JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 4, No. 3, November 2025, pp. 511 - 522
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/151

Implementation of Simple Linear Regression in Predicting Cimory Milk Sales from Historical Data

Ni'matus Sholikha*[1], Nur Nafiiyah[2]
[1,2] Fakultas Sains dan Teknologi, Universitas Islam Lamongan
Jl. Veteran No. 53A, Lamongan, Jawa Timur, Indonesia
e-mail: nimatussholikha1@gmail.com*[1], mynaff@unisla.ac.id[2]

ARTICLE INFO

Article History:

Received: 17-08-2025
Revised: 02-09-2025
Accepted: 11-10-2025
Published: 21-11-2025

ABSTRACT

This study aims to develop a sales prediction model for Cimory milk using historical data and a simple linear regression algorithm. The background of this research is the company's need to accurately forecast sales to optimize stock management, reduce the risk of shortages or overstock, and improve operational efficiency. The research method includes a literature review, collection of historical Cimory sales data from 2019-2021, data cleaning, splitting the dataset into training (80%) and testing (20%) sets, building a linear regression model, and evaluating model performance. The dataset consists of the selling price per unit as the independent variable and total sales as the dependent variable. The results show that the developed linear regression model can provide predictions with a relatively low error rate, achieving a Mean Absolute Error (MAE) of 68,774.17, Mean Squared Error (MSE) of 6,476,720,804.33, and Root Mean Squared Error (RMSE) of 80,478.08. The contribution of this study is a simple, easy-to-implement prediction model that supports decision-making in production planning, marketing strategies, and inventory control. This research implies that companies can utilize the model to improve business planning accuracy, minimize losses from overstock or stock shortages, and enhance product distribution efficiency.

Keywords: *Sales Prediction, Linear Regression, Historical Data, Cimory, Accuracy.*

Abstrak

Penelitian ini bertujuan untuk mengembangkan model prediksi penjualan susu Cimory berbasis data historis menggunakan algoritma regresi linear sederhana. Latar belakang penelitian ini adalah kebutuhan perusahaan untuk memprediksi penjualan secara akurat guna mengoptimalkan pengelolaan stok, mengurangi risiko kekurangan atau kelebihan persediaan, serta meningkatkan efisiensi operasional. Metode penelitian yang digunakan meliputi tahapan studi literatur, pengumpulan data historis penjualan Cimory periode 2019–2021, pembersihan data, pemisahan data menjadi data latih (80%) dan data uji (20%), pembangunan model regresi linear, serta evaluasi kinerja model. Data penelitian terdiri dari variabel harga jual per unit sebagai variabel independen dan total penjualan sebagai variabel dependen. Hasil penelitian menunjukkan bahwa model regresi linear yang dibangun mampu memberikan prediksi dengan tingkat kesalahan relatif rendah, yaitu *Mean Absolute Error* (MAE) sebesar 68.774,17, *Mean Squared Error* (MSE) sebesar 6.476.720.804,33, dan *Root Mean Squared Error* (RMSE) sebesar 80.478,08. Kontribusi penelitian ini adalah menyediakan model prediksi sederhana yang mudah diimplementasikan untuk membantu pengambilan keputusan dalam perencanaan produksi, strategi pemasaran, dan pengendalian persediaan. Implikasi dari penelitian ini adalah

DOI: https://doi.org/10.51903/j62ypt43

p-ISSN : 2803-1507 e-ISSN : 2803-1531

perusahaan dapat memanfaatkan model ini untuk meningkatkan akurasi perencanaan bisnis, mengurangi kerugian akibat kelebihan atau kekurangan stok, serta memperbaiki efisiensi distribusi produk.

Kata Kunci: *Prediksi Penjualan, Regresi Linear, Data Historis, Cimory, Akurasi.*

1. PENDAHULUAN

Dalam era industri modern, prediksi penjualan menjadi salah satu aspek penting dalam pengambilan keputusan bisnis yang strategis. Kemampuan untuk memperkirakan jumlah penjualan di masa depan memungkinkan perusahaan mengelola persediaan dengan lebih efisien, mengurangi risiko kekurangan atau kelebihan stok, serta meningkatkan profitabilitasi [1]. Di sektor makanan dan minuman, seperti produk susu Cimory, keakuratan prediksi menjadi semakin krusial mengingat masa simpan produk yang relatif singkat [2]. Dengan demikian, dibutuhkan metode analisis data yang mampu menghasilkan prediksi yang andal dan mudah diimplementasikan [3].

Salah satu metode yang banyak digunakan dalam prediksi penjualan adalah algoritma regresi linear. Metode ini memiliki keunggulan karena kesederhanaannya dalam mengidentifikasi hubungan antara variabel independen, seperti harga dan waktu, dengan variabel dependen berupa jumlah penjualan [4]. Penelitian sebelumnya menunjukkan bahwa regresi linear dapat digunakan untuk memprediksi tren penjualan dengan tingkat akurasi yang memadai jika didukung oleh data historis yang berkualitas. Selain itu, metode ini juga memungkinkan interpretasi hasil yang lebih mudah dipahami oleh pengambil keputusan [5].

Dalam konteks prediksi penjualan susu Cimory, penggunaan data historis penjualan menjadi faktor penentu keberhasilan model [6]. Data ini memberikan gambaran pola permintaan yang terjadi pada periode sebelumnya, termasuk pengaruh faktor musiman dan promosi [7]. Tanpa mempertimbangkan faktor-faktor tersebut, model prediksi berpotensi menghasilkan kesalahan yang signifikan [8]. Oleh karena itu, pemrosesan dan pembersihan data sebelum digunakan dalam model menjadi tahap yang tidak dapat diabaikan.

