

Estimasi Penjualan Sembako Menggunakan Teknik Data Mining Regresi Linear Berganda

Alya Rahmadani*, Lili Tanti

Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Potensi Utama, Medan, Indonesia

Email: ^{1,*}alyarahmadani76@gmail.com, ²lili@potensi-utama.ac.id,

Email Penulis Korespondensi: alyarahmadani76@gmail.com*

Submitted: 22/10/2025; Accepted: 27/11/2025; Published: 31/12/2025

Abstrak– Penelitian ini bertujuan untuk membangun model prediksi penjualan sembako menggunakan teknik data mining dengan metode regresi linear berganda pada CV. Surya Kencana Sembako. Model ini dirancang untuk membantu perusahaan dalam memperkirakan jumlah penjualan secara lebih akurat berdasarkan data historis. Dataset akhir yang digunakan dalam penelitian ini terdiri dari 867 data penjualan setelah melalui proses pembersihan data, seleksi fitur, dan penanganan outlier. Pengujian dilakukan terhadap dua kelompok data, yaitu 30 data baru dan 174 data testing. Hasil pengujian menunjukkan bahwa model memiliki kinerja prediksi yang sangat baik, dengan nilai Mean Squared Error (MSE) sebesar 628.18 dan 520.7, Root Mean Squared Error (RMSE) sebesar 25.06 dan 22.82, serta Mean Absolute Error (MAE) sebesar 13.91 dan 10.21 pada masing-masing pengujian. Nilai koefisien determinasi (R^2) yang diperoleh sebesar 0.92 pada data baru dan 0.97 pada data testing menunjukkan bahwa model dapat menjelaskan variasi data penjualan dengan sangat baik. Kontribusi utama penelitian ini adalah menghasilkan model estimasi penjualan berbasis regresi linear berganda yang terintegrasi ke dalam sistem informasi penjualan berbasis web. Implikasi dari penelitian ini adalah perusahaan dapat melakukan perencanaan stok dan pengambilan keputusan operasional secara lebih efisien dan tepat sasaran berdasarkan prediksi penjualan yang akurat.

Kata Kunci: Prediksi Penjualan; Data Mining; Regresi Linear Berganda; Sembako; Sistem Informasi Berbasis Web

Abstract– This study aims to develop a sales prediction model for basic necessities (sembako) using data mining techniques with the multiple linear regression method at CV. Surya Kencana Sembako. The model is designed to assist the company in accurately forecasting sales figures based on historical data. The final dataset used in this research consists of 867 cleaned and preprocessed sales records after handling missing values, outliers, and feature selection. Testing was conducted on two data groups, namely 30 new data entries and 174 testing data entries. The results show that the model has excellent predictive performance, with a Mean Squared Error (MSE) of 628.18 and 520.7, a Root Mean Squared Error (RMSE) of 25.06 and 22.82, and a Mean Absolute Error (MAE) of 13.91 and 10.21, respectively. The coefficient of determination (R^2) values of 0.92 for the new data and 0.97 for the testing data indicate that the model can explain the variance in sales data very well. The main contribution of this research is the development of a multiple linear regression-based sales estimation model that can be integrated into a web-based sales information system. The implications of this study show that the company can optimize stock planning and operational decision-making more effectively through the utilization of accurate sales predictions.

Keywords: Sales Prediction; Data Mining; Multiple Linear Regression; Basic Necessities; Web-Based Information System.

1. PENDAHULUAN

Sistem informasi merupakan sekumpulan komponen yang saling berinteraksi untuk menghasilkan informasi yang dibutuhkan manajemen dalam proses pengambilan keputusan [1]. Dalam konteks bisnis sembako, keberadaan informasi penjualan yang akurat sangat penting karena sembako termasuk kebutuhan dasar masyarakat yang harus tersedia secara stabil dan berkesinambungan. Pengelolaan stok yang tidak tepat dapat menyebabkan kekurangan barang, kelebihan stok, hingga kerugian perusahaan [2].

Penentuan jumlah stok adalah hal yang penting dalam suatu proses bisnis usaha yang merupakan pemenuhan CV. Surya Kencana Sembako merupakan perusahaan yang bergerak di bidang distribusi kebutuhan sehari-hari seperti beras, minyak goreng, snack, dan produk pokok lainnya. Aktivitas penjualan dilakukan setiap hari sehingga fluktuasi permintaan secara langsung memengaruhi ketersediaan stok di gudang [3]. Namun, pengelolaan data penjualan pada perusahaan ini masih dilakukan secara manual menggunakan buku catatan, sehingga proses rekapitulasi memakan waktu lama, rawan kehilangan data, dan menyulitkan perusahaan dalam memantau permintaan serta menentukan jumlah persediaan optimal setiap bulan. Kondisi ini menyebabkan sering terjadinya ketidaksesuaian stok dan ketidakmampuan perusahaan memenuhi permintaan pelanggan secara maksimal. Pengaruh ketersediaan produk sembako merupakan landasan penting khususnya untuk jenis sembako yang diminati dan paling banyak terjual. Selama ini pengelolaan data penjualan dan stok produk masih menggunakan sistem secara manual dengan menggunakan rekapan buku dan proses perekapan penjualan dan stok produk dilakukan setiap diakhir bulan. Jadi, hal tersebut memerlukan banyak buku untuk pencatatan penjualan dan stok produk, juga sering mengakibatkan lamanya perekapan data penjualan yang hilang.

Regresi linear berganda adalah model regresi yang melibatkan lebih dari satu variabel independen. Analisis *regresi linear* berganda dilakukan untuk mengetahui arah dan seberapa besar pengaruh variabel independen terhadap variabel dependen [4]. Penelitian yang dilakukan [5] menyimpulkan bahwa kasus yang di bahas adalah *forecast the stock price fluctuations of top-rated companie*, dengan Variabel Independen: *Open*, *High*, *Low*, *Volume* dan Variabel Dependen : *Close*, jumlah data set 98.075, *label class* adalah *forecast the stock price fluctuations of top-rated companie*, algoritma yang digunakan adalah *Linear Regression*, dengan Teknik evaluasi RMSE sebesar 53,5. Penelitian yang dilakukan [6] menyimpulkan bahwa kasus yang di bahas adalah *used car sales in Indonesia*, dengan Variabel Independen: Usia masa pakai, Jarak dan Variabel Dependen : Harga, jumlah

data set 240, *label class* adalah *used car sales in Indonesia*, algoritma yang digunakan adalah Linear Regression, dengan Teknik evaluasi RS-Q sebesar 63.3%. Penelitian yang dilakukan [7] menyimpulkan bahwa kasus yang di bahas adalah *Forecasting Techniques for Sales Budget Determination*, dengan Variabel Independen: *Sales* (unit) Value dan Variabel Dependen : *trend value*, jumlah data set 24.120, *label class* adalah *Forecasting Techniques for Sales Budget Determination*, algoritma yang digunakan adalah *Semi-Average Trend Method*, dengan Teknik evaluasi SKF sebesar 4.1%.

