

Price Trend Prediction and Product Recommendation Using Machine Learning Approach

Prediksi Tren Harga dan Rekomendasi Produk Menggunakan Pendekatan Machine Learning

Fedro Rizkyana Padila¹, Siti Yuliyanti^{2*}, Muhammad Al Husaini³

^{1,2,3} Informatika, Universitas Siliwangi, Indonesia

¹217006060@student.unsil.ac.id, ^{2*}sitiyulianti@unsil.ac.id, ³alhusaini@unsil.ac.id

*: Penulis korespondensi (corresponding author)

Informasi Artikel

Received: December 2026

Revised: December 2025

Accepted: January 2026

Published: February 2026

Abstract

Purpose: This study aims to develop an accurate and relevant machine learning-based model for price trend prediction and product recommendation using retail sales data, in order to support operational efficiency and marketing strategies.

Design/methodology/approach: The study utilizes sales data from two supermarkets spanning the years 2021–2024. Five machine learning algorithms are compared, namely Random Forest, XGBoost, Gradient Boosting, Decision Tree, and Linear Regression, with evaluation metrics including R^2 , MAPE, and RMSE. The best-performing model is then optimized using feature importance and feature selection methods. For the recommendation system, a Content-Based Filtering approach is employed based on similarity in product names and categories.

Findings/result With an R^2 of 80%, MAPE of 25%, and RMSE of 1.07 at the start, the Random Forest model performed the best. Following feature optimization, it greatly improved, reaching R^2 of 92%, MAPE of 13%, and RMSE of 0.84. The Content-Based Filtering recommendation system successfully provided relevant suggestions, with the highest cosine similarity reaching 0.4268.

Originality/value/state of the art: This study is grounded in and extends the findings of previous research.

Abstrak

Keywords: price prediction; product recommendation; machine learning;

Tujuan: Penelitian ini bertujuan untuk mengembangkan model prediksi tren harga dan sistem rekomendasi produk

random forest; content-based filtering
Kata kunci: prediksi harga;
rekomendasi produk; *machine*
learning; random forest; content-based
filtering

pada data penjualan ritel dengan pendekatan machine learning yang akurat dan relevan guna mendukung efisiensi operasional dan strategi pemasaran.

Perancangan/metode/pendekatan: Dataset penjualan dari dua supermarket tahun 2021–2024 digunakan dalam penelitian ini. Lima algoritma machine learning dibandingkan, yaitu Random Forest, XGBoost, Gradient Boosting, Decision Tree, dan Regresi Linier, dengan evaluasi menggunakan metrik R^2 , MAPE, dan RMSE. Model terbaik dioptimasi menggunakan metode feature importance dan feature selection. Untuk sistem rekomendasi, digunakan pendekatan Content-Based Filtering berdasarkan kemiripan nama dan kategori produk. Hasil: Model Random Forest memberikan performa awal terbaik (R^2 sebesar 80%, MAPE 25%, dan RMSE 1,07) dan meningkat signifikan setelah optimasi fitur menjadi R^2 sebesar 92%, MAPE 13%, dan RMSE 0,84. Sistem rekomendasi berbasis Content-Based Filtering berhasil memberikan hasil yang relevan dengan cosine similarity tertinggi mencapai 0,4268.

Keaslian/ *state of the art*: Penelitian ini berlandaskan hasil dari penelitian-penelitian terdahulu.

1. Pendahuluan

Berbagai sektor industri, termasuk ritel dan makanan ringan, telah mengalami transformasi signifikan akibat kemajuan teknologi informasi serta meningkatnya pemanfaatan data dalam pengambilan keputusan [1]. Dalam menghadapi persaingan yang semakin ketat, pelaku bisnis dituntut untuk menerapkan strategi penjualan yang fleksibel guna menjaga daya saing dan meningkatkan keuntungan. Salah satu tantangan utama yang dihadapi oleh sektor ritel adalah fluktuasi harga dan perubahan tren pasar yang sulit diprediksi [2]. Ketidakpastian harga dapat menyebabkan inefisiensi stok, penurunan margin keuntungan, dan menghambat pencapaian target penjualan.

