

PREDIKSI PENJUALAN BERAS HARIAN BERDASARKAN MEREK, HARGA, DAN UKURAN KEMASAN MENGGUNAKAN METODE MACHINE LEARNING

Oleh :

Sherly Maziya Rizqi (221080200122)

Dosen Pembimbing :

Yulian Findawati, ST., M.MT.

Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

2025



Latar Belakang

Penjualan beras merupakan aktivitas utama pada toko beras yang jumlahnya berubah setiap hari. Perubahan tersebut dipengaruhi oleh beberapa faktor, seperti merek, ukuran kemasan, harga, serta waktu penjualan. Kondisi tersebut sering menyulitkan pemilik toko dalam memperkirakan jumlah penjualan pada hari berikutnya sehingga pengelolaan stok menjadi kurang optimal. Oleh karena itu, diperlukan suatu metode yang mampu melakukan prediksi berdasarkan data historis penjualan. Pada penelitian ini digunakan metode Linear Regression dan Random Forest untuk memprediksi jumlah penjualan beras harian sehingga dapat membantu pengambilan keputusan dalam pengelolaan persediaan.

Tujuan Penelitian

- Menerapkan Linear Regression dan Random Forest untuk memprediksi jumlah penjualan beras harian
- Membandingkan tingkat kesalahan pada skenario 7, 5, dan 3 atribut serta pembagian data 80:20, 70:30, 60:40.
- Menganalisis hasil prediksi terpilih berdasarkan merek, periode penjualan, serta kombinasi bulan-merek-ukuran

Data Penelitian

Periode dan Jumlah	Target	7 kandidat atribut
• 02 Agustus 2025 – 31 Mei 2026 • Data awal : 4.420 baris, setelah preprocessing 4.301 baris	• Jumlah Terjual (PCS)	• Merek, Ukuran, Harga/kg, harga/PCS, Hari, Weekday/Weekend, hari Besar

Penelitian Terdahulu

Penelitian A : Prediksi Kebutuhan Stok Barang Menggunakan Algoritma Random Forest (*Sari, Arifin, & Darmanto, 2025*)

RESEARCH GAP: Belum terdapat penelitian yang membandingkan metode Random Forest dan Linear Regression untuk memprediksi penjualan beras harian berdasarkan merek, harga, dan ukuran kemasan.