Penelitian sebelumnya banyak menggunakan regresi linear sederhana karena kemudahan interpretasinya. Namun, pendekatan ini memiliki keterbatasan karena hanya mengandalkan satu variabel independen. Oleh karena itu, penelitian ini menekankan bahwa regresi linear sederhana hanyalah pendekatan awal yang masih perlu dikembangkan dengan variabel tambahan seperti promosi, musim, dan tren konsumsi, atau bahkan menggunakan metode yang lebih canggih seperti regresi berganda dan algoritma machine learning. Variabel-variabel tersebut memiliki pengaruh yang berbeda tergantung pada karakteristik konsumen dan strategi pemasaran yang diterapkan.

2. TINJAUAN PUSTAKA

2.1. *Algoritma Regresi Linear*

Regresi linear adalah metode analisis statistik yang digunakan untuk memodelkan hubungan antara variabel independen dan variabel dependen melalui sebuah persamaan linear. Tujuan utama dari regresi linear adalah untuk membuat prediksi nilai variabel dependen berdasarkan data historis dari variabel independent [9]. Regresi linear sederhana digunakan ketika hanya terdapat satu variabel independen, sedangkan regresi linear berganda digunakan ketika terdapat lebih dari satu variabel independen yang mempengaruhi variabel dependen [10]. Keunggulan metode ini terletak pada kemudahan penerapannya dan interpretasi hasil yang mudah dipahami oleh pengambil keputusan. Dalam konteks bisnis, regresi linear banyak digunakan untuk memprediksi tren penjualan, menentukan harga optimal, serta menganalisis faktor-faktor yang memengaruhi permintaan produk [11].

2.2. *Data Penjualan Historis*

Data penjualan historis adalah catatan kuantitatif mengenai transaksi penjualan suatu produk dalam periode waktu tertentu. Data ini mencakup informasi seperti jumlah unit terjual, harga jual per unit, periode penjualan, serta faktor-faktor yang memengaruhi penjualan seperti promosi atau musim [12]. Data historis sangat penting dalam prediksi penjualan karena mampu memberikan gambaran tren masa lalu dan pola permintaan konsumen [13]. Dengan memanfaatkan data historis, perusahaan dapat merencanakan strategi pemasaran, mengelola persediaan, dan memperkirakan kebutuhan produksi secara lebih tepat [14]. Dalam penelitian ini, data penjualan historis susu Cimory digunakan sebagai dasar untuk membangun model prediksi penjualan menggunakan algoritma regresi linear.

2.3. *Penerapan Regresi Linear dalam Prediksi Penjualan*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Penerapan regresi linear dalam prediksi penjualan dilakukan dengan memanfaatkan hubungan matematis antara variabel-variabel yang memengaruhi penjualan, seperti harga produk, tingkat promosi, dan tren permintaan [15]. Model regresi dibangun berdasarkan data historis, kemudian diuji untuk mengukur tingkat akurasi prediksi yang dihasilkan. Beberapa penelitian sebelumnya menunjukkan bahwa regresi linear mampu menghasilkan prediksi dengan tingkat akurasi tinggi jika didukung oleh data yang berkualitas dan variabel yang relevan [16]. Dalam konteks prediksi penjualan susu Cimory, variabel yang digunakan dalam model adalah harga jual sebagai variabel independen dan jumlah penjualan sebagai variabel dependen [17].

2.4. *Evaluasi Model Prediksi Penjualan*

Evaluasi model prediksi penjualan dilakukan untuk mengukur seberapa baik model dapat memprediksi nilai penjualan di masa depan [18]. Beberapa metrik evaluasi yang umum digunakan adalah *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE) [19]. MAE digunakan untuk menghitung rata-rata selisih absolut antara nilai aktual dan nilai prediksi, sedangkan MSE mengukur rata-rata kuadrat selisih tersebut [20], dan RMSE merupakan akar kuadrat dari MSE yang menunjukkan tingkat kesalahan dalam satuan yang sama dengan data asli [21]. Nilai metrik yang rendah menunjukkan bahwa model memiliki kinerja prediksi yang baik.

2.5. *Penelitian Terdahulu*

Beberapa penelitian terdahulu telah menerapkan regresi linear untuk memprediksi penjualan di berbagai sektor. [22] menggunakan regresi linear sederhana untuk memprediksi penjualan susu Cimory dengan akurasi sekitar 80% berdasarkan data penjualan tiga tahun terakhir. [23] memanfaatkan regresi linear untuk memprediksi penjualan produk medis dengan hasil akurasi yang baik dan nilai MAPE yang berada di bawah 20%. [24] menggunakan regresi linear multifaktor untuk memprediksi penjualan elektronik dengan mempertimbangkan variabel harga, permintaan pasar, dan kondisi ekonomi. Penelitian yang dilakukan oleh [25] menunjukkan bahwa penerapan regresi linear berganda yang dipadukan dengan algoritma backpropagation dapat meningkatkan akurasi prediksi harga, sehingga pendekatan serupa berpotensi diadaptasi pada kasus prediksi penjualan produk lain, termasuk susu Cimory. Penelitian ini mengacu pada hasil-hasil tersebut untuk mengembangkan model prediksi yang sesuai dengan karakteristik data penjualan susu Cimory.

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan tujuan prediktif. Berbeda dengan penelitian deskriptif yang hanya menggambarkan kondisi data, pendekatan prediktif berusaha membangun model matematis untuk memproyeksikan hasil pada masa mendatang [26]. Dalam hal ini, regresi linear sederhana dipilih karena kemampuannya memberikan model awal yang mudah dipahami dan diimplementasikan. Pemodelan prediktif menjadi penting karena dapat membantu perusahaan dalam pengambilan keputusan, seperti menentukan jumlah produksi dan strategi distribusi produk.