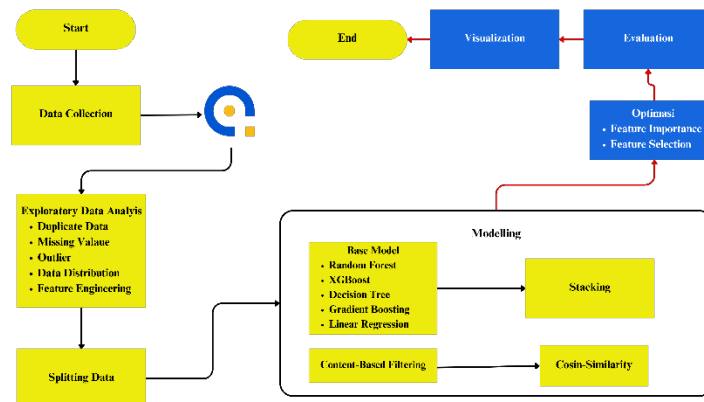
Kemampuan untuk memprediksi harga produk secara akurat adalah komponen penting dalam membuat keputusan strategis dan mengembangkan rencana pemasaran yang tepat sasaran dalam situasi seperti ini. Di sisi lain, sistem rekomendasi produk yang disesuaikan dengan preferensi konsumen juga berperan penting dalam meningkatkan loyalitas pelanggan serta nilai transaksi [3]. Oleh karena itu, ada kemungkinan bahwa sistem prediksi harga dan rekomendasi produk dapat diintegrasikan untuk mendukung efisiensi operasional dan pertumbuhan perusahaan ritel.

Metode pembelajaran mesin telah banyak digunakan untuk membangun strategi pemasaran yang lebih efisien dan proses prediksi penjualan. Algoritma seperti Random Forest, XGBoost, dan Gradient Boosting telah terbukti dapat menghasilkan prediksi yang akurat dan mengidentifikasi pola yang kompleks pada data historis [4], [5]. Studi Warjiyono menemukan bahwa algoritma Random Forest memiliki akurasi sebesar 75%, meskipun model yang

digunakan belum dioptimalkan [6]. tudy Hidayah menemukan bahwa regresi linier mencapai akurasi sebesar 71%, Decision Tree sebesar 59%, dan Random Forest mencapai akurasi sebesar 86% [7]. Anwari mencapai 86% dengan menggunakan model ARIMA [8]. Nasyuli menemukan bahwa model Boosting Gradient dan Linear Regression memiliki akurasi 90% [9]. Selain itu, penelitian tersebut menyarankan penggunaan proporsi antara data pelatihan dan data uji sebesar 80:20, dan penerapan strategi untuk mengoptimalkan model. Sebaliknya, sistem rekomendasi produk telah menjadi bagian penting dari industri ritel kontemporer; rekomendasi dibagi menjadi pendekatan seperti penyaringan berbasis konten, penyaringan kolaboratif, dan penyaringan berbasis item [10]. Studi sebelumnya oleh Yanisa menunjukkan bahwa pendekatan Filter Bersama hanya dapat menghasilkan akurasi 20%, sehingga diperlukan pendekatan alternatif yang lebih kontekstual dan relevan. Dengan mempertimbangkan karakteristik produk dan preferensi pengguna berdasarkan data historis, filtrasi berbasis konten adalah pilihan yang bagus [11].

Penelitian ini bertujuan untuk mengoptimalkan akurasi prediksi harga dan sistem rekomendasi produk dengan membandingkan berbagai algoritma pengajaran mesin, seperti Random Forest, XGBoost, Gradient Boosting, Decision Tree, dan Linear Regression. Metode pemilihan fitur dan pentingnya fitur akan digunakan untuk meningkatkan performa prediksi. Selain itu, sistem rekomendasi produk berbasis filter konten akan dikembangkan. Ini akan membuat rekomendasi yang lebih relevan dan membantu strategi pemasaran yang lebih baik.

2. Metode/Perancangan



Gambar 1. Tahapan Penelitian

Gambar 1 menunjukkan metodologi standar dalam penelitian untuk menerapkan *machine learning*. Penelitian melakukan banyak hal.

2.1. Data Collection

Pengumpulan data berfokus pada mengidentifikasi, mengumpulkan, dan menganalisis kumpulan data yang dapat mencapai tujuan proyek [12]. Pada penelitian ini, dataset didapatkan dari Mitra Startup Campus, yang berisi data historis penjualan dari dua supermarket dalam rentang waktu 2021-2024.

2.2. Exploratory Data Analysis

Exploratory Data Analysis (EDA) bertujuan untuk mempersiapkan kumpulan data akhir untuk implementasi model [13]. Tahapan EDA meliputi proses penanganan data duplikat untuk menangani nilai ganda. Nilai yang hilang, menangani nilai kosong, distribusi data dengan tujuan untuk mendapatkan wawasan dari setiap variabel. Outlier, untuk melihat nilai yang sangat berbeda pada setiap variabel, dan rekayasa fitur untuk membantu proses optimasi.