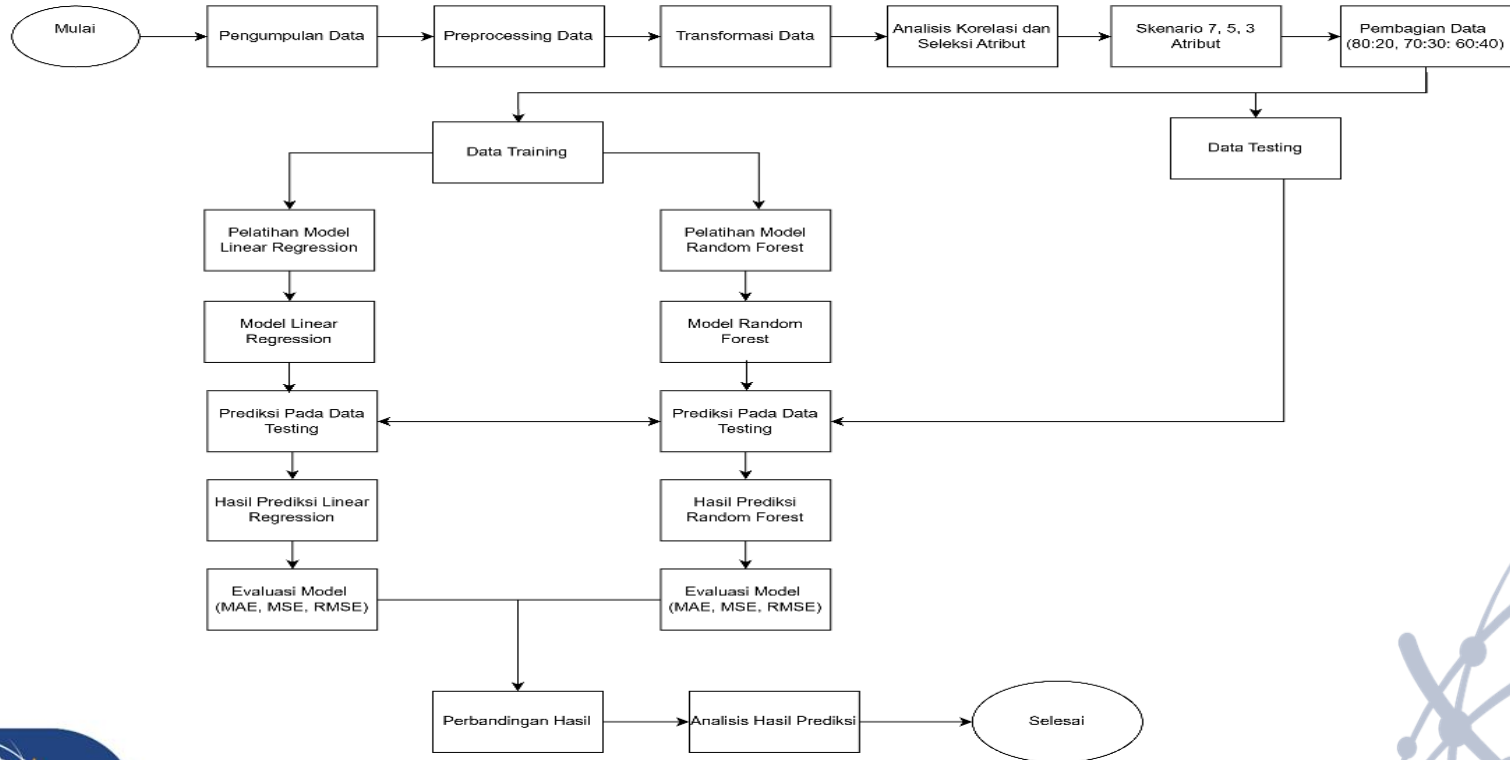
Penelitian C : Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear (*Anggrawan, Hairani, & Azmi, 2022*)

Penelitian B : Prediksi total Penjualan Gorden Menggunakan Random Forest Regressor (*Aszava et al, 2026*)

Penelitian D : Data mining untuk Prediksi Penjualan Menggunakan Simple Linear Regression (*Duran et al., 2024*)

Penelitian ini dilakukan untuk membandingkan performa Random forest dan Linear regression dalam memprediksi penjualan beras harian berdasarkan data historis penjualan.