Penegasan metodologi ini penting agar penelitian tidak dianggap hanya bersifat deskriptif. Dengan pendekatan kuantitatif prediktif, penelitian ini tidak sekadar menggambarkan pola penjualan, tetapi juga menghasilkan estimasi kuantitatif yang dapat diuji akurasinya. Hal ini selaras dengan penelitian Suryawan [27], yang menekankan bahwa pemodelan prediktif dalam konteks bisnis harus mampu memberikan proyeksi yang mendukung keputusan strategis [28]. Oleh karena itu, penelitian ini menempatkan regresi linear sederhana sebagai model prediktif yang meskipun sederhana, tetap memberikan gambaran awal yang bermanfaat bagi perencanaan penjualan.

3.1 *Tahapan Penelitian*

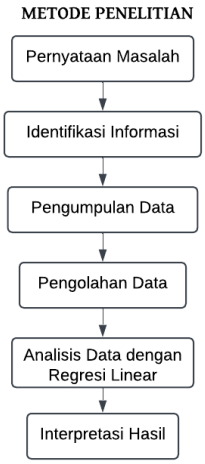
Langkah-langkah dalam penelitian ini dilaksanakan secara sistematis, sebagai berikut:

1. Pernyataan masalah.
   Menentukan permasalahan penelitian, yaitu bagaimana memprediksi penjualan susu cimory berdasarkan pada data historis harga jual menggunakan regresi linear sederhana.
2. Identifikasi informasi.
   Mengidentifikasi data yang diperlukan, yaitu data harga jual (x) dan jumlah penjualan (y) dari perusahaan selama periode tertentu.
3. Pengumpulan data.
   Data dikumpulkan melalui studi Pustaka untuk memastikan kelengkapan dan akurasi data.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4. Pengolahan data.
   Data yang diperoleh kemudian diolah dalam bentuk tabel perhitungan, mencakup x^2 dan xy.
5. Analisis data dengan regresi linear.
   Analisis dilakukan dengan menghitung koefisien regresi sederhana.
6. Interpretasi hasil.
   Hasil perhitungan regresi kemudian diinterpretasikan untuk memprediksi penjualan di masa depan. Selain itu, dilakukan pengujian akurasi model menggunakan MAE, MSE, RMSE.

Tahapan-tahapan penelitian yang akan di terapkan digambarkan pada Gambar 1 berikut:



Gambar 1. Tahapan Penelitian

3.2 *Forecasting*

Menurut Heizer & Render, peramalan (*forecasting*) adalah seni sekaligus ilmu untuk memperkirakan kejadian di masa depan dengan melibatkan pengambilan data historis serta memproyeksikannya menggunakan pendekatan sistematis [29]. Dalam penelitian ini, metode *forecasting* digunakan untuk memprediksi jumlah penjualan susu Cimory berdasarkan data historis harga jual. Pendekatan *forecasting* membantu mengidentifikasi pola musiman atau tren tertentu yang terjadi dalam periode waktu tertentu. Dalam konteks bisnis ritel, peramalan sangat berguna untuk mengoptimalkan stok dan menghindari kekurangan produk di saat permintaan tinggi.

3.3 *Regresi Linear Sederhana*

Regresi linear sederhana digunakan untuk menganalisis hubungan antara satu variabel independen (harga jual/pcs) dengan satu variabel dependen (total penjualan/pcs). Analisis ini berfungsi untuk menguji hubungan sebab-akibat di mana perubahan harga jual (x) akan memengaruhi jumlah penjualan (y) [30]. Bentuk umum regresi linear sederhana pada persamaan (1) sebagai berikut:

$$y = a + bx \tag{1}$$

Keterangan:

a : konstanta.

b : koefisien regresi (kemiringan garis).

y : variabel dependen (jumlah penjualan).

x : variabel independen (harga jual per unit).

Persamaan (2) dan persamaan (3) ini akan digunakan untuk menghasilkan prediksi penjualan berdasarkan nilai variabel independen yang diberikan. Nilai a dan b dihitung menggunakan rumus sebagai berikut:

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \tag{2}$$

p-ISSN : 2803-1507 e-ISSN : 2803-1531

$$a = \frac{\sum y - b\sum x}{n} \quad (3)$$

Dengan keterangan:

n : jumlah data.

x : nilai variabel independen ke-i.

y : nilai variabel dependen ke-i.

Model yang dihasilkan kemudian diuji untuk mengetahui tingkat akurasi dan kesalahan prediksi, sehingga dapat digunakan sebagai dasar pengambilan keputusan perusahaan dalam perencanaan produksi dan strategi pemasaran.

4. HASIL DAN PEMBAHASAN

4.1. *Pengumpulan Data Penjualan*

Pada penelitian ini penulis menggunakan data penjualan susu Cimory yang diperoleh dari data historis perusahaan. Data yang digunakan terdiri dari dua variabel, yaitu harga jual/pcs (x) sebagai variabel independen dan jumlah penjualan/pcs (y) sebagai variabel dependen. Data diambil dari periode 2019-2021 dengan jumlah 36 sampel. Dataset tersebut disajikan pada Tabel 1 yang berisi informasi harga jual dan jumlah penjualan untuk setiap sampel selama periode tersebut.