2.3. Splitting Data

Dataset yang telah melalui tahap *Exploratory Data Analysis*, mengalami proses pemisahan data dengan cara membagi dataset menjadi 80% data latih dan 20% data latih [7].

2.4. Modelling

Dalam penelitian ini, Random Forest, XGBoost, Gradient Boosting, Decision Tree, dan Linear Regression digunakan untuk mencapai akurasi awal yang cukup tinggi untuk mencapai tahap optimasi [14], [15].

2.5. Optimasi

Implementasi model sebelumnya bertujuan untuk mendapatkan akurasi model awal yang cukup akurat, kemudian dilakukan tahapan optimasi untuk meningkatkan performa model [16]. Pada proses optimasi, rekayasa fitur sangat membantu ketika melakukan pengujian variabel dengan menggunakan *feature importance* [17]. Jika nilai batas setiap variabel > 0.01 , maka disarankan untuk menghapusnya karena tidak berpengaruh terhadap nilai akurasi [18]. Kemudian setelah melakukan pengujian variabel, dilakukan seleksi fitur dengan menggunakan metode filter untuk menggunakan variabel yang berpengaruh terhadap hasil akurasi awal dari proses *feature importance* [19].

2.6. Evaluation

2.6.1. R^2 (R-Squared)

R^2 digunakan untuk mengukur keakuratan prediksi algoritma [20]. Kisaran nilai R^2 terdiri dari angka 0 hingga 1, yang menunjukkan seberapa besar kombinasi variabel independen mempengaruhi nilai independen. Rumus persamaan R^2 :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (1)$$

2.6.2. MAPE (Mean Absolute Percentage Error)

MAPE (*Mean Absolute Percentage Error*) dapat membantu mengukur kesalahan dalam prediksi [20]. Rumus persamaan MAPE:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i - \hat{y}_i|}{y_i} \times 100 \right) \quad (2)$$

2.6.3. RMSE (Root Mean Squared Error)

RMSE (*Root Mean Squared Error*) memiliki fungsi untuk mengevaluasi seberapa baik model yang digunakan dalam melakukan prediksi. Rumus persamaan RMSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|} \quad (3)$$

2.7. Visualization

Model yang telah diimplementasikan akan dipresentasikan dengan menggunakan *platform* Tableau, untuk memudahkan masyarakat dalam memahami hasil pemodelan.

3. Hasil dan Pembahasan

3.1. Modelling

Dataset yang digunakan dalam proses pemodelan sebelumnya telah melalui tahap pemisahan data. **Tabel 1**, adalah variabel-variabel yang digunakan dalam implementasi model.

Tabel 1. Variabel Implementasi Model

Variabel	Dtype
nama_toko	int64
harga_pound	float64
harga_per_unit	float64
unit	int64
nama	object
kategori	int64
brand_sendiri	int64
tanggal	object
harga_7d_rolling_mean	float

Implementasi model (*Random Forest, Decision Tree, XGBoost, Gradient Boosting, Regresi Linier*), dengan menggunakan parameter pelatihan. **Tabel 2**, adalah parameter yang digunakan.

