bahwa pendekatan machine learning, khususnya dengan algoritma Random Forest, efektif dalam memprediksi jumlah pesanan makanan secara akurat dan efisien. Model ini dapat menjadi fondasi bagi sistem prediksi permintaan dalam rantai pasok makanan modern berbasis data.

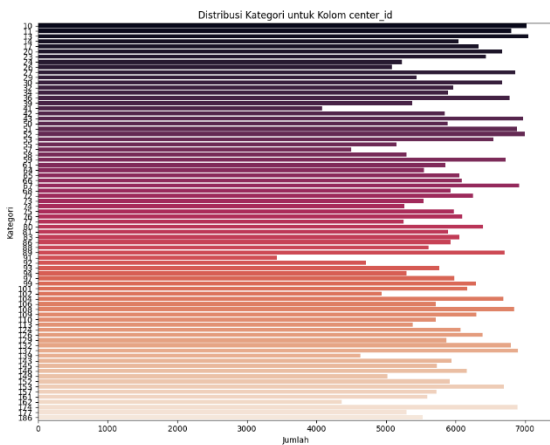
Sebagai bagian dari eksplorasi data sebelum pemodelan, dilakukan visualisasi terhadap fitur-fitur penting:



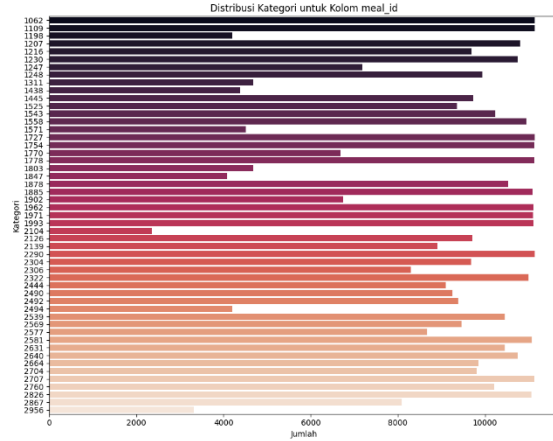
Gambar 1. Histogram Checkout_price. Menampilkan distribusi frekuensi harga akhir produk. Terlihat bahwa mayoritas harga terkonsentrasi di kisaran 250–350.



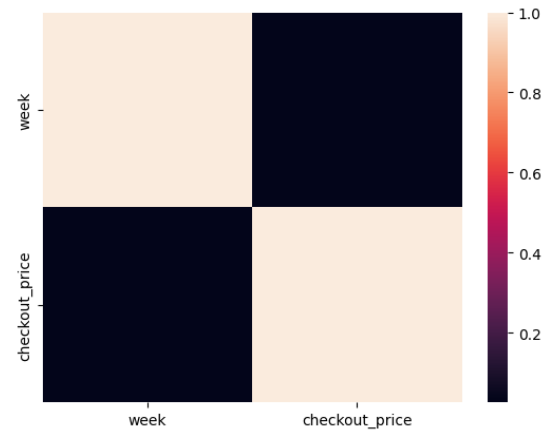
Gambar 2. Histogram Week. Menunjukkan bahwa distribusi data dari minggu ke-1 hingga ke-145 relatif merata, yang menandakan tidak ada bias temporal yang signifikan dalam pengumpulan data.



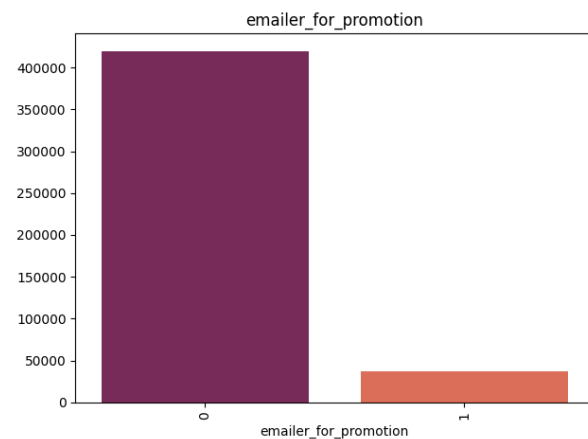
Gambar 3. Distribusi Center_id. Menunjukkan distribusi data yang relatif merata di seluruh pusat distribusi makanan (center_id). Artinya, tidak ada satu pusat pun yang mendominasi dataset secara ekstrem