Penelitian yang dilakukan [8] menyimpulkan bahwa kasus yang di bahas adalah Prediksi Penjualan Buku, dengan Variabel Independen: Kategori, Toko, Tahun Terjual, jumlah dan Variabel Dependen : status terjual, jumlah data set 112, *label class* adalah Laris/ Tidak Laris, algoritma yang digunakan adalah C4.5, dengan Teknik evaluasi Precision sebesar 100%. Penelitian yang dilakukan [9] menyimpulkan bahwa kasus yang di bahas adalah Estimasi Stok Barang, dengan Variabel Independen: Jenis Barang, Satuan, Stok Awal, Stok Terjual dan Variabel Dependen : stok akhir, jumlah data set 503, *label class* adalah Laris/ Tidak Laris, algoritma yang digunakan adalah C4.5, dengan Teknik evaluasi Precision sebesar 100

GAP penelitian terletak pada belum adanya sistem prediksi penjualan yang mampu memanfaatkan data historis secara otomatis untuk menghasilkan estimasi penjualan yang akurat. Meskipun perusahaan memiliki banyak data penjualan, data tersebut belum diolah menggunakan metode analitik atau data mining untuk mendukung perencanaan stok. Hal ini menunjukkan adanya kebutuhan solusi yang lebih kontekstual dan spesifik pada industri sembako, yang memiliki pola penjualan yang lebih fluktuatif dan dipengaruhi banyak variabel simultan.

Berdasarkan permasalahan tersebut, tujuan penelitian ini Adalah Mengembangkan model estimasi penjualan sembako menggunakan regresi linear berganda berdasarkan data historis penjualan. Mengidentifikasi variabel independen yang berpengaruh terhadap jumlah penjualan. Mengevaluasi performa model prediksi menggunakan metrik MSE, RMSE, MAE, dan R^2 [10]. Mengimplementasikan model prediksi ke dalam sistem informasi berbasis web untuk membantu proses operasional perusahaan.

Manfaat penelitian ini Adalah Membantu perusahaan mengoptimalkan perencanaan stok berdasarkan hasil prediksi yang akurat. Meminimalkan risiko kekurangan atau kelebihan stok. Meningkatkan efisiensi operasional melalui otomatisasi analisis data penjualan. Mendukung pengambilan keputusan strategis dalam pengadaan barang, promosi, dan pengelolaan gudang. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam penerapan data mining untuk meningkatkan efektivitas manajemen penjualan sembako pada CV. Surya Kencana Sembako.

2. METODOLOGI PENELITIAN

2.1 Rancangan Penelitian

Metodologi penelitian ini disusun untuk memberikan gambaran yang jelas mengenai tahapan-tahapan dalam membangun model estimasi penjualan sembako menggunakan teknik data mining dengan metode regresi linear berganda. Pendekatan yang digunakan mengacu pada CRISP-DM (Cross Industry Standard Process for Data Mining), karena model ini memberikan alur kerja yang terstruktur dan telah banyak digunakan dalam penelitian maupun industri. Flowchart dari prosedur penelitian dapat ditambahkan oleh peneliti pada bagian ilustrasi alur.

1. Business Understanding (Pemahaman Bisnis)

Tahap ini bertujuan untuk memahami kebutuhan perusahaan, permasalahan utama, dan tujuan akhir yang ingin dicapai. CV. Surya Kencana Sembako mengalami kesulitan dalam memperkirakan jumlah penjualan karena pengelolaan data penjualan serta stok masih dilakukan secara manual. Oleh karena itu diperlukan sistem prediksi untuk membantu perusahaan menentukan jumlah stok optimal setiap periode. Tujuan utama penelitian ini Adalah Membangun model prediksi penjualan, Mengidentifikasi variabel yang berpengaruh dan Mengintegrasikan model ke dalam sistem berbasis web.

2. Data Understanding (Pemahaman Data)

Tahap ini mencakup pengumpulan, eksplorasi, dan analisis awal terhadap data penjualan sembako yang diambil dari database CV. Surya Kencana Sembako. Peneliti mengidentifikasi struktur data, jenis fitur, jumlah data, pola nilai missing, serta kemungkinan outlier.

Variabel Dependen (Y): Jumlah Jual

Variabel Independen (X):

- X1 = Hari
- X2 = Ketersediaan barang
- X3 = Promosi
- X4 = Harga sembako

Alasan Pemilihan Variabel:

- Hari memengaruhi pola belanja konsumen (hari biasa, akhir pekan, hari libur).
- Ketersediaan barang memengaruhi potensi transaksi (stok kurang penjualan tertahan).
- Promosi meningkatkan daya tarik produk dan memicu kenaikan penjualan.
- Harga berpengaruh terhadap keputusan membeli meskipun cenderung elastis rendah untuk sembako.

3. Data Preparation (Persiapan Data)

Tahap ini meliputi proses pembersihan dan transformasi data agar siap digunakan dalam pemodelan.

Langkah-langkah yang dilakukan:

- Handling Missing Value yaitu Menghapus atau mengisi data kosong menggunakan metode statistik dasar (mean/median).
- Penghapusan Duplikasi yaitu Menghindari bias dari data ganda.
- Identifikasi Outlier yaitu Menghapus data yang tidak wajar atau terlalu ekstrem.
- Seleksi Fitur yaitu Menentukan fitur paling relevan untuk model.
- Split Data yaitu Data dibagi menjadi training 80% dan testing 20%, agar model dapat diuji pada data yang belum pernah dilihat. Tahap ini memastikan dataset yang digunakan berkualitas, bersih, dan valid untuk pemodelan statistik.

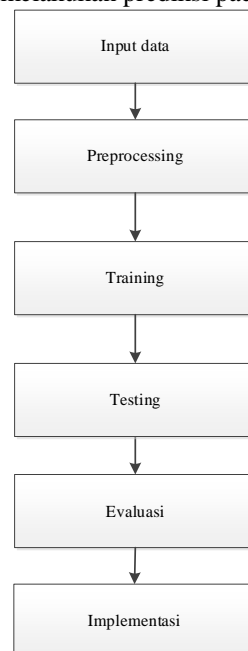
4. Modelling (Pemodelan)

Tahap ini berfokus pada prosedur pembuatan model tanpa melakukan hitungan manual. Metode yang digunakan adalah Regresi Linear Berganda, yaitu teknik statistik yang menganalisis hubungan antara satu variabel dependen dan beberapa variabel independen. Model kemudian dilatih menggunakan data training untuk menghasilkan parameter (koefisien regresi). Alasan Pemilihan Metode:

- Sederhana dan mudah diinterpretasikan perusahaan dapat memahami faktor-faktor apa yang paling memengaruhi penjualan.
- Cocok untuk prediksi numerik — jumlah penjualan merupakan variabel kontinu.
- Efektif untuk hubungan linear — cocok dengan karakteristik variabel seperti harga, stok, dan promosi.
- Banyak digunakan dalam penelitian serupa, sehingga mudah dibandingkan dengan studi terdahulu.
- Dapat digunakan untuk peramalan jangka pendek, sesuai kebutuhan perusahaan.