Tabel 2. Parameter Model

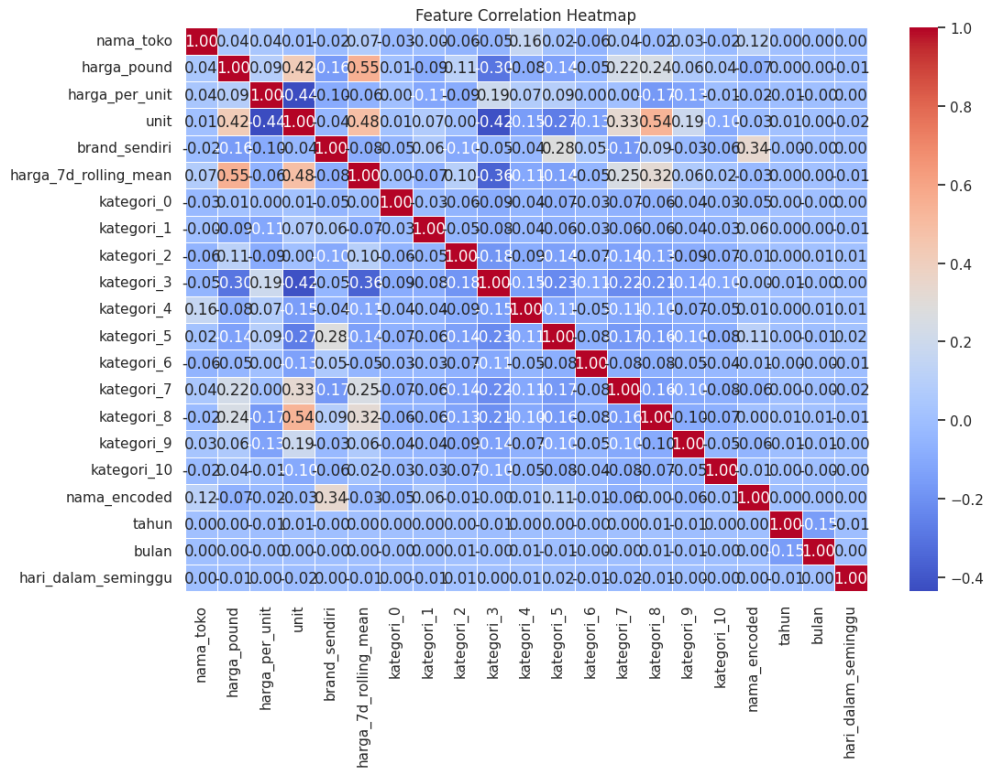
n_estimator	random_state
100	42

Tabel 2, konfigurasi parameter n_estimator (100) dan random_state (42), berfungsi untuk memastikan hasil yang konsisten ketika dijalankan kembali, dan untuk menangani data yang kompleks untuk menghasilkan prediksi yang stabil dan mengurangi *overfitting* [21], [22].

3.2. Optimasi

3.2.1. Feature Importance

Fitur ini penting berfungsi untuk pengujian variabel yang dibantu dengan menggunakan heatmap untuk memudahkan proses analisis. **Gambar 2**, adalah heatmap pengujian variabel.



Gambar 2. Heatmap Feature Importance

3.2.2. Feature Selection

Hasil dari pengujian variabel dengan menggunakan *feature importance*, selanjutnya akan dilakukan seleksi fitur. Variabel yang memiliki nilai < 0.01 (tahun, bulan, hari_dalam_seminggu, dan harga_per_unit), akan dihilangkan karena tidak berpengaruh terhadap nilai akurasi.

3.3. Evaluation

3.3.1. Prediksi

Implementasi model bertujuan untuk memprediksi harga suatu produk. **Tabel 3**, merupakan hasil akurasi awal model. Hasil akurasi awal pemodelan tersebut akan dilakukan tahap lebih lanjut terhadap model yang mempunyai akurasi awal yang cukup akurat untuk ditingkatkan performa model dengan dilakukannya tahapan optimasi terhadap model terbaik.

Tabel 3. Hasil Akurasi Awal Prediksi

Metric	Random Forest	Decision Tree	XGBoost	Gradient Boosting	Regresi Linear
R^2	0.802325	0.750271	0.716892	0.622142	0.622142
MAPE	0.254961	0.262946	0.393922	0.479370	0.479370
RMSE	1.076.416	1.141.194	1.177.552	1.265.681	1.265.681

Pada Tabel 3, hasil penilaian menggunakan metrik R^2 , MAPE, dan RMSE menunjukkan bahwa *Random Forest* merupakan model dengan performa terbaik. Model ini mencatat nilai R^2 sebesar 80%, MAPE terendah sebesar 25%, dan RMSE sebesar 1,076, yang menunjukkan

akurasi dan stabilitas prediksi yang unggul. Dengan demikian, Random Forest dipilih untuk melanjutkan ke tahap optimasi.

3.3.2. Optimasi Prediksi

Pada pemodelan awal, algoritma Random Forest memiliki akurasi awal yang cukup baik, dan dilakukan proses optimasi untuk meningkatkan performa model. **Tabel 4**, menunjukkan hasil optimasi tersebut.

Tabel 4. Optimasi Prediksi

Metric	R ²	MAPE	RMS E
Random Forest	0.9251	0.1382	0.8445

Tabel 4, menyajikan hasil optimasi model *Random Forest* yang dilakukan dengan menggunakan metode *feature importance* dan *feature selection*. Hasilnya menunjukkan peningkatan yang signifikan pada performa model. Nilai R² meningkat menjadi 92%, MAPE menurun menjadi 13%, dan RMSE menurun menjadi 0.84, mengindikasikan bahwa rata-rata kesalahan prediksi menjadi lebih kecil. Sebelumnya, model mencatatkan nilai R² sebesar 80% (kategori sangat kuat), MAPE sebesar 25% (kategori cukup), dan RMSE sebesar 1,07. Setelah proses optimasi, performa meningkat sebesar 15,30% untuk R², dengan MAPE dan RMSE yang juga menunjukkan peningkatan, masing-masing masuk ke dalam kategori baik [23].