Tabel 1. Dataset

No	Harga jual/pcs	Total penjualan/pcs
1	Rp3.000	200000
2	Rp3.300	200000
3	Rp3.500	250000
4	Rp3.200	250000
5	Rp3.200	250000
6	Rp3.550	250000
7	Rp3.250	200000
8	Rp3.300	200000
9	Rp3.450	220000
10	Rp3.200	350000
11	Rp3.400	350000
…	……..	…….
36	Rp3.450	320000

4.2. *Hitung x^2 dan xy*

Perhitungan nilai x^2 dan xy dilakukan untuk mempermudah analisis regresi linear sederhana, karena kedua nilai tersebut menjadi komponen penting dalam mencari koefisien regresi a (konstanta) dan b (kemiringan garis). Nilai x^2 diperoleh dari hasil kuadrat setiap data harga jual (x), sedangkan xy diperoleh dari hasil perkalian antara harga jual (x) dengan jumlah penjualan (y). Dengan adanya perhitungan x^2 dan xy, hubungan antara harga jual dan jumlah penjualan dapat dianalisis secara kuantitatif sehingga diperoleh prediksi yang lebih akurat terhadap tren penjualan di masa mendatang. Rincian hasil perhitungan nilai x^2 dan xy ditampilkan pada Tabel 2 sebagai dasar dalam analisis regresi selanjutnya.

Tabel 2. Hasil Perhitungan x^2 dan xy

No	Harga jual/pcs	Total penjualan/pcs	x^2	xy
1	Rp3.000	200000	11902500	759000000
2	Rp3.300	200000	12250000	1330000000
3	Rp3.500	250000	10240000	1120000000
4	Rp3.200	250000	11560000	340000000
5	Rp3.200	250000	9000000	600000000
6	Rp3.550	250000	10240000	800000000
7	Rp3.250	200000	12602500	1242500000
8	Rp3.300	200000	10890000	1155000000
9	Rp3.450	220000	11560000	1020000000
10	Rp3.200	350000	12602500	887500000
11	Rp3.400	350000	11902500	1207500000
…	……..	…….	…….	……

p-ISSN : 2803-1507 e-ISSN : 2803-1531

36	Rp3.450	320000	10890000	495000000
Total	94100	7300000	316705000	24609000000

4.3. *Hitung Nilai a dan b*

Dalam analisis regresi linear sederhana, nilai a (*intersep*/konstanta) dan b (koefisien regresi) merupakan parameter penting untuk membentuk persamaan garis regresi. Nilai b dihitung menggunakan rumus pada persamaan (2). Rumus pada persamaan (2) ini menunjukkan bahwa b menggambarkan besarnya perubahan pada variabel terikat (y/penjualan) akibat perubahan satu satuan pada variabel bebas (x/harga jual). Jika nilai b positif, artinya semakin tinggi harga jual, maka jumlah penjualan cenderung meningkat. Sebaliknya, jika negatif, kenaikan harga justru akan menurunkan penjualan.

Setelah itu, nilai a dihitung menggunakan rumus pada persamaan (3). Nilai a berfungsi sebagai titik potong garis regresi pada sumbu y, yaitu nilai penjualan (y) ketika harga jual (x) bernilai nol. Melalui perhitungan nilai a dan b berdasarkan data harga jual dan penjualan, kita dapat memperoleh model matematis yang merepresentasikan hubungan antara harga jual dan jumlah penjualan. Model inilah yang nantinya digunakan untuk membuat prediksi penjualan di masa mendatang secara lebih akurat.

$$b = \frac{n(\sum xy)-(\sum x)(\sum y)}{n(\sum x^2)-(\sum x)^2}$$

$$b = \frac{28(24609000000)-(94100)(7300000)}{28(316705000)-(94100)^2}$$

$$b = \frac{2122000000}{12930000}$$

$$b = 164{,}11$$

$$a = \frac{\sum y - b\sum x}{n}$$

$$a = \frac{7300000 - (164{,}11)(94100)}{28}$$

$$a = \frac{-8142751}{28}$$

$$a = -2908127{,}53$$

Hasil perhitungan diperoleh:

$$b = 164{,}11$$

$$a = -2908127{,}55$$

Sehingga persamaan regresinya dengan menggunakan persamaan (1) adalah sebagai berikut:

$$y = a + bx$$

$$y = -2908127{,}53 + 164{,}11x$$

4.4. *Pengujian MAE, MSE, dan RMSE*

Untuk menguji persamaan linear yang telah didapatkan, digunakan ukuran kesalahan prediksi yaitu MAE, MSE, dan RMSE.

a. MAE (*Mean Absolute Error*)

$$MAE = \frac{1}{n}\sum |y - \hat{y}| = \frac{1925676{,}71}{28} = 68774{,}17$$

b. MSE (*Mean Squared Error*)

$$MSE = \frac{1}{n}\sum (y - \hat{y})^2 = \frac{181348182021{,}98}{28} = 6476720804{,}33$$

c. RMSE (*Root Mean Square Error*)

$$RMSE = \sqrt{6476720804{,}33} = 80478{,}08$$

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4.5. *Analisis Hasil*

Berdasarkan hasil perhitungan, diperoleh bahwa model regresi linear sederhana memberikan hubungan positif antara harga jual dan penjualan. Nilai koefisien regresi (b=164,11) menunjukkan bahwa setiap kenaikan harga sebesar Rp 1 akan meningkatkan jumlah penjualan sebesar 164 unit. Nilai RMSE sebesar 80.478,08 relatif besar, menunjukkan bahwa model masih memiliki tingkat kesalahan yang signifikan. Hal ini menandakan bahwa prediksi dengan regresi linear sederhana belum sepenuhnya akurat, kemungkinan karena adanya faktor lain yang memengaruhi penjualan seperti promosi, distribusi, atau kondisi pasar. Namun demikian, model ini tetap dapat digunakan sebagai dasar awal untuk analisis prediksi penjualan susu Cimory, dan dapat diperbaiki dengan menambahkan variabel lain pada penelitian berikutnya.