Prosedur Pemodelan:

- Memasukkan variabel X1–X4 sebagai input.
- Menjalankan algoritma regresi pada data training.
- Sistem menghasilkan nilai koefisien regresi dan persamaan model.
- Model kemudian digunakan untuk melakukan prediksi pada data testing dan data baru.



Gambar 1. Flowchart Penelitian

5. Evaluation (Evaluasi Model)

Model dievaluasi menggunakan empat metrik utama:

- MSE (Mean Squared Error)
- RMSE (Root Mean Squared Error)
- MAE (Mean Absolute Error)
- R^2 (Koefisien Determinasi)

Tujuan Evaluasi Mengukur seberapa jauh prediksi model mendekati data actual, mengetahui stabilitas model antara data training dan testing, menentukan apakah model layak digunakan dalam sistem riil.

6. Deployment (Implementasi Sistem)

Tahap akhir adalah menerapkan model yang telah dilatih ke dalam sistem informasi berbasis web.

Prosesnya meliputi:

- Integrasi model regresi ke backend (PHP).

2. Input variabel oleh admin atau sistem (stok, harga, hari, promosi).
3. Sistem menjalankan perhitungan prediksi otomatis.
4. Hasil prediksi ditampilkan pada dashboard untuk membantu pengambilan keputusan.

Tahap ini memastikan model dapat digunakan secara langsung oleh perusahaan dan menghasilkan manfaat nyata dalam operasi bisnis.

2.2 Split Data

Split Data adalah proses membagi dataset menjadi beberapa bagian untuk keperluan pelatihan (*training*) dan pengujian (*testing*) model *machine learning*. Tujuan *Split Data*:

- a) Belajar dari sebagian data (*training set*)
- b) Diuji performanya pada data yang belum pernah dilihat sebelumnya (*testing set*)
- c) Menghindari *overfitting*, yaitu model terlalu cocok dengan data latihan tapi buruk saat diuji [11]

2.3 Regresi Linear Berganda

Istilah ‘regresi’ pertama kali diperkenalkan oleh seorang ahli yang bernama Fancis Galton pada tahun 1886. Menurut Galton, analisis regresi berkenaan dengan studi ketergantungan dari suatu variabel yang disebut variabel tak bebas (*dependet variable*), pada satu atau variabel yang menerangkan dengan tujuan untuk memperkirakan ataupun meramalkan nilai-nilai dari variabel tak bebas apabila nilai variabel yang menerangkan sudah diketahui [12]

Regresi linier berganda merupakan salah satu metode yang ada di dalam Data Mining yang bisa melakukan kajian terhadap kebutuhan dari satu variabel dengan variabel lain yang memiliki tujuan untuk membuat estimasi rata-rata nilai variabel yang sudah diketahui. [13]

Persamaan umum regresi linier berganda:

$$\sum Y^2 = \sum Y^2 - n Y^2 \quad (1)$$

$$\sum X1^2 = \sum X1^2 - n X1^2 \quad (2)$$

$$\sum X2^2 = \sum X2^2 - n Y^2 \quad (3)$$

$$\sum X1Y = \sum X1Y - n X1Y \quad (4)$$

$$\sum X2Y = \sum X2Y - n X2Y \quad (5)$$

$$\sum X1X2 = \sum X1X2 - n X1X2 \quad (5)$$

$$b1 = (\sum X2^2)(\sum X1Y) - (\sum X1X2)(\sum X2Y) / (\sum X1^2)(\sum X2^2) - (\sum X1X2)^2 \quad (7)$$

$$b2 = (\sum X1^2)(\sum X2Y) - (\sum X1X2)(\sum X1Y) / (\sum X1^2)(\sum X2^2) - (\sum X1X2)^2 \quad (8)$$

$$a = Y - b1X1 - b2X2$$

Persamaan umum regresi linier berganda:

$$Y = a + b1X1 + b2X2 + \dots bnXn \quad (9)$$

Keterangan:

Y = Variabel terikat (*dependen*)

X = Variabel tidak terikat (*independen*)

a = Konstanta (*intercept*)

b = Koefisien regresi (kemiringan) besaran response yang ditimbulkan oleh variable [14]

3. HASIL DAN PEMBAHASAN

3.1. Pembahasan

1. Business Understanding

Tahap ini berfokus pada pemahaman awal yaitu mengenai masalah dan tujuan penelitian. Penulis telah mengidentifikasi permasalahan dalam penelitian ini, diantaranya :

- 1) CV. Surya Kencana Sembako memiliki data penjualan namun tidak diolah untuk memprediksi penjualan yang akan datang.
- 2) CV. Surya Kencana Sembako belum memiliki kejelasan mengenai variabel apa saja yang harus diperhitungkan dalam model prediksi penjualan
- 3) Bagaimana metode regresi linear dapat diterapkan untuk menemukan pengetahuan baru yang dapat mendukung pengambilan keputusan bisnis.

Tujuan penelitian ini adalah untuk mengolah data penjualan historis CV. Surya Kencana Sembako dengan menerapkan konsep data mining menggunakan metode regresi linear untuk memprediksi hasil penjualan, mengidentifikasi variabel yang relevan dalam model prediksi, serta mengevaluasi efektivitas metode regresi linear dalam memprediksi hasil penjualan.

2. Data Understanding (Pemahaman Data)

Variabel yang digunakan dalam pemodelan:

X1 – Hari

X2 – Ketersediaan Barang

X3 – Promosi

X4 – Harga Sembako

Y – Jumlah Jual

3.1. Hasil Penelitian

1. Preprocessing Data

Dari 1000 data awal, proses cleaning (menghapus missing value, outlier, dan seleksi fitur) menghasilkan 867 data final, yang kemudian dibagi menjadi:

Tabel 2. Data Final

No	Jenis data	Jumlah
1	Training	693
2	Testing	174

2. Modelling

Pada tahap pemodelan, algoritma *Linear Regression* diterapkan pada *dataset* yang telah disiapkan. Model *Linear Regression* dibangun dengan menggunakan variabel independen sebagai fitur dan variabel dependen sebagai target. Dari tabel diatas maka diperoleh hasil Keseluruhan Variabel Data Menjumlahkan keseluruhan variabel data dengan variabel x1,x2,x3,X4,y, dari tabel diatas maka diketahui sebagai berikut:

Tabel 3. Tabel Data Hasil Sigma

Keterangan	Total
X1	1.348
X2	54.442
X3	950
X4	13.729.900
Y	29.360
X1X2	106.175
X1X3	1.819
X1X4	27.003.100
X2X3	74.497
X2X4	936.403.450
X3X4	18.695.500
X1Y	57.269
X2Y	2.661.085
X3Y	39.959
X4Y	534.165.000
X1^2	2.894
X2^2	5.104.826
X3^2	1.464
X4^2	1.030.086.440.000
Y^2	1.642.128
Rata - Rata X1	1,95
Rata - Rata X2	78,56
Rata - Rata X3	1,37
Rata - Rata X4	19812,27
Rata Y	42,37

