

PENERAPAN ALGORITMA RANDOM FOREST UNTUK PREDIKSI KETERSEDIAAN STOK DAN OPTIMALISASI PEMASARAN PRODUK PAKAIAN

Roy Untung Pratama Saing¹, Parasian D.P. Silitonga²

Universitas Katolik Santo Thomas, Indonesia

e-mail: ¹royuntung8@gmail.com, ²parasian.silitonga@ust.ac.id

Abstract: Inventory management is a major challenge in the clothing retail business. Toko Saing as a clothing retail business still performs manual stock recording, causing risks of overstock and stockout. This study aims to apply the Random Forest algorithm to predict stock availability and optimize marketing of clothing products. Historical sales data for 6 months (October 2024 - March 2025) was used as the training dataset. The prediction system was built web-based using PHP and MySQL. The test results show that the Random Forest algorithm is able to predict stock needs with an average accuracy of 77.3% (R^2) and Mean Absolute Error (MAE) of 0.6 pcs. The developed system has successfully helped Toko Saing in managing stock more efficiently and providing data-based marketing recommendations. This study proves that Random Forest is effectively applied to stock prediction problems in retail goods.

Keyword: Random Forest; stock prediction; data mining; machine learning; clothing retail.

Abstrak: Pengelolaan persediaan stok merupakan tantangan utama dalam usaha ritel pakaian. Toko Saing sebagai usaha ritel pakaian masih melakukan pencatatan stok secara manual, menyebabkan risiko kelebihan stok (overstock) dan kekurangan stok (stockout). Penelitian ini bertujuan menerapkan algoritma Random Forest untuk memprediksi ketersediaan stok dan mengoptimalkan pemasaran produk pakaian. Data historis penjualan selama 6 bulan (Oktober 2024 - Maret 2025) digunakan sebagai dataset pelatihan. Sistem prediksi dibangun berbasis web menggunakan PHP dan MySQL. Hasil pengujian menunjukkan algoritma Random Forest mampu memprediksi kebutuhan stok dengan akurasi rata-rata 77.3% (R^2) dan Mean Absolute Error (MAE) sebesar 0.6 pcs. Sistem yang dikembangkan berhasil membantu Toko Saing dalam mengelola stok secara lebih efisien dan memberikan rekomendasi pemasaran berbasis data. Penelitian ini membuktikan bahwa Random Forest efektif diterapkan pada permasalahan prediksi stok barang ritel.

Kata kunci: Random Forest; prediksi stok; data mining; machine learning; ritel pakaian.

PENDAHULUAN

Perkembangan teknologi informasi saat ini mendorong pelaku usaha untuk memanfaatkan sistem berbasis data dalam mendukung kegiatan operasional dan pengambilan keputusan. Salah satu aspek penting dalam kegiatan usaha ritel pakaian adalah pengelolaan persediaan barang (stok). Ketersediaan stok yang tidak terkontrol dapat menimbulkan berbagai permasalahan, seperti kelebihan stok (overstock) maupun kekurangan stok

(stockout). Kelebihan stok menyebabkan meningkatnya biaya penyimpanan dan risiko barang tidak terjual, sedangkan kekurangan stok dapat mengakibatkan hilangnya peluang penjualan serta menurunnya kepuasan pelanggan [1].

Toko Saing merupakan salah satu usaha penjualan produk pakaian yang masih melakukan pengelolaan stok secara manual. Pencatatan persediaan dilakukan berdasarkan pengalaman dan perkiraan pemilik toko tanpa analisis data historis penjualan secara sistematis. Kondisi ini

menyebabkan beberapa permasalahan, seperti penumpukan barang yang kurang diminati pelanggan serta kekosongan stok ketika permintaan meningkat. Selain itu, strategi promosi sering kali tidak mempertimbangkan ketersediaan stok sehingga kurang efektif dalam meningkatkan penjualan [2].