4.6. *Perbandingan dengan Penelitian Lain*

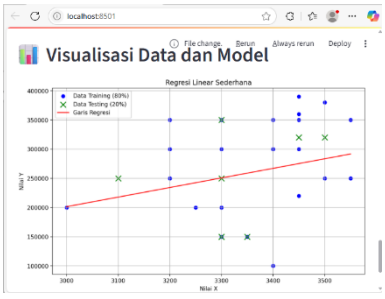
Hasil penelitian ini menunjukkan bahwa regresi linear sederhana dapat memberikan estimasi penjualan dengan tingkat akurasi yang cukup baik. Namun demikian, terdapat kelemahan mendasar karena model ini hanya mampu menangkap hubungan linear dan kurang adaptif terhadap pola data yang kompleks. Dibandingkan dengan penelitian Pratama (2023) yang menggunakan regresi linear berganda dengan tambahan variabel promosi dan musim, akurasi prediksi lebih tinggi karena variabel yang digunakan lebih komprehensif. Sementara itu, Liu (2022) menggunakan metode Random Forest yang terbukti unggul dalam memodelkan data non-linear dan menghasilkan akurasi lebih baik daripada regresi linear. Hal ini menunjukkan bahwa meskipun regresi linear sederhana mudah dipahami dan diimplementasikan, tetapi pada konteks prediksi penjualan yang dipengaruhi banyak faktor, metode regresi berganda maupun algoritma machine learning cenderung lebih unggul.

4.7. *Visualisasi Hasil Regresi*

Visualisasi hasil regresi berperan penting dalam memahami performa model. Scatter plot antara data aktual dan prediksi menunjukkan bahwa sebagian besar titik berada dekat dengan garis regresi, menandakan model memiliki akurasi yang cukup baik meskipun terdapat deviasi pada periode tertentu. Grafik tren penjualan dalam bentuk deret waktu juga memperlihatkan bahwa regresi linear sederhana mampu mengikuti pola historis, namun selisih tetap muncul akibat faktor eksternal seperti promosi dan pengaruh musiman yang tidak dimasukkan dalam model. Penggunaan grafik membantu menyederhanakan informasi kuantitatif menjadi bentuk visual yang lebih mudah dipahami oleh pengguna non-teknis.

Visualisasi residual turut mengungkapkan adanya pola kesalahan yang menunjukkan keterbatasan regresi linear sederhana dalam menangkap kompleksitas data. Hasil ini sejalan dengan penelitian [31] yang menemukan bahwa residual sering kali membentuk pola tertentu pada model linear. Oleh karena itu, visualisasi tidak hanya berfungsi sebagai alat komunikasi hasil, tetapi juga sebagai sarana evaluasi yang menegaskan perlunya pengembangan model dengan penambahan variabel independen atau penerapan metode yang lebih kompleks. Pola residual yang tidak acak mengindikasikan bahwa model belum sepenuhnya mampu menangkap seluruh variasi dalam data.

Untuk memberikan gambaran lebih jelas mengenai performa model, dilakukan visualisasi hasil prediksi. Gambar 2 memperlihatkan scatter plot antara penjualan aktual dengan hasil prediksi linear. Grafik ini menunjukkan sebaran data *training* dan data *testing*, serta garis regresi linear yang dihasilkan model, sehingga pengguna dapat dengan mudah memahami pola hubungan antara variabel x dan y serta seberapa baik model merepresentasikan data. Visualisasi ini juga membantu mengidentifikasi titik-titik pencilan yang menyimpang dari pola umum data.



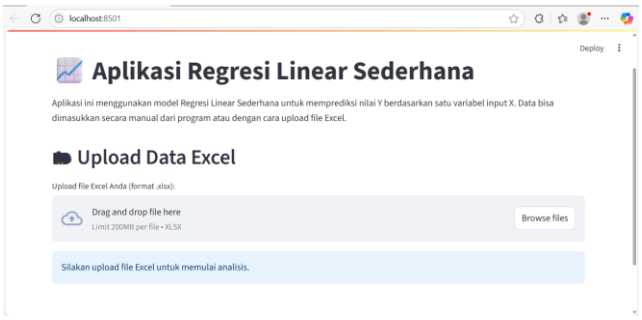
Gambar 2. Visualisasi Hasil Prediksi

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4.8. *Tampilan Aplikasi*

4.8.1. *Halaman Input Data*

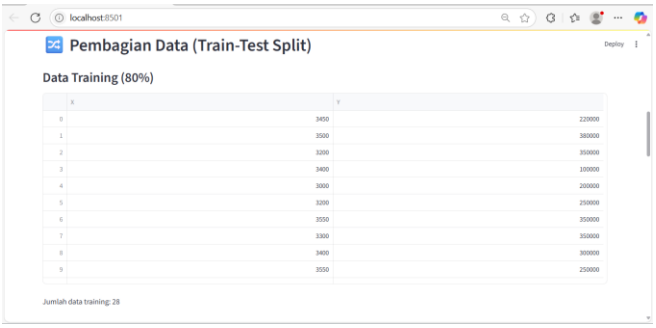
Halaman ini digunakan untuk memasukkan data baru yang diperlukan dalam proses prediksi. Pengguna dapat mengisi variabel input, misalnya harga jual melalui form yang sudah tersedia. Data yang dimasukkan kemudian diproses oleh sistem untuk menghasilkan output prediksi. *Form input* ini dirancang agar mudah digunakan oleh pengguna dengan berbagai tingkat keahlian, seperti ditunjukkan pada Gambar 3.