Menyederhanakan persamaan Regresi Linear Dari tabel diatas maka dilakukan perhitungan proses Regresi Linier Berganda dengan menggunakan rumus persamaan sebagai berikut :

$$\sum Y = n \cdot a + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 + b_4 \sum X_4 \quad (P1)$$

$$\sum Y \cdot X_1 = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 \cdot X_2 + b_3 \sum X_1 \cdot X_3 + b_4 \sum X_1 \cdot X_4 \quad (P2)$$

$$\sum Y \cdot X_2 = a \sum X_2 + b_1 \sum X_1 \cdot X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 \cdot X_3 + b_4 \sum X_2 \cdot X_4 \quad (P3)$$

$$\sum Y \cdot X_3 = a \sum X_3 + b_1 \sum X_1 \cdot X_3 + b_2 \sum X_2 \cdot X_3 + b_3 \sum X_3^2 + b_4 \sum X_3 \cdot X_4 \quad (P4)$$

$$\sum Y * X_4 = a \sum X_4 + b_1 \sum X_1 * X_4 + b_2 \sum X_2 * X_4 + b_3 \sum X_3 * X_4 + b_4 \sum X_4^2 \text{ (P5)}$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (2) adalah sebagai berikut:

$$\sum Y = n * a + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 + b_4 \sum X_4 \text{ (P1)}$$

$$\sum Y * X_1 = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 * X_2 + b_3 \sum X_1 * X_3 + b_4 \sum X_1 * X_4 \text{ (P2)}$$

$$\begin{array}{rcllclclcl} 29360 & = & 693a & + & 1348b1 & + & 54442b2 & + & 950b3 & + & 13729900b4 \\ 57269 & = & 1348a & + & 2894b1 & + & 106175b2 & + & 1819b3 & + & 27003100b4 \\ \hline 39.577.280 & = & 934.164a & + & 1.817.104b1 & + & 73.387.816b2 & + & 1.280.600b3 & + & 18.507.905.200b4 \\ 39.687.417 & = & 934.164a & + & 2.005.542b1 & + & 73.579.275b2 & + & 1.260.567b3 & + & 18.713.148.300b4 \\ \hline - 110.137 & = & & - & 188.438b1 & + & - 191.459b2 & + & 20.033b3 & + & -205.243.100b4 \end{array} \quad (-)$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (3) adalah sebagai berikut:

$$\begin{array}{rcllclclcl} 29360 & = & 693a & + & 1348b1 & + & 54442b2 & + & 950b3 & + & 13729900b4 \\ 2661085 & = & 54442a & + & 106175b1 & + & 5104826b2 & + & 74497b3 & + & 936403450b4 \\ \hline - & = & & & & + & & + & & + & \\ 45.714.785 & = & & - & 191.459b1 & + & -573.713.054b2 & + & 93.479b3 & + & 98.555.624.950b4 \end{array} \quad (-)$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (4) adalah sebagai berikut:

$$\begin{array}{rcllclclcl} 29360 & = & 693a & + & 1348b1 & + & 54442b2 & + & 950b3 & + & 13729900b4 \\ 39959 & = & 950a & + & 1819b1 & + & 74497b2 & + & 1464b3 & + & 19695500b4 \\ \hline & = & & & & + & & + & & + & \\ 200.413 & = & & 20.033b1 & + & 93.479b2 & + & - 112.052b3 & + & 87.423.500b4 \end{array} \quad (-)$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (5) adalah sebagai berikut:

$$\begin{array}{rcllclclcl} 29360 & = & 693a & + & 1348b1 & + & 54442b2 & + & 950b3 & + & 13729900b4 \\ & = & 1372990 & + & 27003100b1 & + & & + & 18695500 & + & \\ 534165000 & = & 0a & + & 1 & + & 936403450b2 & + & b3 & + & 1030086440000b4 \\ \hline (-) & = & & + & - & + & + & + & + & - & \\ 32.933.519.000 & = & & 205.243.100b1 & + & 98.555.624.950b2 & + & 87.423.500b3 & + & 525.339.748.910.000b3 \end{array}$$

Setelah melakukan proses eliminasi antara persamaan [1] hingga persamaan [5], maka diperoleh persamaan baru yaitu sebagai berikut :

$$\begin{array}{rcllclclcl} - 110.137 & = & - 188.438b1 & + & - 191.459b2 & + & 20.033b3 & + & -205.243.100b4 & & \text{(P6)} \\ - 45.714.785 & = & - 191.459b1 & + & -573.713.054b2 & + & 93.479b3 & + & 98.555.624.950b4 & & \text{(P7)} \\ 200.413 & = & 20.033b1 & + & 93.479b2 & + & - 112.052b3 & + & 87.423.500b4 & & \text{(P8)} \\ 32.933.519.000 & = & - & 205.243.100b1 & + & 98.555.624.950b2 & + & 87.423.500b3 & + & 525.339.748.910.000b4 & & \text{(P9)} \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [6] dengan persamaan [7] adalah sebagai berikut :

$$\begin{array}{rcllclclcl} - 110.137 & = & - 188.438b1 & + & - 191.459b2 & + & 20.033b3 & + & -205.243.100b4 \\ - 45.714.785 & = & - 191.459b1 & + & -573.713.054b2 & + & 93.479b3 & + & 98.555.624.950b4 \\ \hline & = & & & & + & - & + & & & (-) \end{array}$$

$$\begin{array}{rcllclclcl} 245602882,3 & = & & 573518525,6b2 & + & 73124,83695b3 & + & -98764158446b4 \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [7] dengan persamaan [8] adalah sebagai berikut :

$$\begin{array}{rcllclclcl} - 45.714.785 & = & - 191.459b1 & + & -573.713.054b2 & + & 93.479b3 & + & 98.555.624.950b4 & & \text{(P7)} \end{array}$$

$$\begin{array}{rclclcl}
 & = & & + & & + & 87.423.500 \text{ b4} \\
 200.413 & & 20.033 \text{ b1} & & 93.479 \text{ b2} & & - 112.052 \text{ b3} & & (P8) \\
 \hline
 & & & & & & & & (-) \\
 - & & & & & & & & \\
 243799438 & & & & -572819675,2 \text{ b2} & & -977401,964\text{b3} & & 99391131340\text{b4}
 \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [8] dengan persamaan [9] adalah sebagai berikut :

$$\begin{array}{rclclcl}
 & = & & + & & + & 87.423.500 \text{ b4} \\
 200.413 & & 20.033 \text{ b1} & & 93.479 \text{ b2} & & - 112.052 \text{ b3} & & (P8) \\
 & = & - & & + & & + & & - \\
 32.933.519.00 & & 205.243.100\text{b} & & 98.555.624.950\text{b} & & & & 525.339.748.910.0006 \\
 0 & & 1 & & 2 & & 87.423.50063 & & 4 (P9) \\
 \hline
 & & & & & & & & (-) \\
 - & & & & & & & & \\
 -34986798284 & & & & -99513339740 & & 1060576132 & & 524.444.074.170.860 \\
 & & & & \text{b2} & & \text{b3} & & \text{b4}
 \end{array}$$