Tahapan Penelitian



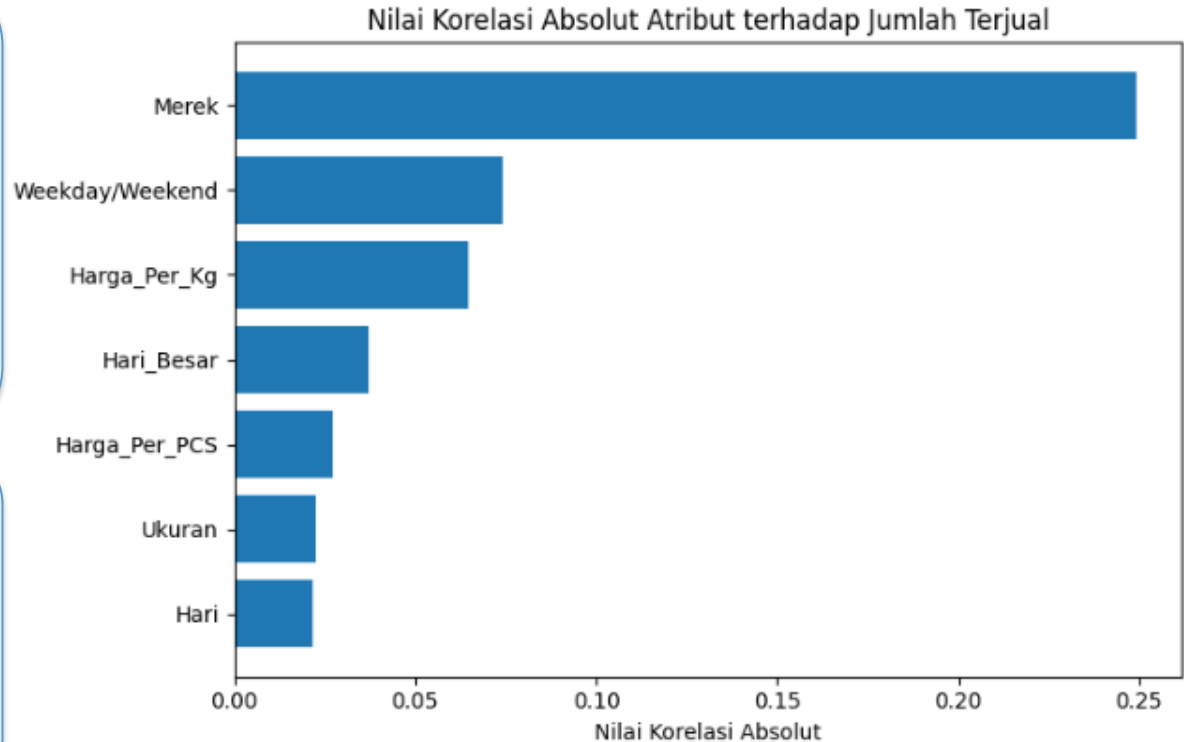
Transformasi dan Seleksi Atribut

Transformasi Data

Atribut kategorikal Merek, Hari, Weekday/Weekend, dan Hari Besar diubah menjadi numerik menggunakan Label Encoding.

Seleksi Atribut

Nilai korelasi absolut digunakan untuk membentuk skenario 7 atribut, 5 atribut, dan 3 atribut.



Skenario pengujian

Skenario Atribut

7 atribut
5 atribut
3 atribut

Pembagian Data

80:20
70:30
60:40

Metode

Linear Regression
Random Forest
Regression

3 skenario atribut × 3 pembagian data × 2 metode = 18 konfigurasi pengujian

Evaluasi menggunakan MAE, MSE, dan RMSE. Semakin kecil nilainya, semakin rendah tingkat kesalahan prediksi.

Skenario pengujian

Skenario Atribut

7 atribut
5 atribut
3 atribut

Pembagian Data

80:20
70:30
60:40

Metode

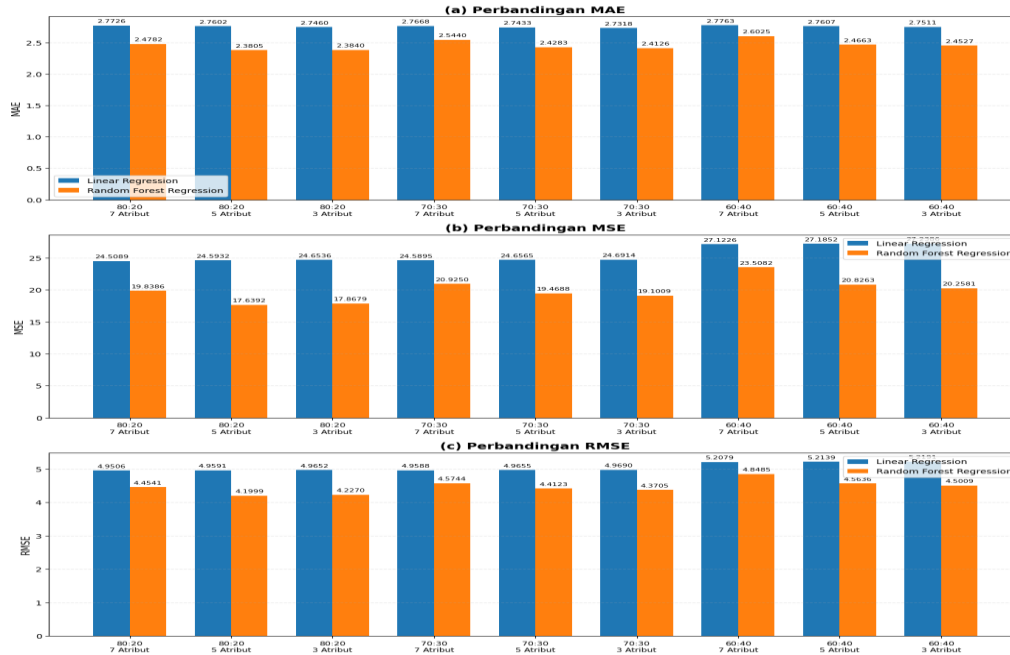
Linear Regression
Random Forest
Regression