Hasil optimasi menggunakan *feature importance* dan *feature selection* menunjukkan performa yang baik. Dapat dilihat pada **Tabel 5**, adalah hasil perbandingan model dengan penelitian terdahulu.

Tabel 5. Perbandingan Akurasi Model

Model	Article	Accuracy
Random Forest	[6]	75%
ARIMA	[8]	86%
Gradient Boosting + Regresi Linear	[9]	90%
Decision Tree + Feature Scalling	[24]	91%
Random Forest+Feature Importance and Feature Selection	our	92%

Tabel 5, menyajikan perbandingan hasil implementasi model *machine learning* antara penelitian ini dengan beberapa penelitian sebelumnya. Hasil tersebut menunjukkan bahwa optimasi fitur memberikan kontribusi yang signifikan dalam meningkatkan performa prediksi dibandingkan dengan pendekatan pada penelitian-penelitian sebelumnya, model *Random Forest* yang telah dioptimasi dengan menggunakan *feature importance* dan *feature selection* mencapai akurasi tertinggi yaitu sebesar 92%.

3.3.3. Rekomendasi

Hasil rekomendasi menggunakan filter berbasis konten, memberikan rekomendasi kepada pengguna berdasarkan kemiripan nama dan kategori produk, dapat diamati di **Tabel 6**.

Tabel 6. Hasil Rekomendasi dengan Content-Based Filtering Produk Results *Carte D'Or Baileys Ice Cream Dessert Tub*

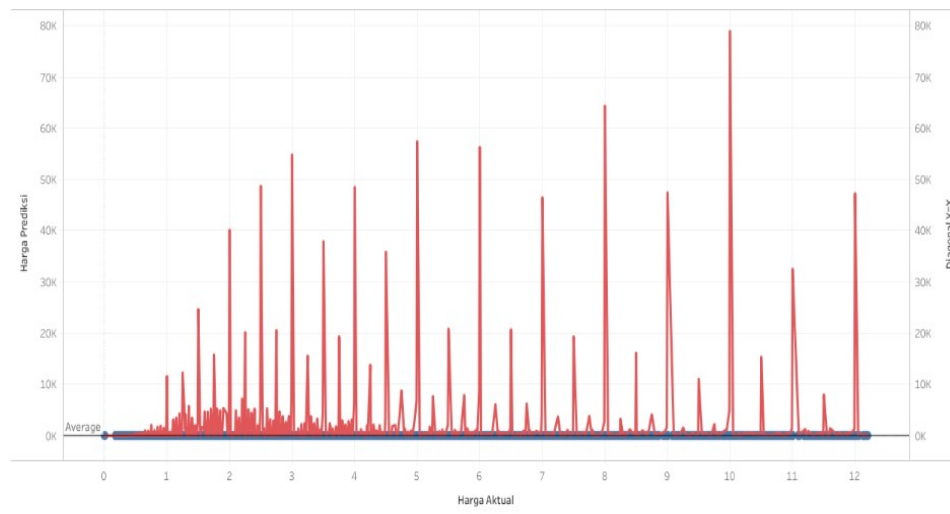
No	Nama Produk	Kategori	Harga Per Unit	Nama Toko	Cosine Similarity
1.	Ben & Jerry's Half Baked Chocolate Ice Cream Tub	6	11.8	0	0.4268
2.	Little Moons Mochi Ice Cream Italian Hazelnut	6	19.5	0	0.2885
3.	Sainsbury's Vanilla Flavour Ice Cream 1L	4	1.5	1	0.2730
4.	Carousel Ice Cream Cones x21	3	20.5	1	0.2582
5	Morrisons Tropical Swirl Ice Creams	6	5.3	0	0.2157

Berdasarkan **Tabel 6**, produk yang direkomendasikan memiliki nilai kemiripan (*cosine similarity*) paling tinggi dengan produk acuan, yaitu produk *Carte D'Or Baileys Ice Cream Dessert Tub*. Hasil perhitungan *cosine similarity* berdasarkan kemiripan nama dan kategori produk, yaitu produk *Ben & Jerry's Half Baked Chocolate Ice Cream Tub* memiliki tingkat kemiripan yang tinggi dengan produk acuan (0.4268). Oleh karena itu, produk *Ben & Jerry's Half Baked Chocolate Ice Cream Tub* sangat direkomendasikan karena memiliki tingkat kemiripan yang cukup tinggi dengan produk acuan.