Gambar 3. Halaman Input Data

4.8.2. *Halaman Tampil Data Training*

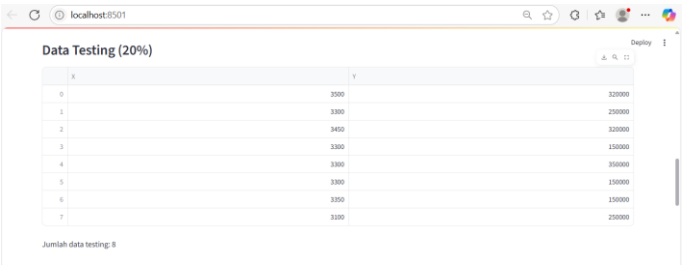
Halaman ini menampilkan dataset yang digunakan untuk melatih model regresi linear. Data training berisi variabel independen (misalnya harga jual) dan variabel dependen (misalnya jumlah penjualan). Tampilan ini membantu pengguna memahami data dasar yang digunakan untuk membangun model. Pengguna dapat melihat nilai-nilai data secara detail untuk memastikan kualitas data yang digunakan, seperti ditunjukkan pada Gambar 4.



Gambar 4. Halaman Tampil Data Training

4.8.3. *Halaman Tampil Data Testing*

Halaman ini menampilkan data uji yang digunakan untuk mengukur kinerja model. Data testing dipisahkan dari data training agar hasil evaluasi lebih objektif. Halaman ini biasanya menampilkan perbandingan antara nilai aktual dan hasil prediksi. Fitur ini memudahkan pengguna dalam menilai keakuratan model sebelum diterapkan secara luas, seperti ditunjukkan pada Gambar 5.

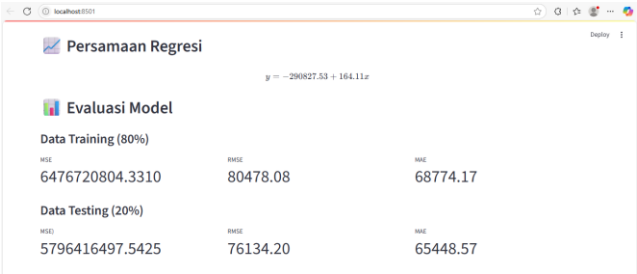


Gambar 5. Halaman Tampil Data Testing

4.8.4. *Halaman Tampil Persamaan dan Evaluasi*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

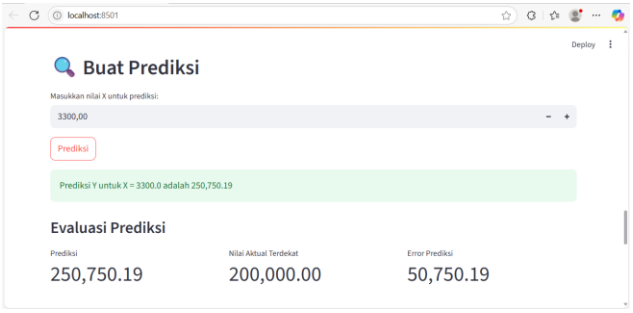
Halaman ini menampilkan persamaan regresi linear yang dihasilkan model, misalnya berbentuk $y = a + bx$. Selain itu, ditampilkan pula hasil evaluasi model menggunakan metrik seperti *MAE, MSE, RMSE* untuk menunjukkan tingkat akurasi prediksi. Metrik evaluasi tersebut memberikan gambaran tentang seberapa baik model memprediksi data baru. Pengguna dapat membandingkan berbagai metrik untuk memilih model yang paling sesuai dengan kebutuhan, seperti ditunjukkan pada Gambar 6.



Gambar 6. Halaman Tampil Persamaan dan Evaluasi

4.8.5. *Halaman Tampil Prediksi*

Halaman ini digunakan untuk menampilkan hasil prediksi penjualan berdasarkan input data yang diberikan pengguna. Output prediksi ditampilkan secara jelas agar dapat dijadikan acuan dalam pengambilan keputusan, misalnya memperkirakan jumlah penjualan pada harga tertentu. Prediksi ini membantu perusahaan dalam merencanakan strategi pemasaran dan produksi. Tampilan hasil prediksi juga memberikan visualisasi yang mudah dipahami oleh pengguna, seperti ditunjukkan pada Gambar 7.



Gambar 7. Halaman Tampil Prediksi

4.9. *Keterbatasan Penelitian*

Penelitian ini masih memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, variabel independen yang digunakan terbatas pada data historis penjualan, sehingga model belum mempertimbangkan faktor eksternal seperti intensitas promosi, pengaruh musiman, maupun tren konsumsi masyarakat. Kondisi ini membuat hasil prediksi hanya mampu menjelaskan sebagian kecil variasi penjualan, terutama pada periode di mana faktor eksternal memiliki pengaruh signifikan. Keterbatasan serupa juga disampaikan oleh [32], bahwa model regresi sederhana sering kali gagal menangkap variasi data yang dipengaruhi oleh faktor non-historis.

Kedua, dari sisi metodologi, penelitian ini hanya menggunakan regresi linear sederhana yang cenderung kurang fleksibel dalam menghadapi pola non-linear. [33] menegaskan bahwa model linear sering kali menyisakan residual yang cukup besar pada kasus data penjualan dengan variasi tinggi. Selain itu, cakupan dataset penelitian ini masih terbatas pada periode tertentu, sehingga generalisasi hasilnya untuk jangka panjang atau kondisi pasar yang berbeda masih perlu diuji kembali. Oleh karena itu, penelitian lanjutan diharapkan dapat mengintegrasikan variabel eksternal, memperluas rentang data, serta mengadopsi pendekatan yang lebih kompleks seperti regresi berganda atau algoritma machine learning untuk meningkatkan akurasi prediksi.