Dari persamaan diatas maka diperoleh hasil sebagai berikut ;

$$\begin{array}{rclclcl}
 & = & & + & & + & & & (P10) \\
 245602882,3 & & 573518525,6\text{b2} & & 73124,83695\text{b3} & & -98764158446\text{b4} & &
 \end{array}$$

$$\begin{array}{rclclcl}
 -243799438 & = & -572819675,2 \text{ b2} & + & -977401,964\text{b3} & + & 99391131340\text{b4} & & (P11)
 \end{array}$$

$$\begin{array}{rclclcl}
 -34986798284 & = & -99513339740 \text{ b2} & + & 1060576132 \text{ b3} & + & 524.444.074.170.860 \text{ b4} & & (P12)
 \end{array}$$

Setelah melakukan proses eliminasi antara persamaan [10] hingga persamaan [11], maka diperoleh persamaan baru yaitu sebagai berikut :

$$\begin{array}{rclclcl}
 & = & & + & & + & & & (P10) \\
 245602882,3 & & 573518525,6\text{b2} & & 73124,83695\text{b3} & & -98764158446\text{b4} & &
 \end{array}$$

$$\begin{array}{rclclcl}
 -243799438 & = & -572819675,2 \text{ b2} & + & -977401,964\text{b3} & + & 99391131340\text{b4} & & (P11)
 \end{array}$$

$$\begin{array}{rclclcl}
 \hline
 & & & & & & & & (-) \\
 2269353,025 & = & & & 1091096,409\text{b3} & & -1231395494 \text{ b4} & &
 \end{array}$$

Setelah melakukan proses eliminasi antara persamaan [11] hingga persamaan [12], maka diperoleh persamaan baru yaitu sebagai berikut :

$$\begin{array}{rclclcl}
 -243799438 & = & -572819675,2 \text{ b2} & + & -977401,964\text{b3} & + & 99391131340\text{b4} & & (P11)
 \end{array}$$

$$\begin{array}{rclclcl}
 -34986798284 & = & -99513339740 \text{ b2} & + & 1060576132 \text{ b3} & + & 524.444.074.170.860 \text{ b4} & & (P12)
 \end{array}$$

$$\begin{array}{rclclcl}
 \hline
 & & & & & & & & (-) \\
 -77209256852 & & & & 1262226066 & + & 541.367.754.109.026 & &
 \end{array}$$

Dari persamaan diatas maka diperoleh hasil sebagai berikut ;

$$\begin{array}{rclclcl}
 2269353,025 & = & 1091096,409\text{b3} & & -1231395494 \text{ b4} & & \text{P13}
 \end{array}$$

$$\begin{array}{rclclcl}
 -77209256852 & = & 1262226066 & + & 541.367.754.109.026 & & \text{P14}
 \end{array}$$

Selanjutnya melakukan proses eliminasi antara persamaan [13] dengan persamaan [14] adalah sebagai berikut :

$$\begin{array}{rclclcl}
 2269353,025 & = & 1091096,409\text{b3} & & + & -1231395494 \text{ b4} & \text{P13}
 \end{array}$$

$$\begin{array}{rclclcl}
 -77209256852 & = & 1262226066\text{b3} & & + & 541.367.754.109.026\text{b4} & \text{P14}
 \end{array}$$

$$\begin{array}{rclclcl}
 \text{b4} & = & -0.0001470812
 \end{array}$$

Maka nilai b4 adalah -0.0001470812.

Kemudian untuk menghasilkan nilai b3 maka hasil dari b4 dimasukan ke persamaan [13] atau persamaan [14], sebagai berikut:

$$\begin{array}{rclclcl}
 2269353,025 & = & 1091096,409\text{b3} & + & -1231395494\text{b4} & & \text{P13}
 \end{array}$$

$$\begin{array}{rclclcl}
 2269353,025 & = & 1091096,409\text{b3} & + & -1231395494\text{b4} \times -0.0001470812
 \end{array}$$

$$\begin{array}{rclclcl}
 \text{b3} & = & 1.91389
 \end{array}$$

Kemudian untuk menghasilkan nilai b2 maka hasil dari b3 dan b4 dimasukan ke persamaan [10], persamaan [11] atau persamaan [12], sebagai berikut:

$$\begin{array}{rclclcl}
 245602882,3 & = & 573518525,6\text{b2} & + & -73124,83695\text{b3} & + & -98764158446\text{b4} & & (P10)
 \end{array}$$

$$\begin{array}{rclclcl}
 & = & & + & - & + & -98764158446\text{b4} \times -
 \end{array}$$

$$\begin{array}{rclclcl}
 245602882,3 & & 573518525,6\text{b2} & & 73124,83695 \times 1.91389 & & 0.0001470812
 \end{array}$$

$$b_2 = 0.453323$$

Kemudian untuk menghasilkan nilai b_1 maka hasil dari b_2 , b_3 dan b_4 dimasukan ke persamaan [6] sebagai berikut:

$$-110.137 = -188.438b_1 - 191.459b_2 + 20.033b_3 + -205.243.100b_4 \quad (P6)$$

$$-110.137 = -188.438b_1 - 191.459b_2 * 0.45 + 20.033 * 1.91389 + 0.0001470812$$

$$b_1 = 0.48754$$

Maka dengan demikian dapat diperoleh hasil nilai b_1 , b_2 , b_3 dan b_4 adalah sebagai berikut :

$$b_1 = 0.48754$$

$$b_2 = 0.453323$$

$$b_3 = 1.91389$$

$$b_4 = -0.0001470812$$

Kemudian untuk mendapatkan nilai a maka hasil dari b_1 , b_2 , b_3 dimasukan ke persamaan [1] atau [2] atau [3] ataupun [4], dalam hal ini menggunakan persamaan [1] sebagai berikut

$$2936 = 0 + 693a + 1348b_1 + 54442b_2 + 950b_3 + 13729900b_4$$

$$2936 = 0 + 693a + 1348 * 0.48754 + 54442 * 0.45332 + 950 * 1.91389 + 13729900 * -0.0001470812$$

$$a = 6.0954$$

Berdasarkan hasil regresi:

a) Intercept (a) = 6.0954.

b) b_1 (Hari) = 0.48754

Hari memiliki pengaruh positif kecil. Pada hari tertentu (misalnya akhir pekan), permintaan cenderung meningkat.

c) b_2 (Ketersediaan Barang) = 0.453323

Semakin besar stok yang tersedia, semakin besar pula potensi jumlah barang yang terjual. Ini terlihat sebagai faktor kunci dalam perputaran produk.

d) b_3 (Promosi) = 1.91389

Paling Berpengaruh Positif: setiap kenaikan 1 unit pada pengukuran promosi diasosiasikan dengan kenaikan ~1.91 unit penjualan (kontrol variabel lain). Ini menandakan efektivitas promosi pada peningkatan penjualan sembako prioritas strategi yang jelas.

e) b_4 (Harga) = -0.000147

Harga memiliki pengaruh negatif kecil terhadap penjualan. Artinya, peningkatan harga menurunkan penjualan, namun dampaknya tidak drastis (produk sembako termasuk kebutuhan pokok).