3 skenario atribut × 3 pembagian data × 2 metode = 18 konfigurasi pengujian

Evaluasi menggunakan MAE, MSE, dan RMSE. Semakin kecil nilainya, semakin rendah tingkat kesalahan prediksi.

Hasil Pengujian 18 Konfigurasi

Perbandingan Hasil Evaluasi Linear Regression dan Random Forest Regression



Diketahui dari hasil disamping bahwa Metode Random Forest Rgression menghasilkan nilai kesalahan lebih rendah daripada Linear Regression pada seluruh kombinasi yang sama

Konfigurasi dengan tingkat kesalahan terendah

Random Forest Regression • 5 Atribut • Split 80:20

MAE

2,3805
≈ rata-rata selisih
absolut 2,38 pcs

MSE

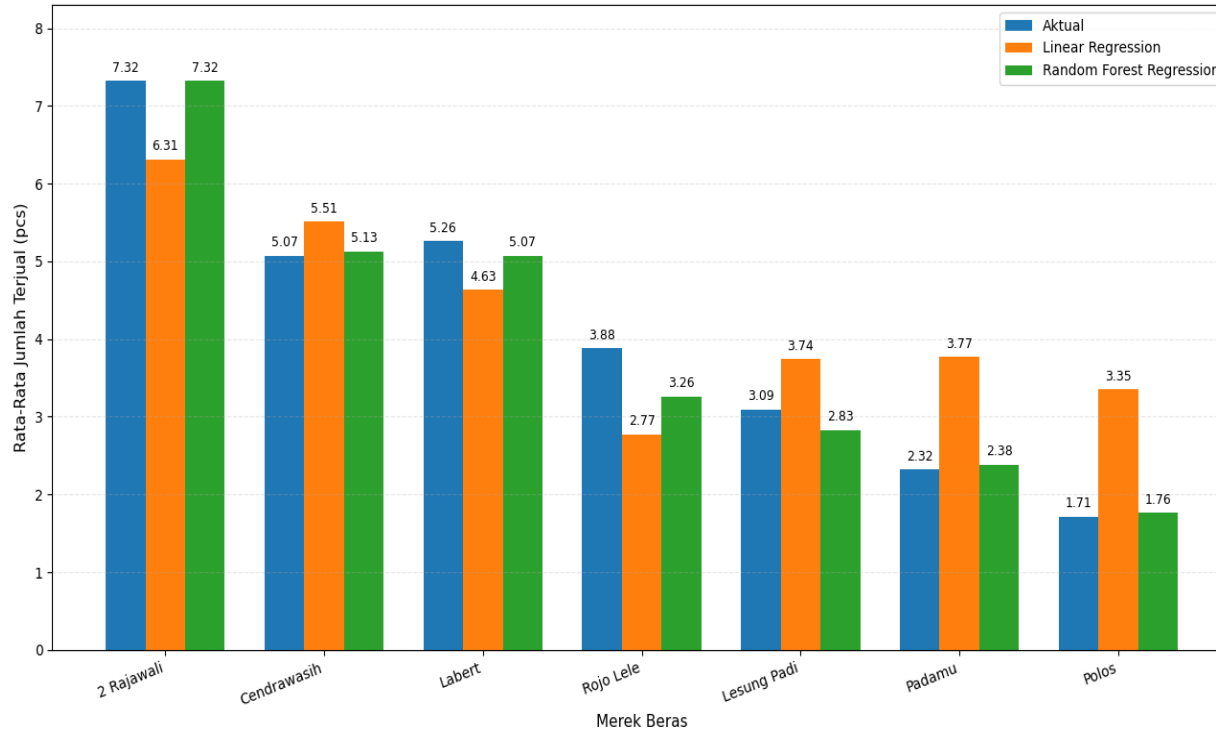
17,6392

RMSE

4,1999

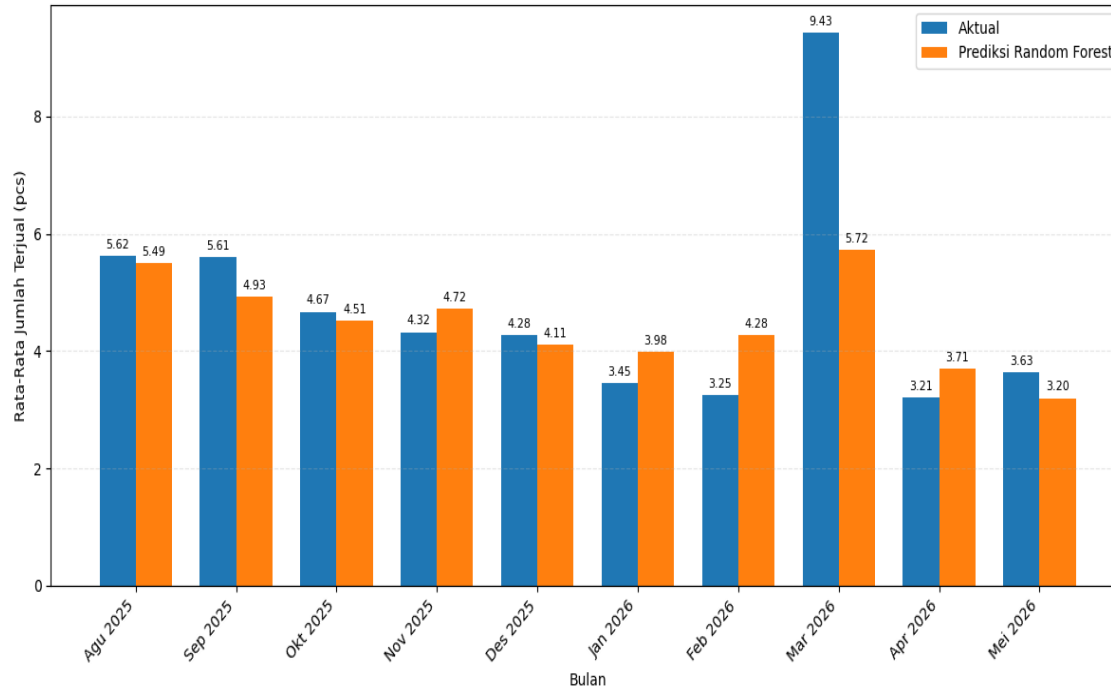
5 atribut: Merek • Weekday/Weekend • Harga Per Kg • Hari Besar • Harga Per PCS