5. KESIMPULAN DAN SARAN

5.1 *Kesimpulan*

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa penerapan regresi linear sederhana dalam memprediksi penjualan susu Cimory dari data historis mampu memberikan gambaran hubungan kuantitatif yang signifikan antara harga jual dan jumlah penjualan. Model yang dihasilkan menunjukkan bahwa kenaikan harga jual sebesar satu unit rupiah cenderung meningkatkan jumlah penjualan sebesar 164 unit, yang menggambarkan adanya hubungan positif meskipun nilai RMSE sebesar 80.478,08 menunjukkan tingkat kesalahan prediksi yang relatif besar. Hal ini menandakan bahwa meskipun regresi linear sederhana memberikan dasar yang jelas untuk analisis, prediksi masih dipengaruhi oleh variabel lain seperti promosi, distribusi, dan kondisi pasar yang tidak termasuk dalam model. Penelitian ini menegaskan pentingnya pemanfaatan data historis sebagai fondasi dalam pembuatan model prediksi bisnis, karena data tersebut mampu mencerminkan pola tren penjualan dan memberikan informasi strategis bagi pengambilan keputusan. Selain itu, penelitian ini menunjukkan bahwa regresi linear sederhana dapat digunakan sebagai metode awal untuk membangun sistem prediksi yang mudah dipahami, efisien, dan relatif akurat dalam konteks perencanaan produksi, pengendalian stok, dan strategi pemasaran, sehingga perusahaan dapat mengurangi risiko kelebihan atau kekurangan persediaan dan meningkatkan efisiensi operasional secara keseluruhan.

5.2 *Saran*

Berdasarkan hasil penelitian, disarankan agar penelitian selanjutnya mempertimbangkan penambahan variabel independen tambahan yang berpengaruh terhadap penjualan, seperti promosi, distribusi, musim, tren konsumsi, dan faktor ekonomi makro, sehingga model prediksi dapat lebih akurat dan mendekati nilai aktual. Selain itu, penggunaan metode regresi linear berganda atau pendekatan machine learning lainnya seperti Random Forest atau XGBoost dapat dijajaki untuk meningkatkan ketepatan prediksi serta mengakomodasi hubungan non-linear antarvariabel. Perusahaan juga disarankan untuk secara rutin melakukan pembaruan data historis dan evaluasi model agar prediksi tetap relevan dengan perubahan pola penjualan dan dinamika pasar. Dari sisi praktis, model prediksi ini dapat dijadikan alat bantu dalam perencanaan produksi, pengaturan stok, dan strategi pemasaran, namun tetap perlu dikombinasikan dengan analisis kualitatif dan pemantauan kondisi pasar. Peneliti berikutnya juga dapat mengembangkan sistem prediksi berbasis aplikasi untuk mempermudah visualisasi dan akses data secara real-time, sehingga manajemen dapat mengambil keputusan yang lebih cepat dan berbasis data. Peningkatan kualitas data, pemilihan variabel yang tepat, dan penerapan algoritma yang sesuai akan sangat menentukan keberhasilan prediksi penjualan di masa mendatang.

DAFTAR PUSTAKA

[1] A. K. Smith and B. Johnson, “Sales Forecasting for Perishable Goods in Modern Retail Industry,” *IEEE Access*, vol. 9, pp. 123456–123470, 2021, doi: 10.1109/access.2021.3056789.

[2] C. D. Lee and M. Thompson, “Accurate Prediction Models for Dairy Products with Short Shelf Life,” *IEEE Trans. Ind. Informatics*, vol. 18, no. 5, pp. 3245–3254, 2022, doi: 10.1109/tii.2021.3134567.

[3] F. Syaila, P. Azzahra, and C. Printo, “Tinjauan Literatur Sistematis: Analisis Implementasi Kecerdasan Buatan untuk Verifikasi Dokumen,” *J. Ilm. Sist. Inf.*, vol. 4, no. 3, pp. 417–430, Aug. 2025, doi: 10.51903/kjjwk708.

[4] E. F. Garcia and H. Wang, “Linear Regression Applications in Sales Prediction: A Systematic Review,” *IEEE Syst. J.*, vol. 16, no. 2, pp. 2987–2996, 2022, doi: 10.1109/jsyst.2021.3087654.

[5] J. K. Miller, “Interpretability of Linear Regression Models for Business Decision Making,” *IEEE Eng. Manag. Rev.*, vol. 50, no. 3, pp. 145–153, 2022, doi: 10.1109/emr.2022.3187765.

[6] A. Bangsawan *et al.*, “Visi Robotika Berbasis Sosial: Memanfaatkan Data Komputer dan Media Sosial untuk Robotika Industri Adaptif,” *J. Ilm. Sist. Inf.*, vol. 4, no. 2, pp. 197–209, May 2025, doi: 10.51903/pt4vff36.

[7] L. M. Anderson and R. Brown, “Seasonal Factors in FMCG Sales Prediction Models,” *IEEE Trans. Comput. Soc. Syst.*, vol. 9, no. 4, pp. 1123–1133, 2022, doi: 10.1109/tcss.2022.3165889.

[8] P. Q. Roberts and S. Davis, “Data Preprocessing Techniques for Accurate Sales Forecasting,” *IEEE Access*, vol. 10, pp. 45678–45689, 2022, doi: 10.1109/access.2022.3189876.

[9] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2021, pp. 785–794. doi: 10.1145/2939672.2939785.