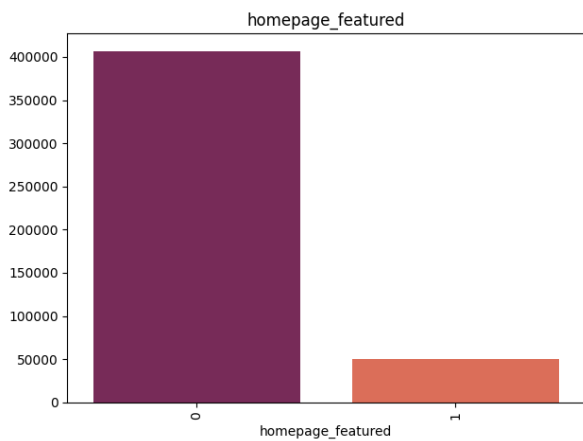
Gambar 4. Distribusi Meal_id. Menunjukkan distribusi jumlah pesanan untuk setiap jenis makanan (meal_id) dalam dataset. Terlihat adanya variasi yang signifikan dalam popularitas antar makanan. Beberapa meal_id memiliki frekuensi pesanan yang sangat tinggi (di atas 10.000), sementara yang lain jauh lebih rendah (di bawah 4.000).



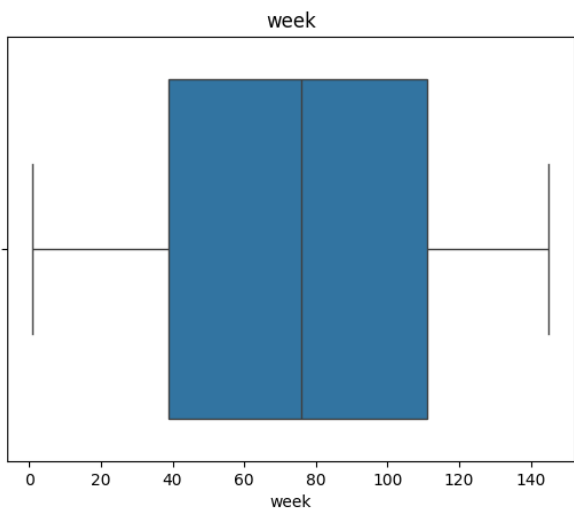
Gambar 5. Heatmap Korelasi. Menunjukkan matriks korelasi visual antara variabel week dan checkout price. Gambar ini mengindikasikan hubungan yang sangat lemah antara kedua variabel tersebut.



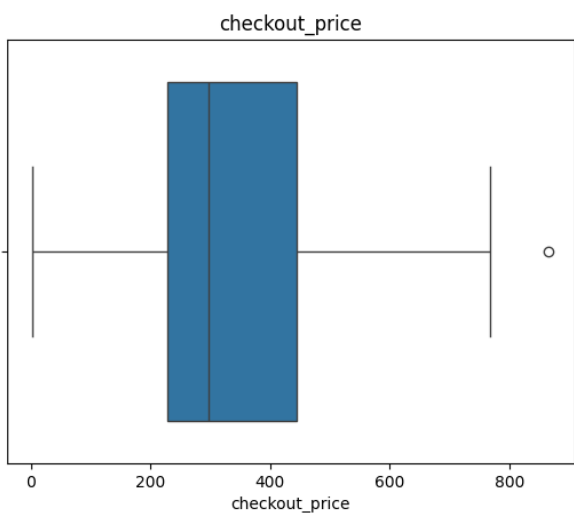
Gambar 6. Diagram emailer_for_promotion. Menunjukkan bahwa mayoritas data tidak menggunakan promosi melalui email (nilai 0) dibandingkan yang menggunakan (nilai 1).



Gambar 7. Diagram Homepage_Featured. Menunjukkan bahwa sebagian besar makanan tidak ditampilkan di halaman utama (homepage) (nilai 0) dibandingkan yang ditampilkan (nilai 1).



Gambar 8. Boxplot Week. Mnunjukkan distribusi variabel week yang seimbang dan simetris, dengan rentang data yang konsisten dari minggu ke-1 hingga ke-145. Tidak adanya outlier yang terdeteksi menandakan bahwa fitur waktu ini bersih dan dapat langsung digunakan untuk pemodelan tanpa memerlukan penanganan khusus.



