

ANALISIS PREDIKSI PENJUALAN TISU MENGUNAKAN REGRESI LINEAR

Ichlas Ardhanur¹, Martanto², Arif Rinaldi Dikananda³, Mulyawan⁴

¹Teknik Informatika, STMIK IKMI CIREBON, Jl. Perjuangan No.10B Cirebon, 45135

²Manajemen Informatika, STMIK IKMI CIREBON, Jl. Perjuangan No.10B Cirebon, 45135

³Rekayasa Perangkat Lunak, STMIK IKMI CIREBON, Jl. Perjuangan No.10B Cirebon, 45135

⁴Program Studi Sistem Infomasi, STMIK IKMI CIREBON, Jl. Perjuangan No.10B Cirebon, 45135

Received: 2 Maret 2025

Accepted: 25 Maret 2025

Published: 14 April 2025

Keywords:

Prediksi, Regresi Linear,
Penjualan.

Correspondent Email:

ichlasardhanur@gmail.com

Abstrak. Penelitian ini dilakukan untuk mendukung pengambilan keputusan dalam strategi produksi dan distribusi penjualan tisu yang efisien. Dalam dunia bisnis yang semakin kompetitif, prediksi penjualan yang akurat menjadi elemen penting untuk memastikan ketersediaan produk sesuai dengan permintaan pasar sekaligus menghindari kelebihan atau kekurangan stok. Model regresi linear dipilih dalam penelitian ini karena kesederhanaannya, penerapannya yang luas, dan kemampuannya untuk memberikan gambaran hubungan antara variabel-variabel tertentu terhadap penjualan. Penelitian ini bertujuan untuk memprediksi penjualan tisu menggunakan regresi linear dan mengevaluasi keakuratan modelnya melalui metrik R-squared (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan Mean Absolute Error (MAE). Hasil penelitian menunjukkan bahwa model regresi linear dapat menjelaskan 83% variasi data penjualan tisu ($R^2 = 0,83$), menunjukkan efektivitas model ini dalam menggambarkan hubungan variabel. Namun, nilai RMSE sebesar 78,34 dan MAE sebesar 56,69 menunjukkan adanya kesalahan prediksi yang signifikan. Oleh karena itu, disarankan untuk menambahkan variabel prediktor lain, seperti faktor musiman atau promosi, serta menggunakan model lebih kompleks, seperti regresi non-linear atau Random Forest, untuk hasil yang lebih akurat. Penelitian ini menyimpulkan bahwa regresi linear merupakan metode dasar yang bermanfaat, tetapi pengembangan model dan data yang lebih komprehensif diperlukan untuk meningkatkan akurasi prediksi.

Abstract. This study was conducted to support decision-making in efficient production and distribution strategies for tissue sales. In an increasingly competitive business world, accurate sales forecasting is an essential element to ensure that product availability matches market demand while avoiding overstocking or stockouts. Linear regression was chosen for this study due to its simplicity, widespread application, and ability to provide insights into the relationship between specific variables and sales. The study aims to predict tissue sales using linear regression and evaluate the accuracy of the model through metrics such as R-squared (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The results show that the linear regression model explains 83% of the variation in tissue sales ($R^2 = 0.83$), demonstrating the model's effectiveness in capturing the relationship between variables. However, the RMSE value of 78.34 and MAE of 56.69 indicate significant prediction errors. Therefore, it is recommended to add additional predictor variables, such as seasonal factors or promotions, and to use more complex models, such as non-linear regression or Random Forest, for more

accurate results. This study concludes that linear regression is a useful baseline method, but further model development and more comprehensive data are needed to improve prediction accuracy.

1. PENDAHULUAN

Penelitian ini dimulai dengan latar belakang yang menjelaskan bagaimana perkembangan teknologi informasi telah membawa dampak signifikan di berbagai sektor, termasuk bisnis. Dalam era modern, kemampuan analisis data menjadi salah satu kebutuhan utama untuk mendukung pengambilan keputusan strategis, terutama dalam memprediksi penjualan [1]. *Regresi linear* dipilih sebagai metode karena sifatnya yang sederhana namun efektif dalam menganalisis hubungan antara variabel-variabel [2].

Pemilihan objek penelitian, yaitu PT Panjunan, didasarkan pada relevansinya sebagai perusahaan kecil dan menengah (UKM) yang menghadapi tantangan dalam mengelola data penjualan dan memanfaatkan teknik prediktif untuk mendukung strategi operasional dan keuangan [3]. UKM seperti PT Panjunan sering kali terbatas dalam hal sumber daya, baik teknologi maupun finansial, yang membuat kebutuhan akan model prediksi yang sederhana namun akurat menjadi sangat penting [4].