3.4. Visualization

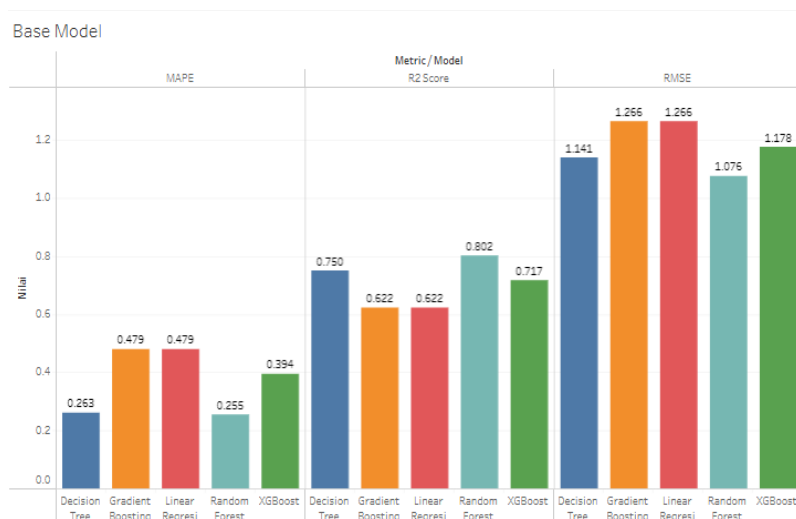
Visualisasi merupakan bagian dari proses analisis yang berperan penting dalam membantu memahami pola, tren, dan mendapatkan insight dari sebuah data. Pada penelitian ini, visualisasi digunakan pada beberapa tahapan, diantaranya untuk menggambarkan kondisi awal data, mengevaluasi performa model, dan menyajikan hasil sistem rekomendasi yang dihasilkan.

Gambar 3, menunjukkan interpretasi hasil, mayoritas titik prediksi berada di sepanjang garis diagonal $Y = X$ (biru), menghasilkan banyak nilai prediksi yang mendekati nilai aktual, yang berarti model ini cukup akurat dalam memprediksi harga.



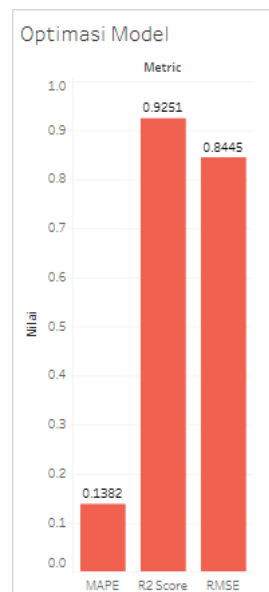
Gambar 3. Visualisasi Perbandingan Harga Aktuan dan Prediksi

Hasil pemodelan awal (*base model*) dilakukan dengan menggunakan parameter-parameter yang dibutuhkan pada masing-masing model. Gambar 4, merupakan hasil pemodelan awal dalam memprediksi harga yang menunjukkan model *Random Forest* memiliki akurasi awal yang baik dengan nilai R2 sebesar 80%, MAPE 25%, dan RMSE 1.07.



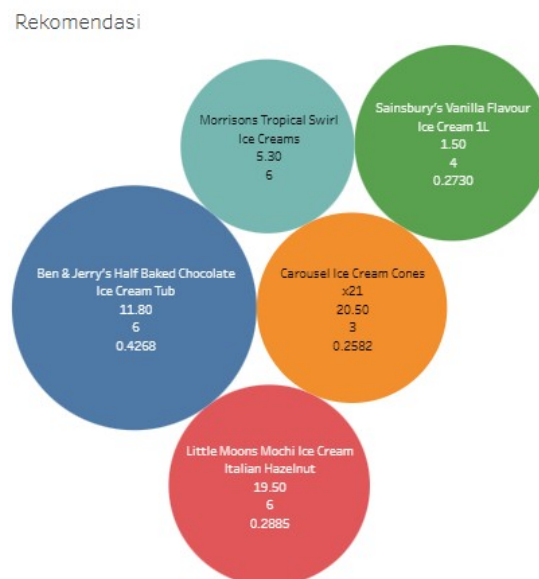
Gambar 4. Perbandingan Akurasi Awal Model

Hasil optimasi model *Random Forest* dengan menggunakan *feature importance* dan *feature selection* mengalami peningkatan sebesar 15%, dengan nilai R2 sebesar 92%, MAPE sebesar 13%, dan RMSE sebesar 0.84%. Dapat dilihat pada Gambar 5, merupakan visualisasi hasil optimasi model.