Maka :

$$Y = a + b_1 * X_1 + b_2 * X_2 + b_3 * X_3 + b_4 * X_4$$

$$= 6,0954 + (0,48754 * 1) + (0.453323 * 20) + (1.91389 * 2) + (-0.0001470812 * 48000)$$

$$= 12.42$$

Jadi menurut perhitungan diatas maka estimasi penjualan dengan diketahui metode regresi linear berganda, maka yang akan di dapatkan pada kedepannya diestimasi penjualan adalah 12.42.

3.2.Hasil

Nilai Error adalah selisih antara nilai pengamatan yang sebenarnya dengan nilai prediksi. Untuk prediksi yang diukur menggunakan MAE, MSE, dan RMSE nilai yang paling baik adalah nilai yang paling kecil. Berbeda dengan R2-Score nilai koefisien determinan yang mendekati 1 mengartikan bahwa variabel independennya memberikan semua informasi yang diperlukan dalam mengprediksi variabel dependennya.

Tabel 4. Hasil Evaluasi

No	X1	X2	X3	X4	Predicted	Data Aktual (Jumlh jual)	Selisih	Selisih \wedge^2	Nilai Variasi
1	1	20	2	48000	12,42	15	2,58	6,67	1228,2
2	1	120	2	18000	62,16	68	5,84	34,08	322,3
3	1	40	2	12000	26,78	30	3,22	10,38	401,8
4	1	180	2	5000	91,27	98	6,73	45,24	2299,6
5	1	130	2	15000	67,14	70	2,86	8,20	398,2
...
172	2	35	1	36000	19,56	21	1,44	2,09	843,7
173	2	25	1	70000	10,02	13	2,98	8,87	1372,4

No	X1	X2	X3	X4	Predicted	Data Aktual (Jumlah jual)	Selisih	Selisih ^2	Nilai Variasi
174	2	30	1	45000	15,97	18	2,03	4,14	1026,9
Total					6931.42	8708	1776.58	90601.82	60691.63
Rata – Rata					50.04				

1. MSE

MSE (*Mean Squared Error*) adalah salah satu teknik evaluasi yang digunakan untuk mengukur seberapa baik model regresi memprediksi nilai numerik. MSE mengukur rata-rata dari kuadrat selisih antara nilai aktual (asli) dan nilai prediksi.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = 90601.82 / 174 = 520.70$$

2. RMSE

Root Mean Squared Error (RMSE), merupakan nilai rata-rata dari jumlah kuadrat kesalahan, juga dapat menyatakan ukuran besarnya kesalahan yang dihasilkan oleh suatu model prediksi. [15]

$$RMSE : \frac{\sum_{i=1}^n \sqrt{y_i - \hat{y}_i^2}}{n} = \sqrt{520.70} = 22.82$$

3. Mean Absolute Error (MAE)

MAE (*Mean Absolute Error*) adalah teknik evaluasi yang digunakan untuk mengukur rata-rata kesalahan absolut antara nilai aktual dan prediksi pada model regresi.

$$MAE = \frac{1}{n} \sum_{i=1}^n (A_i - F_i) = 1776.58 / 174 = 10.21$$

4. R²

R-Square (R²) atau Koefisien Determinasi adalah teknik evaluasi yang digunakan untuk mengukur seberapa baik model regresi menjelaskan variabilitas data target.

$$R^2 = 1 - \frac{1776.58}{60691.63}$$

$$R^2 = 1 - 0.0292$$

R² = 0.97 atau 97% yang menunjukkan bahwa model cukup baik dalam menjelaskan variabilitas data. menunjukkan model menjelaskan sebagian besar variasi penjualan dalam dataset perusahaan. Ini berarti variabel yang dipilih (hari, ketersediaan, promosi, harga) sangat relevan untuk kasus ini. Maka dengan demikian Nilai R² mencapai 0.97 menunjukkan model sangat baik dalam menjelaskan variasi data. Error rendah (MAE 10–13) menunjukkan prediksi mendekati nilai aktual. Hasil ini menunjukkan bahwa metode regresi linear berganda cocok digunakan untuk kasus prediksi penjualan sembako, terutama ketika variabel bersifat numerik dan hubungan antarvariabel cukup linier. Variabel paling berpengaruh Promosi dan Ketersediaan barang Variabel berpengaruh kecil adalah Hari dan Harga

Berikut ini merupakan hasil pengujian sistem pada data penjualan baru sebanyak 30 data penjualan dan dapat diperoleh hasil prediksi jumlah jual sembako dan menentukan nilai evaluasi seperti pada Tabel di bawah ini:

Tabel 5. Pengujian Sistem pada data penjualan baru

Data Aktual									
No	Produk	X1	X2	X3	X4	Predicted (Jumlah Jual)	Selisih	Selisih^2	Nilai Variasi
1	Saus Sambal 500g	1	18	2	8000	17.39	66	48.61	2362.93
2	Beras 5kg	1	120	2	80000	53.04	30	-23.04	530.84
3	Minyak Goreng 1L	1	100	2	18000	53.1	56	2.9	8.41
4	Gula Pasir 1kg	1	70	2	15000	39.94	54	14.06	197.68
5	Tepung Terigu 1kg	1	90	2	10000	49.74	34	-15.74	247.75
25	Mizone Laci Lemon	1	110	2	3500	59.76	66	6.24	38.94
26	You C1000 Apple	2	115	1	3950	60.54	63	2.46	6.05
27	Gery toya-toya stick coklat	2	82	1	500	46.08	50	3.92	15.37
28	Gery Toya Toya Beries	2	80	1	500	45.18	56	10.82	117.07
29	Gery Toya Toya pisang	2	90	1	500	49.71	67	17.29	298.94
30	Gery 2 Choco Roll Keju	2	94	1	500	51.52	66	14.48	209.67
Total						1169.56	1587	417.44	18845.13
Rata – Rata						38.99	52.90	13.91	628.71

Maka untuk menghitung nilai evaluasi adalah sebagai berikut :

Menghitung MSE

$$MSE = 18845.13 / 30$$

$$MSE = 628.18$$

Menghitung RMSE

$$RMSE = \sqrt{628.18}$$

$$RMSE = 25.06$$

Menghitung MAE

$$MAE = 417.44 / 30$$

$$MAE = 13.91$$

Menghitung R2

$$R^2 = 1 - (417.44 / 5398.7)$$

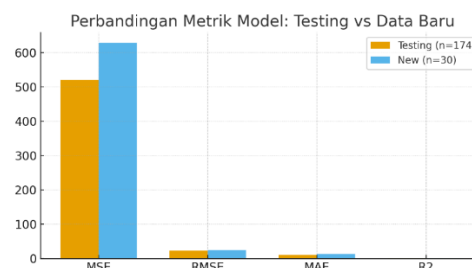
$$R^2 = 0.92 = 92\%$$

Berikut ini merupakan ringkasan hasil perhitungan evaluasi :