Perbandingan Output Linear Regression dan Random Forest



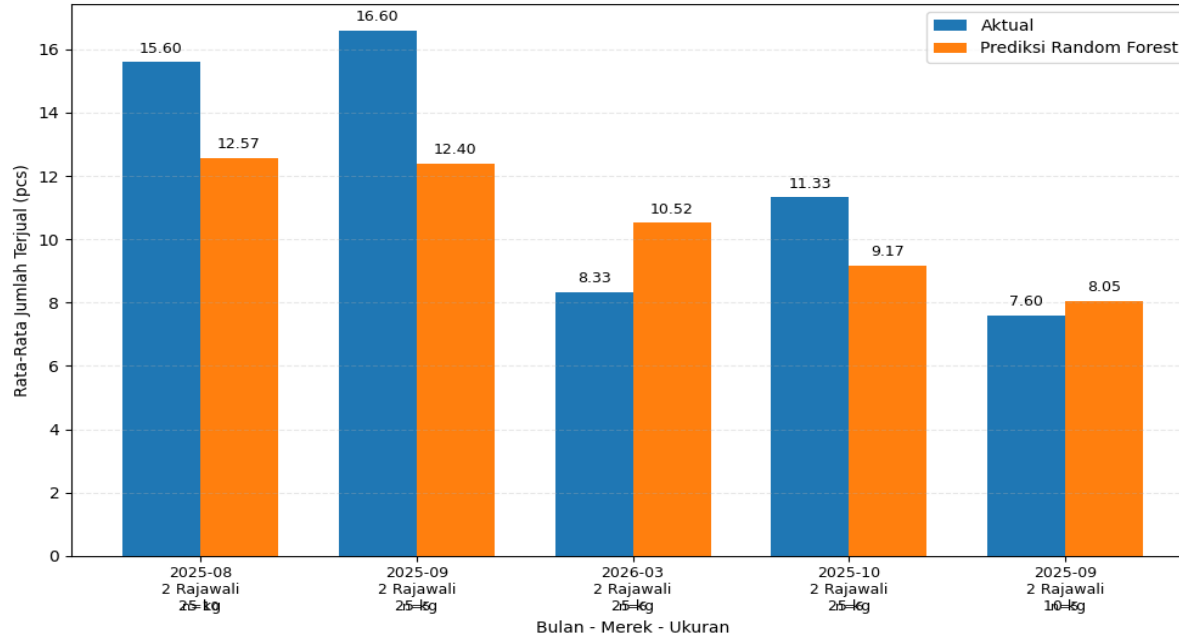
Pada hasil disamping menunjukkan bahwa kedua metode memberi rata-rata prediksi tertinggi pada merek 2 Rajawali. Dengan LR + 6,31 PCS, RF + 7,32 PCS, dengan Aktual = 7,32 PCS.

Analisis hasil Prediksi Berdasarkan Bulan



Pada gambar grafik disamping menunjukkan bahwa rata- rata prediksi tertinggi terdapat pada bulan Maret 026 sebesar 5,72 PCS, sedangkan rata – rata aktualnya 9,43 PCS.

Analisis Kombinasi Bulan, Merek, dan Ukuran



Pada gambar disaming menunjukkan bahwa Di antara kelompok dengan ≥ 3 observasi pada data testing, 2 Rajawali 25 kg pada Agustus 2025 memiliki rata-rata prediksi tertinggi 12,57 pcs (aktual 15,60 pcs; $n=10$).

Kesimpulan

- Dari 18 konfigurasi, Random Forest Regression dengan 5 atribut dan split 80:20 menghasilkan tingkat kesalahan terendah: MAE 2,3805; MSE 17,6392; RMSE 4,1999.
- Berdasarkan merek, rata-rata prediksi Random Forest tertinggi terdapat pada 2 Rajawali sebesar 7,32 pcs.
- Berdasarkan periode, rata-rata prediksi tertinggi terdapat pada Maret 2026 sebesar 5,72 pcs.
- Analisis kombinasi menunjukkan 2 Rajawali 25 kg pada Agustus 2025 memiliki rata-rata prediksi 12,57 pcs pada kelompok yang memiliki sedikitnya tiga observasi.
- Untuk prediksi periode masa depan, penelitian selanjutnya dapat mempertimbangkan pendekatan time series dan penambahan variabel relevan.