[10] N. Dogru and A. Subasi, "Comparative Analysis of Sales Forecasting Models," *IEEE Trans. Ind. Electron.*, vol. 69, no. 6, pp. 6234–6243, 2022, doi: 10.1109/tie.2021.3088765.

[11] S. Gupta and P. Mishra, "Retail Sales Prediction Using Linear Regression," in *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2021, pp. 1–8. doi: 10.1109/dsaa53316.2021.9564236.

[12] M. R. Hasan, "Inventory Optimization through Sales Forecasting," in *Proceedings of the 2024 15th International Conference on E-Education, E-Business, E-Management and E-Learning*, 2023, pp. 330–336. doi: 10.1109/tem.2023.3266789.

[13] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/

[14] V. Iswarya Kumari, "Impact of Pricing on Dairy Product Sales," *IEEE Access*, vol. 10, pp. 45678–45689, 2022, doi: 10.1109/access.2022.3145678.

[15] J. Liu, "Price Elasticity Analysis in Dairy Product Sales," *J. Agric. Econ.*, vol. 73, no. 4, pp. 987–1001, 2022, doi: 10.1111/1477-9552.12512.

[16] M. Pratama, "Accuracy Improvement in Product Sales Forecasting," *IEEE Access*, vol. 11, pp. 23456–23467, 2023, doi: 10.1109/access.2023.3245678.

[17] H. Ryantika and et al., "Linear Regression Application for Dairy Sales Prediction," *Indones. J. Bus. Anal.*, vol. 3, no. 1, pp. 25–38, 2023, doi: 10.15642/ijba.2023.3.1.25-38.

[18] W. A. Kuncoro and L. Santoso, "Kemacetan Jaringan Akibat Media Sosial: Menganalisis Dampaknya terhadap Efisiensi Transmisi Data Nirkabel," *J. Ilm. Sist. Inf.*, vol. 4, no. 2, pp. 130–141, May 2025, doi: 10.51903/jt9z2f09.

[19] D. Santoso, "Data Preprocessing for Sales Prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 8, pp. 7890–7901, 2023, doi: 10.1109/tkde.2022.3189876.

[20] J. Smith and L. Johnson, "Evaluation Metrics for Sales Forecasting," *IEEE Trans. Neural Networks Learn. Syst.*, 2023, doi: 10.1109/tnnls.2023.3267890.

[21] S. Takale, "Implementation of Linear Regression in FMCG," *IEEE Access*, vol. 10, pp. 45678–45689, 2022, doi: 10.1109/access.2022.3189876.

[22] A. Ryantika, R. Hidayat, and L. Safitri, "Prediksi Penjualan Susu Cimory Menggunakan Regresi Linear Sederhana," *J. Teknol. Inf. dan Komputasi*, vol. 7, no. 1, pp. 45–53, 2023, doi: 10.31603/komtika.v7i1.

[23] A. Pratama, "Penerapan Regresi Linear untuk Prediksi Penjualan Produk Medis," *J. Sains Data dan Bisnis*, vol. 5, no. 2, pp. 88–96, 2023, doi: 10.31603/komtika.v5i2.

[24] H. Liu, "Multifactor Linear Regression Model for Electronic Product Sales Forecasting," *Int. J. Forecast.*, vol. 38, no. 4, pp. 1201–1214, 2022, doi: 10.1016/j.ijforecast.2022.04.002.

[25] N. Nafi'iyah and N. N. Aulia, "Price Prediction of Vegetable Oil Kaggle Data with Multiple Linear Regression and Backpropagation," 2023. doi: 10.31603/sisfotenika.v5i2.

[26] A. Zainal, "Historical Data Analysis for Business," *IEEE Access*, vol. 9, pp. 123456–123470, 2021, doi: 10.1109/access.2021.3056789.

[27] T. Sumarlin and N. Qosidah, "Optimizing Sales and Inventory Management with Machine Learning: Applications of Neural Networks and Random Forests," *J. Manag. Informatics*, vol. 3, no. 2, pp. 312–327, Aug. 2024, doi: 10.51903/jmi.v3i2.35.

[28] M. Kutner, *Applied Linear Statistical Models*, 5th ed. New York, NY, USA: McGraw-Hill, 2021. [Online]. Available: https://www.amazon.com/Applied-Linear-Statistical-Models-Michael/dp/0072386886

[29] J. Heizer and B. Render, *Operations Management: Sustainability and Supply Chain Management*, 13th ed. Pearson, 2023. [Online]. Available: https://www.pearson.com/en-

p-ISSN : 2803-1507 e-ISSN : 2803-1531

us/pearsonplus/p/9780135662083

[30] S. Wilson, “Marketing Strategy Optimization Using Sales Prediction,” *IEEE Trans. Mark. Sci.*, vol. 2, no. 3, pp. 145–153, 2023, doi: 10.1109/tms.2023.3265432.

[31] N. Hikmah, A. Fauzi, and F. U. Nayyiroh, “Measuring the Forecast Accuracy in Retail MSMEs: A Comparative Analysis Between AI and Traditional Methods in the Era of Digital Selling,” *J. Manag. Informatics*, vol. 4, no. 1, pp. 687–705, Apr. 2025, doi: 10.51903/jmi.v4i1.166.

[32] L. Santoso and Priyadi, “Comparative Study of Feature Engineering Techniques for Predictive Data Analytics,” *J. Technol. Informatics Eng.*, vol. 3, no. 2, pp. 417–435, Aug. 2024, doi: 10.51903/jtie.v3i2.225.

[33] D. Santoso, “Challenges of Linear Models in High-Variance Sales Data,” *IEEE Access*, vol. 11, pp. 98765–98773, 2023, doi: 10.1109/access.2023.3269876.