Tabel 6. Ringkasan hasil perhitungan evaluasi

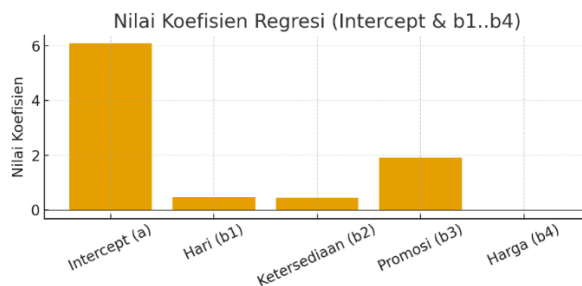
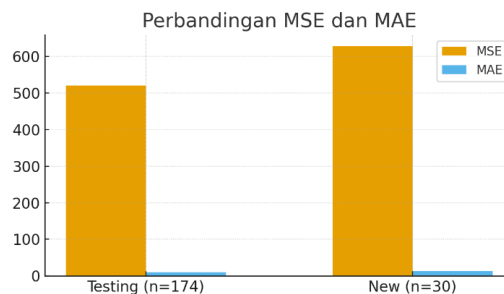
No	Hasil Pengujian Data Testing (174 Data)		Hasil Pengujian Data Baru (30 Data)	
1	Mektrik	Nilai	Mektrik	Nilai
2	MSE	520.70	MSE	628.18
3	RMSE	22.82	RMSE	25.06
	MAE	10.21	MAE	13.91
5	R ²	0.97	R ²	0.92

Model terbukti konsisten menunjukkan kinerja sangat baik pada kedua skenario pengujian. Berbagai penelitian sebelumnya menunjukkan bahwa metode data mining, khususnya regresi linear, telah banyak digunakan dalam proses peramalan di berbagai bidang. Penelitian oleh [16] studi harga saham (RMSE 53.5) bekerja di domain yang memiliki volatilitas tinggi dan unit pengukuran berbeda; banding langsung RMSE tanpa mempertimbangkan skala target (mean/median target) dapat menyesatkan. Dalam konteks Anda, MAE ~10–14 dan RMSE ~23–25 relatif kecil terhadap rata-rata penjualan (rata-rata $Y \approx 42.37$), sehingga performa model pada skala lokal ini sangat baik. Penelitian lain oleh [17] menerapkan algoritma C4.5 untuk memprediksi penjualan buku dan memperoleh tingkat akurasi tinggi dalam bentuk precision 100%, namun pendekatan tersebut lebih cocok untuk klasifikasi kategori seperti “Laris” atau “Tidak Laris”, bukan untuk estimasi jumlah penjualan secara numerik. Pada bidang otomotif, studi yang dilakukan [18] studi dengan algoritma C4.5 menghasilkan precision tinggi untuk klasifikasi “laris/tidak laris” (metrik berbeda). Untuk estimasi jumlah kuantitatif, regresi adalah metode yang lebih tepat; perbandingan yang adil harus menggunakan studi regresi pada domain ritel atau retail FMCG dan pada penelitian yang dilakukan menunjukkan R² jauh lebih tinggi (0.92–0.97) dibanding R² ~0.63 pada studi mobil bekas, menandakan bahwa variabel yang Anda pilih relevan dan hubungan linear cocok untuk dataset ini. Sementara itu, penelitian lain seperti oleh [19] menggunakan metode K-Means untuk mengelompokkan stok barang, tetapi model tersebut tidak menghasilkan nilai estimasi penjualan sehingga tidak dapat digunakan dalam konteks prediksi kuantitatif. Jika dibandingkan dengan penelitian-penelitian sebelumnya, penelitian ini memberikan kontribusi lebih spesifik pada industri sembako dengan memanfaatkan kombinasi variabel seperti harga, promosi, ketersediaan barang, dan faktor harian, sehingga menghasilkan model regresi linear berganda dengan performa yang jauh lebih tinggi (R² mencapai 0.92–0.97). Hal ini menunjukkan bahwa pendekatan regresi linear berganda sangat relevan dan efektif ketika variabel input dipilih secara tepat serta sesuai konteks operasional perusahaan, sehingga penelitian ini memperkuat bahwa model statistik sederhana tetap dapat menghasilkan prediksi yang sangat akurat bila diterapkan pada data yang tepat. Beberapa faktor yang mempengaruhi nilai R² lebih tinggi di bandingkan dengan penelitian sebelumnya adalah Variabel independen yang kuat (promosi dan ketersediaan) yang berhubungan langsung dengan penjualan harian, Dataset domain sempit pada distribusi sembako) pola lebih homogen dibanding dataset lintas-sektor lainnya. Pra-pemrosesan (pembersihan, outlier handling, seleksi fitur) yang ketat dengan data 1000 di awal hingga 867 final records membantu stabilitas model. Dan berikut ini merupakan grafik visual :



Gambar 2. Grafik perbandingan data testing dengan uji baru

Nilai R² tinggi (0.92–0.97) menunjukkan model menjelaskan sebagian besar variasi penjualan dalam dataset perusahaan. Ini berarti variabel yang dipilih (hari, ketersediaan, promosi, harga) sangat relevan untuk kasus ini. Perbedaan antara hasil testing (lebih baik) dan data baru (sedikit menurun) adalah wajar: n-testing lebih besar (174) dan estimasi metrik lebih stabil. Data baru (n=30) menunjukkan performa sedikit lebih rendah yaitu pantauan lebih lanjut atau validasi silang (cross-validation) direkomendasikan untuk memastikan generalisasi.

**Gambar 3.** Grafik Koefisien Regresi**Gambar 4.** Grafik perbandingan Evaluasi

MAE (10–14 unit) mengatakan bahwa rata-rata prediksi meleset sekitar 10–14 unit penjualan. Secara operasional, perusahaan dapat membandingkan angka ini terhadap ukuran batch barang seperti satuan pack untuk menilai signifikansi kesalahan.