Data transaksi penjualan Toko Saing sebenarnya tersimpan setiap hari dan berpotensi dimanfaatkan sebagai sumber informasi yang bernilai. Melalui teknik data mining, data historis tersebut dapat dianalisis untuk menemukan pola permintaan produk. Han, Kamber, dan Pei (2012) menyatakan bahwa data mining mampu menggali pola tersembunyi dari data besar untuk mendukung pengambilan keputusan bisnis [3]. Seiring perkembangan teknologi, pendekatan machine learning dapat digunakan untuk melakukan prediksi secara otomatis berdasarkan data historis. Machine learning memungkinkan sistem belajar dari pola penjualan masa lalu untuk memperkirakan kebutuhan stok di masa mendatang [4].

Salah satu algoritma yang memiliki performa baik dalam prediksi adalah Random Forest. Breiman (2001) menjelaskan bahwa Random Forest mampu meningkatkan akurasi prediksi serta mengurangi risiko overfitting karena menggunakan banyak pohon keputusan secara ensemble [5]. Beberapa penelitian terdahulu menunjukkan bahwa Random Forest efektif dalam prediksi penjualan dan persediaan barang. Efendi dkk. (2024) melaporkan peningkatan akurasi prediksi hingga 85% pada prediksi penjualan produk UMKM [6]. Sari dkk. (2025) memperoleh nilai R^2 hingga 98-99% pada prediksi kebutuhan stok ritel modern [7].

Berdasarkan permasalahan tersebut, diperlukan suatu sistem prediksi berbasis machine learning untuk membantu Toko

Saing dalam memperkirakan ketersediaan stok sekaligus mendukung optimalisasi pemasaran produk. Penelitian ini bertujuan menerapkan algoritma Random Forest untuk memprediksi ketersediaan stok produk pakaian berdasarkan data historis penjualan pada Toko Saing, serta merancang dan membangun sistem berbasis web yang dapat membantu proses prediksi kebutuhan stok produk pakaian menggunakan algoritma Random Forest.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan algoritma Random Forest untuk memprediksi ketersediaan stok produk pakaian di Toko Saing. Subjek penelitian terdiri dari 5 produk pakaian yang menjadi objek penjualan. Teknik pengambilan sampel yang digunakan adalah purposive sampling, di mana produk yang dipilih merupakan produk yang paling sering muncul pada catatan penjualan Toko Saing. Pengumpulan data dilakukan melalui observasi, wawancara, dan dokumentasi untuk memperoleh informasi mengenai proses penjualan, data produk, serta data penjualan historis.

Data yang digunakan merupakan data historis penjualan produk pakaian pada Toko Saing selama periode 6 bulan, yaitu Oktober 2024 sampai Maret 2025. Data diperoleh dari catatan transaksi penjualan harian yang ditulis secara manual di buku penjualan toko. Karena sebagian pencatatan hanya menampilkan total harga tanpa jumlah pcs, maka dilakukan proses konversi untuk memperkirakan jumlah barang terjual berdasarkan kisaran harga produk yang umum dijual di toko. Hasil konversi tersebut kemudian direkap menjadi data penjualan bulanan.

Tabel 1 Data Produk

id produk	Nama Produk	Kategori	Harga Rata-rata
1	Kaos	Pakaian	Rp 50.000 - 80.000
2	Celana	Pakaian	Rp 70.000 - 150.000
3	Jaket	Pakaian	Rp 120.000 - 250.000

4	Baju Wanita	Pakaian	Rp 80.000 - 200.000
5	Produk Lainnya	Pakaian	Bervariasi

Tabel 2 Data Penjualan Bulanan

Produk	Okt 2024	Nov 2024	Des 2024	Jan 2025	Feb 2025	Mar 2025
Kaos	4	5	9	8	5	6
Celana	3	4	6	5	4	3
Jaket	4	6	8	5	5	4
Baju Wanita	2	3	5	4	2	1
Produk Lainnya	5	8	3	6	3	3

Sebelum dataset digunakan dalam proses pembentukan model Random Forest, data terlebih dahulu dilakukan proses transformasi. Transformasi dataset bertujuan untuk mengubah bentuk data penjualan menjadi format supervised learning. Dalam format ini, beberapa periode sebelumnya digunakan sebagai variabel input (X), sedangkan periode berikutnya digunakan sebagai variabel target (Y) yang akan diprediksi.