Gambar 9. Boxplot checkout_price. Menunjukkan distribusi harga akhir (checkout_price), yang mengidentifikasi adanya outlier atau data pencilan dengan nilai yang sangat tinggi. Kehadiran outlier ini dapat mempengaruhi akurasi model prediksi, sehingga perlu ditangani terlebih dahulu melalui teknik seperti penghapusan atau normalisasi data sebelum pemodelan.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengevaluasi model prediksi permintaan makanan berbasis algoritma Random Forest dalam konteks manajemen inventori layanan pengiriman makanan. Model ini dibangun dengan pendekatan CRISP-DM dan menggunakan dataset historis dari platform Kaggle yang mencakup variabel-variabel numerik, kategori, dan biner. Berdasarkan hasil evaluasi dengan metrik MSE, RMSE, dan R², Random Forest menunjukkan performa terbaik dibandingkan dengan Decision Tree dan Linear Regression, dengan MSE sebesar 109,570.40, RMSE sebesar 331.01, dan R² mencapai 0.26. Meskipun menjadi model terbaik dalam perbandingan ini, skor R² sebesar 0.26 mengindikasikan bahwa model baru mampu menjelaskan sekitar 26% dari variabilitas permintaan. Hal ini menunjukkan bahwa meskipun model dapat menangkap beberapa pola, masih terdapat faktor-faktor lain yang signifikan dalam data yang belum teridentifikasi. Meskipun demikian, penelitian ini memberikan kontribusi penting dalam membangun fondasi metodologi untuk sistem manajemen inventori berbasis data. Sistem yang dikembangkan mampu memproyeksikan jumlah pesanan mingguan dan menghitung kebutuhan stok, membuktikan bahwa pendekatan machine learning dapat diintegrasikan secara praktis untuk mendukung efisiensi operasional. Keunggulan Random Forest dalam menangani data non-linear dan ketahanannya terhadap overfitting menjadikannya titik awal yang kuat untuk pengembangan lebih lanjut. Keterbatasan utama dari model saat ini adalah skor R² yang masih rendah, yang kemungkinan disebabkan oleh tidak adanya variabel-variabel penting seperti data cuaca, hari libur nasional, atau event lokal yang dapat mempengaruhi permintaan secara signifikan. Untuk pengembangan selanjutnya, disarankan untuk

memperkaya dataset dengan fitur-fitur eksternal tersebut, serta mengeksplorasi algoritma yang lebih canggih seperti Gradient Boosting atau model time-series (misalnya LSTM) untuk meningkatkan akurasi prediksi. Implementasi dashboard interaktif dan integrasi data real-time juga tetap menjadi langkah pengembangan yang relevan untuk membuat sistem lebih adaptif di lapangan.

Secara keseluruhan, penelitian ini yang didukung oleh visualisasi data untuk mengungkap karakteristik dataset menegaskan bahwa pendekatan machine learning, khususnya Random Forest, dapat menjadi fondasi yang kuat. Implementasi sistem ini berpotensi besar untuk dikembangkan menjadi alat prediksi yang lebih presisi dan responsif, seiring dengan penambahan data dan fitur yang lebih kaya di masa depan.

REFERENSI

- [1] A. Alhidayatullah, S. Putri, N. Hastarita, and A. R. Regina, "ANALISIS SUPPLY CHAIN PADA USAHA MIKRO: STUDI KASUS RICE BOWL TRUSTMEE DI KOTA METRO," *Jurnal Inspirasi Ilmu Manajemen*, vol. 3, no. 2, pp. 98–104, Feb. 2025, doi: 10.32897/JIIM.2025.3.2.4086.
- [2] V. Ikpe and M. Shamsuddoha, "Functional Model of Supply Chain Waste Reduction and Control Strategies for Retailers—The USA Retail Industry," *Logistics 2024, Vol. 8, Page 22*, vol. 8, no. 1, p. 22, Feb. 2024, doi: 10.3390/LOGISTICS8010022.
- [3] S. Teerasoponpong and A. Sopadang, "Decision support system for adaptive sourcing and inventory management in small- and medium-sized enterprises," *Robot Comput Integr Manuf*, vol. 73, p. 102226, Feb. 2022, doi: 10.1016/J.RCIM.2021.102226.
- [4] A. Alsanad, "Forecasting Daily Demand of Orders Using Random Forest Classifier," 2018.
- [5] R. M. van Steenberg, "New Product Forecasting with Analogous Products: Applying Random Forest and Quantile Regression Forest to forecasting and inventory management," 2019.
- [6] S. K. Panda and S. N. Mohanty, "Time Series Forecasting and Modeling of Food Demand Supply Chain Based on Regressors Analysis," *IEEE Access*, vol. 11, pp. 42679–42700, 2023, doi: 10.1109/ACCESS.2023.3266275.
- [7] C. Schröder, F. Kruse, and J. M. Gómez, "A Systematic Literature Review on Applying CRISP-DM Process Model," *Procedia Comput Sci*, vol. 181, pp. 526–534, Jan. 2021, doi: 10.1016/J.PROCS.2021.01.199.
- [8] P. Appiahene, Y. M. Missah, and U. Najim, "Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks," *Advances in Fuzzy Systems*, vol. 2020, no. 1, p. 8581202, Jan. 2020, doi: 10.1155/2020/8581202.
- [9] J. Torres, D. Carpio, and V. Parasi, "Model to Predict Inventory Demand in Retail SMEs Using CRISP-DM and Machine Learning," *Proceedings of the 2024 IEEE 31st International Conference on Electronics, Electrical Engineering and Computing, INTERCON 2024*, 2024, doi: 10.1109/INTERCON63140.2024.10833461.
- [10] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput Sci*, vol. 7, pp. 1–24, Jul. 2021, doi: 10.7717/PEERJ-CS.623/SUP