4. KESIMPULAN

Penelitian ini berhasil membangun model prediksi penjualan sembako menggunakan metode regresi linear berganda pada CV. Surya Kencana Sembako. Hasil pengujian menunjukkan bahwa model memiliki performa prediksi yang sangat baik dan konsisten. Pada pengujian 30 data baru, diperoleh nilai MSE 628.18, RMSE 25.06, MAE 13.91, dan R^2 0.92. Sementara pada pengujian 174 data testing, performa meningkat dengan MSE 520.7, RMSE 22.82, MAE 10.21, dan R^2 0.97. Nilai error yang relatif rendah dan nilai R^2 yang tinggi membuktikan bahwa model mampu menjelaskan variasi penjualan secara akurat serta layak digunakan sebagai alat bantu estimasi penjualan. Kontribusi utama penelitian ini adalah menghasilkan model estimasi penjualan berbasis regresi linear berganda yang memanfaatkan data historis perusahaan secara optimal dan telah diintegrasikan ke dalam sistem informasi penjualan berbasis web. Implementasi ini memberikan manfaat praktis bagi perusahaan dalam perencanaan pengadaan stok, pengendalian persediaan, dan penentuan strategi penjualan secara lebih efisien dan berbasis data. Namun, penelitian ini masih memiliki keterbatasan karena model hanya menggunakan variabel internal dan mengasumsikan hubungan linier, sehingga belum mempertimbangkan faktor eksternal maupun pola non-linear. Oleh karena itu, penelitian lanjutan disarankan menambahkan variabel eksternal (misalnya harga pasar, inflasi, musiman) serta membandingkan performa model dengan metode non-linear atau time series seperti Random Forest, XGBoost, ARIMA, atau LSTM untuk meningkatkan akurasi prediksi.

REFERENCES

- [1] M. Lase, D. Saripurna, and V. W. Sari, "Estimasi Penjualan Ice Cream Walls Menggunakan Metode Regresi Linear Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 5, p. 625, 2022, doi: 10.53513/jursi.v1i5.5146.
- [2] M. J. Zaqy, R. Mentari, and M. Iqbal, "Implementasi Data Mining Untuk Memprediksi Penjualan Toko Roti Mawar Menggunakan Regresi Linier Berganda," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 6, no. 2, p. 385, 2023, doi: 10.53513/jsk.v6i2.8534.
- [3] A. Damayanti, F. D. Marleny, and A. A. Ningrum, "Implementasi Regresi Linear Berganda Untuk Prediksi Penjualan Pada Pt Trimandiri Sarana Propetindo Banjarmasin," *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, 2025, doi: 10.23960/jitet.v13i3.6679.
- [4] N. C. Florensa Nainggolan, A. F. Boy, and E. Elfritriani, "Penerapan Data Mining Untuk Prediksi Export Penjualan Produk Kerajinan Rotan Menggunakan Metode Regresi Linear Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 2, no. 5, p. 743, 2023, doi: 10.53513/jursi.v2i5.6779.
- [5] A. P. B. Eka, A. A. Bakri, and L. Yuliyani, "Utilizing linear regression to forecast the stock price fluctuations of top-rated companies," *J. Info Sains Inform. dan Sains*, vol. 14, no. 01, pp. 551–559, 2024, doi: 10.54209/infosains.v14i01.
- [6] C. K. Puteri and L. N. Safitri, "Analysis of linear regression on used car sales in Indonesia," *J. Phys. Conf. Ser.*, vol. 1469, no. 1, 2020, doi: 10.1088/1742-6596/1469/1/012143.
- [7] R. Ravitilova, D. Amelia, I. Saputra, and R. Rinda, "Forecasting Sales: A Comprehensive Analysis of Forecasting Techniques for Sales Budget Determination," *J. Audit. Pajak, Akunt. Publik*, vol. 2, no. 2, p. 97, 2023, doi:

- 10.32897/ajib.2023.2.2.3217.
- [8] M. D. Wahyudi, "Penerapan Data Mining Dengan Algoritma C4. 5 Dalam Prediksi Penjualan Buku," *J. Teknorama (Informatika dan ...)*, vol. 1, no. 1, pp. 1–6, 2023, [Online]. Available: <https://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/view/1%0Ahttps://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/download/1/1>
- [9] F. Zafira, B. Irawan, and A. Bahtiar, "Penerapan Data Mining Untuk Estimasi Stok Barang Dengan Metode K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 156–161, 2024, doi: 10.36040/jati.v8i1.8319.
- [10] N. Prayogi and L. Tanti, "Statistical Parabolic Projection Dalam Forecasting Jumlah Penjualan Mesin Otopcopy Pada CV. Indah Bersama Copier Berbasis Android Statistical parabolic projection in forecasting number of machinery sales photocopy on cv. Beautiful together android-based copier," *Jid*, no. 2, pp. 612–624, 2023.
- [11] Amalia Yunia Rahmawati, "Analisis Prediksi Dengan Metode Regresi Liineardi Pt. Eagle Industry Indonesia," vol. 5, no. July, pp. 1–23, 2020.
- [12] S. Lailiyah, A. Yusnita, and L. Hariri, "Prediksi Persediaan Bahan Baku Untuk Produksi Makanan Olahan 'Sanggar Krispi' Menggunakan Metode Regresi Linear Berganda," *Simkom*, vol. 8, no. 2, pp. 84–94, 2023, doi: 10.51717/simkom.v8i2.141.
- [13] C. Jhony Mantho Sianturi, M. Dayan Sinaga, N. Sari Br Sembiring, E. Ginting, U. K. Potensi Utama JIKL Yos Sudarso, and T. Mulia -Medan, "Metode Regresi Linear Berganda Dalam Prediksi Penjualan Produk Berbasis Web," *J. Inform. Manaj. dan Komput.*, vol. 16, no. 1, pp. 236–242, 2024.
- [14] A. A. Kurniawan, D. Mustikasari, A. Triyanto, and P. Raharjo, "Implementasi Metode Regresi Linear Berganda untuk Prediksi Harga Penjualan Material Paving Block pada CV . Difa Jaya Abadi," no. 10, 2025, doi: 10.52620/sainsdata.v3i2.281.
- [15] D. Miftahul Huda, G. Dwilestari, and A. Rizki Rinaldi, "Jurnal Informatika dan Rekayasa Perangkat Lunak Prediksi Harga Mobil Bekas Menggunakan Algoritma Regresi Linear Berganda," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 150–157, 2024.
- [16] A. Hurifiani, A. Irma Purnamasari, and I. Ali, "Penerapan Algoritma Regresi Linear Untuk Prediksi Penjualan Alat Tulis Kantor (Atk) Di Bumdes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, pp. 266–273, 2024, doi: 10.36040/jati.v8i1.8305.
- [17] F. Karina and D. A. Dermawan, "Prediksi Peningkatan Omzet Penjualan dengan Menggunakan Metode Regresi Linier Berganda (Studi Kasus: UB Makmur Surabaya)," *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 5, no. 1, pp. 27–34, 2024, doi: 10.26740/jeisbi.v5i1.58197.
- [18] A. A. Saputra, M. Munir, and Z. D. Rizki, "Peramalan Pendapatan dari Penjualan Bawang Merah Menggunakan Metode Regresi Linier Berganda," *Stain. (Seminar Nas. Teknol. Sains)*, vol. 2, no. 1, pp. 383–389, 2023, [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/stains/article/view/2900>
- [19] R. C. Sijabat, M. Hutasuht, P. O. Kronis, and R. C. Sijabat, "Jurnal sistem informasi tgd," vol. 1, no. November, pp. 763–769, 2022.