Random Forest merupakan salah satu algoritma dalam machine learning yang termasuk dalam kategori ensemble learning. Secara matematis, untuk masalah regresi, Random Forest menggunakan rata-rata (averaging) dari semua pohon:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T f_t(x)$$

Di mana $\hat{y}$ adalah prediksi akhir, T adalah jumlah pohon keputusan, $f_t(x)$ adalah prediksi dari pohon ke- t , dan x adalah data input.

Random Forest bekerja melalui empat tahapan utama: Bootstrap Sampling

(pengambilan sampel acak dengan pengembalian), Random Feature Selection (pemilihan fitur secara acak pada setiap node), pembangunan Decision Tree, dan agregasi hasil prediksi.

Evaluasi model dilakukan untuk mengukur seberapa baik model dalam melakukan prediksi. Metrik evaluasi yang digunakan adalah Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), dan Koefisien Determinasi (R^2) [8].

HASIL DAN PEMBAHASAN

Penelitian ini menggunakan 5 produk pakaian sebagai alternatif yang dinilai berdasarkan data penjualan historis selama 6 bulan. Setiap produk memiliki data penjualan per bulan yang digunakan sebagai variabel input dalam proses prediksi. Nilai penjualan tersebut menjadi dasar dalam pembentukan dataset sebelum dilakukan proses perhitungan menggunakan metode Random Forest.

Tabel 3 Perbandingan Nilai Aktual dan Prediksi

Produk	Nilai Aktual	Nilai Prediksi	Error	Produk	Nilai Aktual	Nilai Prediksi
Kaos	6	6	0	Kaos	6	6
Celana	3	4	1	Celana	3	4
Jaket	4	5	1	Jaket	4	5
Baju Wanita	1	2	1	Baju Wanita	1	2
Produk Lainnya	3	3	0	Produk Lainnya	3	3

Berdasarkan Tabel 3, setiap produk memiliki nilai aktual dan prediksi yang berbeda. Error terbesar terjadi pada produk Celana, Jaket, dan Baju Wanita dengan

selisih 1 pcs, sedangkan Kaos dan Produk Lainnya memiliki error 0.

Perhitungan evaluasi model dilakukan terhadap seluruh produk yang

telah diprediksi. Hasil perhitungan evaluasi model adalah sebagai berikut:

Tabel 4 Hasil Evaluasi Model

Metrik	Nilai	Interpretasi
MAE	0.6	Rata-rata kesalahan prediksi = 0.6 pcs
RMSE	0.775	Kesalahan kuadrat rata-rata = 0.775
MAPE	31.67%	Persentase kesalahan rata-rata = 31.67% (cukup)
R ²	0.773	Model menjelaskan 77.3% variasi data

Perhitungan MAE:

$$MAE = \frac{0 + 1 + 1 + 1 + 0}{5} = \frac{3}{5} = 0.6$$

Perhitungan RMSE:

$$RMSE = \sqrt{\frac{0 + 1 + 1 + 1 + 0}{5}} = \sqrt{\frac{3}{5}} = \sqrt{0.6} = 0.775$$

Perhitungan MAPE:

$$MAPE = \frac{100\%}{5} \times (0\% + 33.33\% + 25\% + 100\% + 0\%) = 20\% \times 158.33\% = 31.67\%$$

Perhitungan R²:

$$\bar{Y} = \frac{6 + 3 + 4 + 1 + 3}{5} = \frac{17}{5} = 3.4$$

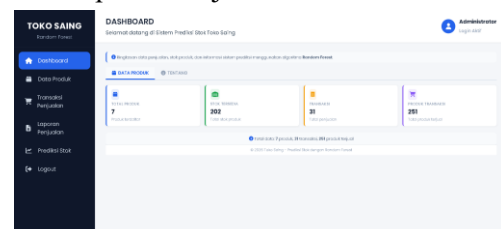
$$SS_{total} = (6 - 3.4)^2 + (3 - 3.4)^2 + (4 - 3.4)^2 + (1 - 3.4)^2 + (3 - 3.4)^2 = 13.2$$

$$SS_{res} = 0 + 1 + 1 + 1 + 0 = 3$$

$$R^2 = 1 - \frac{3}{13.2} = 1 - 0.227 = 0.773$$

Berdasarkan hasil evaluasi tersebut, model Random Forest dapat dikatakan cukup baik dalam melakukan prediksi penjualan produk. Nilai MAE sebesar 0.6 menunjukkan bahwa rata-rata kesalahan prediksi hanya sekitar 0.6 pcs per produk, sedangkan nilai R² sebesar 0.773 menunjukkan bahwa model mampu menjelaskan 77.3% variasi data penjualan.