Referensi

- N. Fatimah and Y. Elmasari, “Perancangan Sistem Informasi Perpustakaan Berbasis Web Untuk Sma Islam Sunan Gunung Jati,” *JIPi (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 3, no. 2, pp. 130–137, 2018, doi: 10.29100/jipi.v3i2.783.
- G. Krisnawati, Isharijadi, and E. Astuti, “Analysis of the Implementation of the Affari Retail System Inventory Accounting Information System in Sty Store,” *J. Sustain.*, vol. 03, no. 01, pp. 75–90, 2023, [Online]. Available: <http://journal.um-surabaya.ac.id/index.php/sustainable/index>
- P. A. Duran, A. V. Vitianingsih, M. S. Riza, A. L. Maukar, and S. F. A. Wati, “Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression,” *Teknika*, vol. 13, no. 1, pp. 27–34, 2024, doi: 10.34148/teknika.v13i1.712.
- B. A. Nugroho, A. K. A. Pradana, and E. Nurfarida, “Prediksi Waktu Kedatangan Pelanggan Servis Kendaraan Bermotor Berdasarkan Data Historis menggunakan Support Vector Machine,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 25, 2021, doi: 10.26418/jp.v7i1.42964.
- S. R. Febrian *et al.*, “PREDIKSI PENJUALAN SUKU CADANG MOTOR DENGAN PENERAPAN RANDOM FOREST DI PT TERUS JAYA SENTOSA MOTOR,” vol. 8, no. 5, pp. 10507–10513, 2024.

Referensi

N. Fatimah and Y. Elmasari, “Perancangan Sistem Informasi Perpustakaan Berbasis Web Untuk Sma Islam Sunan Gunung Jati,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 3, no. 2, pp. 130–137, 2018, doi: 10.29100/jipi.v3i2.783.

G. Krisnawati, Isharijadi, and E. Astuti, “Analysis of the Implementation of the Affari Retail System Inventory Accounting Information System in Sty Store,” *J. Sustain.*, vol. 03, no. 01, pp. 75–90, 2023, [Online]. Available: <http://journal.um-surabaya.ac.id/index.php/sustainable/index>

P. A. Duran, A. V. Vitianingsih, M. S. Riza, A. L. Maukar, and S. F. A. Wati, “Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression,” *Teknika*, vol. 13, no. 1, pp. 27–34, 2024, doi: 10.34148/teknika.v13i1.712.

B. A. Nugroho, A. K. A. Pradana, and E. Nurfarida, “Prediksi Waktu Kedatangan Pelanggan Servis Kendaraan Bermotor Berdasarkan Data Historis menggunakan Support Vector Machine,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 25, 2021, doi: 10.26418/jp.v7i1.42964.

S. R. Febrian *et al.*, “PREDIKSI PENJUALAN SUKU CADANG MOTOR DENGAN PENERAPAN RANDOM FOREST DI PT TERUS JAYA SENTOSA MOTOR,” vol. 8, no. 5, pp. 10507–10513, 2024.

Referensi

- G. N. Ayuni and D. Fitriana, “Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ,” *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2020, doi: 10.61769/telematika.v14i2.321.
- T. Indarwati, T. Irawati, and E. Rimawati, “Using the Linear Regression Method to Predict Smartphone Sales,” *J. Inf. Commun. Technol.*, vol. 6, no. 2, pp. 2–7, 2019.
- A. Sari, M. Arifin, and E. Darmanto, “Prediksi Kebutuhan Stok Barang Menggunakan Algoritma Random Forest Untuk Meningkatkan Efisiensi Penjualan,” *J. Inf. Syst. Informatics Eng.*, vol. 9, no. 2, pp. 339–351, 2025.
- A. Aszava *et al.*, “Prediksi Total Penjualan Gorden Berdasarkan Data Historis Bulanan Menggunakan Random Forest Regressor,” vol. 5, no. 1, pp. 786–792, 2026, doi: 10.70247/jumistik.v5i1.292.
- Suwarno, “JITE (Journal of Informatics and Telecommunication Engineering) (Case Study : Ministry of Religious Affairs of Batam),” vol. 4, no. January, 2021.
- A. Manurung and A. Rahman, “Implementasi Algoritma Random Forest untuk Memprediksi Penjualan Produk Pada Toko Ritel Elektronik di Kota Palembang,” no. 2026, pp. 1337–1344.
- M. Hardika, Martono, Arif Rinaldy Dikananda, “Meningkatkan Model Prediksi Penjualan Pada Toko Devanjayaban,” *JITET (Jurnal Inform. dan Tek. Elektro Ter.*, vol. 13, no. 2, pp. 913–921, 2025.

Sekian,
Terimakasih