Gambar 5. Hasil Optimasi Random Forest

Hasil rekomendasi dengan menggunakan metode *Content-Based Filtering* yang ditunjukkan pada **Gambar 6** merupakan visualisasi rekomendasi produk dengan mengacu pada produk *Carte D'Or Ice Cream Dessert Tub*. Hasil rekomendasi berdasarkan perhitungan *cosine similarity* produk *Ben & Jerry's Half Baked Chocolate Ice Cream Tub* memiliki nilai kemiripan paling tinggi (0.4268), sehingga menjadi rekomendasi utama meskipun harganya relatif tinggi.



Gambar 6. Visualisasi Hasil Rekomendasi Produk *Carte D'Or Baileys Ice Cream Dessert Tub*

4. Kesimpulan dan Saran

Penelitian ini menunjukkan bahwa Random Forest efektif dalam memprediksi tren harga dan memberikan rekomendasi produk pada data ritel. Model ini mencatat performa awal terbaik

(R^2 : 80%, MAPE: 25%, RMSE: 1,07) dan meningkat signifikan setelah optimasi fitur (R^2 : 92%, MAPE: 13%, RMSE: 0,84), menegaskan pentingnya optimasi fitur dalam meningkatkan akurasi. Rekomendasi berbasis Content-Based Filtering juga mampu memberikan hasil yang relevan berdasarkan nama dan kategori produk. Untuk pengembangan lebih lanjut, disarankan penerapan optimasi hyperparameter serta mempertimbangkan variabel eksternal dalam rekomendasi seperti kandungan gizi, dan komposisi untuk lebih optimal.

Daftar Pustaka

- [1] S. Maesaroh, R. R. Lubis, L. N. Husna, R. Widyaningsih, and R. Susilawati, "Efektivitas Implementasi Manajemen Business Intelligence pada Industri 4.0," *ADI Bisnis Digit. Interdisiplin J.*, vol. 3, no. 2, pp. 1–8, 2022, doi: 10.34306/abdi.v3i2.764.
- [2] Y. Lisdarti, "MACHINE LEARNING UNTUK ANALISA POLA TREN PASAR DAN REDIKSI PERMINTAAN PRODUK MENGGUNAKAN METODE," pp. 62–68, 2023.
- [3] S. Sibuea and Y. B. Widodo, "Pengembangan Model Machine Learning untuk Rekomendasi Produk Berdasarkan Analisis Pola Pembelian," vol. 10, no. 2, pp. 567–583, 2024, [Online]. Available: <https://journal.thamrin.ac.id/index.php/jtik/article/view/2354/2254>
- [4] P. E-commerce and M. Firanti, "Penggunaan Algoritma Machine Learning dalam Prediksi," pp. 1–7, [Online]. Available: <http://www.circle-archive.com/index.php/carc/article/view/306/323>
- [5] T. Algoritma, "Optimisasi Harga menggunakan Machine Learning."
- [6] Warjiyono, A. Nur Rais, I. Alfarobi, S. Wira Hadi, and W. Kurniawan, "Analisa Prediksi Harga Jual Rumah Menggunakan Algoritma Random Forest Machine Learning," *JURISTEKNI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 6, no. 2, pp. 416–423, 2024.
- [7] F. Hidayah, S. J. Angesti, and Y. P. Widyastuti, "Prediksi Harga Rumah di Boston Menggunakan Metode Linear Regression, SVR, Decision Tree dan Random Forest Regression," pp. 1–9, 2024.
- [8] I. Anwari and I. Nurma Yulita, "Penggunaan Machine Learning Untuk Prediksi Harga Telur Ayam Ras Di Kota Bandung," *Communnity Dev. J.*, vol. 4, no. 3, pp. 6001–6005, 2023.
- [9] L. P. Nasyuli, I. Lubis, and A. M. Elhanafi, "Penerapan Model Machine Learning Algoritma Gradient Boosting dan Linear Regression Melakukan Prediksi Harga Kendaraan Bekas Application Of Machine Learning Models and Gradient Boosting Algorithms Doing Linear Regression Vehicle Price Prediction Used," *J. Ilmu Komput. dan Sist. Informasi(JIRSI)*, vol. 2, no. 2, pp. 299–310, 2023, [Online]. Available: <https://jurnal.unity-academy.sch.id/index.php/jirsi/index%0Ahttp://creativecommons.org/licenses/by-sa/4.0/>
- [10] Yaya Suharya, Y. Herdiana, N. Indah Putri, and Z. Munawar, "Sistem Rekomendasi Untuk Toko Online Kecil Dan Menengah," *Tematik*, vol. 8, no. 2, pp. 176–185, 2021,

doi: 10.38204/tematik.v8i2.683.