Sistem prediksi ketersediaan stok dibangun berbasis web menggunakan bahasa pemrograman PHP dan database MySQL. PHP digunakan karena mampu mengolah data secara dinamis, sedangkan MySQL berfungsi untuk menyimpan data produk, penjualan, dan hasil prediksi. Sistem yang dibangun memiliki fitur utama yaitu Halaman Login dan Logout, Dashboard, Kelola Data Produk, Kelola Transaksi Penjualan, Prediksi Stok menggunakan algoritma Random Forest, dan Laporan Penjualan.



Gambar 1 Halaman Dashboard



Gambar 2 Halaman Prediksi Stok

Pengujian sistem dilakukan menggunakan metode Black Box Testing untuk mengetahui apakah fungsi utama sistem dapat berjalan sesuai dengan kebutuhan yang telah dirancang. Pengujian mencakup proses login, pengelolaan data produk, penjualan, perhitungan prediksi, hingga laporan penjualan. Seluruh fungsi yang diuji dapat dijalankan sesuai dengan hasil yang diharapkan. Pengujian tersebut menunjukkan bahwa sistem telah dapat menjalankan fungsi utama dalam proses pengelolaan data, penjualan, prediksi, hingga laporan penjualan.

SIMPULAN

Penerapan algoritma Random Forest berhasil digunakan untuk memprediksi ketersediaan stok produk pakaian berdasarkan data historis

penjualan pada Toko Saing. Berdasarkan pengolahan terhadap 5 produk menggunakan data penjualan 6 bulan, sistem menghasilkan prediksi dengan nilai MAE sebesar 0.6 yang menunjukkan rata-rata kesalahan prediksi hanya sekitar 0.6 pcs per produk, dan nilai R^2 sebesar 0.773 yang menunjukkan bahwa model mampu menjelaskan 77.3% variasi data penjualan.

Implementasi algoritma Random Forest ke dalam sistem berbasis web menggunakan PHP dan MySQL memungkinkan proses prediksi dilakukan secara otomatis. Sistem yang dikembangkan memiliki fitur-fitur yang memudahkan Toko Saing dalam mengelola data produk, mencatat transaksi penjualan, melakukan prediksi stok, serta menghasilkan laporan penjualan. Hasil prediksi yang dihasilkan oleh sistem dapat digunakan sebagai bahan pertimbangan bagi pemilik Toko Saing dalam mengambil keputusan terkait pengelolaan stok dan pemasaran produk.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Toko Saing yang telah memberikan izin dan data untuk penelitian ini, serta kepada dosen pembimbing yang telah memberikan arahan dan bimbingan selama proses penelitian berlangsung.

DAFTAR PUSTAKA

- J. Heizer, B. Render, and C. Munson, *Operations Management: Sustainability and Supply Chain Management*, 12th ed. Pearson, 2017.
- A. Sutrisno and B. Nugroho, "Prediksi Stok Barang Menggunakan Algoritma Random Forest pada Sistem Ritel," *Jurnal Ilmu Komputer dan Teknologi*, vol. 11, no. 3, pp. 150-160, 2020.
- J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2012.
- S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica*, vol. 31, no. 3, pp. 249-268, 2007.
- L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
- A. Efendi, B. Cahyono, and S. Sulisty, "Penerapan Algoritma Random Forest untuk Prediksi Penjualan Produk UMKM," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 12, no. 1, pp. 45-55, 2024.
- R. P. Sari, T. Hidayat, and W. Nugroho, "Implementasi Random Forest untuk Prediksi Stok Barang pada Sistem Ritel Modern," *Jurnal Sistem Informasi*, vol. 18, no. 1, pp. 20-30, 2025.
- G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, 2nd ed. Springer, 2021.