- [11] K. S. Yanisa Putri, I Made Agus Dwi Suarjaya, and Wayan Oger Vihikan, “Sistem Rekomendasi Skincare Menggunakan Metode Content Based Filtering dan Collaborative Filtering,” *Decod. J. Pendidik. Teknol. Inf.*, vol. 4, no. 3, pp. 764–774, 2024, doi: 10.51454/decode.v4i3.601.
- [12] A. Pratama, A. C. Nurcahyo, and L. Firgia, “Penerapan Machine Learning dengan Algoritma Logistik Regresi untuk Memprediksi Diabetes,” *Pros. CORISINDO 2023*, pp. 116–121, 2023, [Online]. Available: <https://stmikpontianak.org/ojs/index.php/corisindo/article/view/30%0Ahttps://stmikpontianak.org/ojs/index.php/corisindo/article/download/30/22>
- [13] M. Rizki and M. Mulyawan, “Penerapan Metode K-Means Clustering Pada Data Penjualan Optik Chantika,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1303–1307, 2023, doi: 10.36040/jati.v7i2.6562.
- [14] F. A. H. Ş. A. M. F. A. Y. Adıgüzel, A. M. Priyatno, W. F. R. Sudirman, and R. J. Musridho, “A New Hybrid Approach for Real Estate Price Prediction Using Outlier Detection, Feature Selection, and Clustering Techniques,” *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, pp. 1906–1915, 2023, doi: <https://doi.org/10.1109/UBMK59864.2023.10286673>.
- [15] A. M. Priyatno, W. F. R. Sudirman, and R. J. Musridho, “Feature selection using non-parametric correlations and important features on recursive feature elimination for stock price prediction,” *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, pp. 1906–1915, 2024, doi: 10.11591/ijece.v14i2.pp1906-1915.
- [16] E. Setia Budi, A. Nofriyaldi Chan, P. Priscillia Alda, and M. Arif Fauzi Idris, “RESOLUSI: Rekayasa Teknik Informatika dan Informasi Optimasi Model Machine Learning untuk Klasifikasi dan Prediksi Citra Menggunakan Algoritma Convolutional Neural Network,” *Media Online*, vol. 4, no. 5, p. 509, 2024, [Online]. Available: <https://djournals.com/resolusi>
- [17] M. N. Faiz, O. Somantri, A. R. Supriyono, and A. W. Muhammad, “Impact of Feature Selection Methods on Machine Learning-based for Detecting DDoS Attacks : Literature Review,” *J. Informatics Telecommun. Eng.*, vol. 5, no. 2, pp. 305–314, 2022, doi: 10.31289/jite.v5i2.6112.
- [18] S. Budiman, A. Sunyoto, and A. Nasiri, “Analisa Performa Penggunaan Feature Selection untuk Mendeteksi Intrusion Detection Systems dengan Algoritma Random Forest Classifier,” *Sistemasi*, vol. 10, no. 3, p. 753, 2021, doi: 10.32520/stmsi.v10i3.1550.
- [19] H. Rahman, R. Agusman, and T. Sutabri, “MODEL PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN MACHINE LEARNING,” *Sport. Cult.*, vol. 15, no. 1, pp. 72–86, 2024, doi: 10.25130/sc.24.1.6.
- [20] P. Triya, N. Suarna, and N. Dienwati Nuris, “Penerapan Machine Learning Dalam Melakukan Prediksi Harga Saham Pt. Bank Mandiri (Persero) Tbk Dengan Algoritma Linear Regression,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 1207–1214, 2024,

doi: 10.36040/jati.v8i1.8958.

- [21] P. Studi *et al.*, “KLASIFIKASI DATA MAHASISWA LAMPAU MENGGUNAKAN METODE DECISION TREE DAN SUPPORT VECTOR MACHINE Oleh : KURNIAWATI,” 2024, [Online]. Available: <http://etheses.uin-malang.ac.id/72754/>
- [22] M. K. Suryadi, R. Herteno, S. W. Saputro, M. R. Faisal, and R. A. Nugroho, “A Comparative Study of Various Hyperparameter Tuning on Random Forest Classification with SMOTE and Feature Selection Using Genetic Algorithm in Software Defect Prediction,” *J. Electron. Electromed. Eng. Med. Informatics*, vol. 6, no. 2, pp. 137–147, 2024, doi: 10.35882/jeeemi.v6i2.375.
- [23] A. T. Nurani, A. Setiawan, and B. Susanto, “Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma,” *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, 2023, doi: 10.24246/juses.v6i1p34-43.
- [24] S. Radithya, H. Santoso, P. A. Bhirawa, D. Fibianto, G. Fauzan, and B. Rahmat, “Prediksi Target Penjualan Mobil Menggunakan Decision Tree,” vol. 4, pp. 14–18, 2024.