Dalam penelitian ini mencakup kesenjangan antara kebutuhan UKM untuk memprediksi penjualan secara akurat dan keterbatasan alat atau model yang tersedia [5]. Ketidakakuratan dalam memprediksi penjualan dapat berdampak pada pengambilan keputusan yang salah, yang pada akhirnya memengaruhi kesehatan finansial perusahaan [6].

Tujuan penelitian adalah untuk mengembangkan model prediksi penjualan berbasis regresi linear yang dapat diterapkan pada kondisi spesifik PT Panjunan [7]. Model ini diharapkan mampu memberikan hasil prediksi yang akurat berdasarkan data historis perusahaan dan faktor-faktor ekonomi relevan, sehingga membantu perusahaan dalam perencanaan strategis [8].

Penelitian ini difokuskan pada penggunaan data historis dan variabel ekonomi yang relevan, dengan mengakui keterbatasan sumber daya PT Panjunan dalam hal pengelolaan data [9]. Penelitian ini juga hanya menggunakan regresi linear sebagai

pendekatan utama tanpa mengintegrasikan metode prediktif yang lebih kompleks [10].

Data pendukung yang digunakan dalam penelitian ini mencakup data penjualan historis dari PT Panjunan serta literatur tentang metode prediksi berbasis regresi linear [11]. Kesenjangan analisis (*gap analysis*) terlihat dari kondisi realita bahwa banyak UKM kesulitan menggunakan metode prediksi yang canggih, sementara kebutuhan prediksi akurat semakin mendesak dalam menghadapi ketidakpastian ekonomi [12].

Penelitian ini dibandingkan dengan studi sebelumnya, seperti penelitian oleh [2] yang menunjukkan regresi linear efektif tetapi menghadapi tantangan dalam data dengan variabilitas tinggi [2]. Penelitian oleh [9] juga menunjukkan keberhasilan *regresi linear* dalam prediksi penjualan, namun tidak mencakup pengujian asumsi model yang dapat memengaruhi validitas hasilnya [9]. Studi ini berfokus pada adaptasi model *regresi linear* untuk kondisi spesifik UKM dengan data terbatas dan tidak terstruktur, sehingga menawarkan kontribusi baru dalam penerapan metode prediksi sederhana namun efektif [13].

2. TINJAUAN PUSTAKA

Regresi linear sederhana digunakan dalam data mining untuk menganalisis hubungan antara variabel bebas dan terikat guna memprediksi nilai di masa depan. Dalam bisnis, metode ini membantu memproyeksikan omset berdasarkan faktor seperti harga, volume penjualan, dan tanggal. Melalui pendekatan Knowledge Discovery in Database (KDD), proses analisis dilakukan secara sistematis untuk mengoptimalkan strategi bisnis dan pengambilan keputusan.[14]

Penelitian ini menggunakan beberapa ukuran kebaikan model seperti Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), dan R-squared (R^2) untuk menilai performa kedua metode tersebut dalam berbagai proporsi data uji. Temuan ini menunjukkan bahwa regresi linear berganda cenderung memberikan prediksi yang lebih akurat dibandingkan dengan regresi decision tree dalam

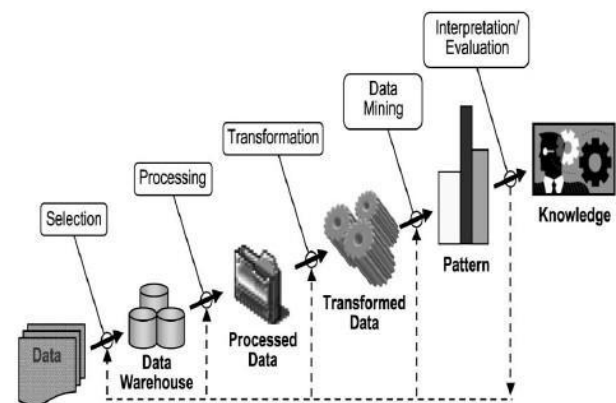
konteks dataset yang digunakan dalam penelitian ini. [13]

Focus artikel ini adalah mengestimasi laju pertumbuhan penduduk di Kabupaten Simalungun menggunakan metode regresi linier berganda. Penelitian ini dilakukan untuk membantu Badan Pusat Statistik (BPS) Simalungun dalam memprediksi pertumbuhan penduduk, yang penting untuk berbagai kebutuhan strategis, termasuk ekonomi dan infrastruktur. Temuan dari penelitian ini menunjukkan bahwa metode regresi linier berganda dapat digunakan untuk memprediksi laju pertumbuhan penduduk di BPS Simalungun dengan menggunakan data jumlah penduduk laki-laki dan perempuan sebagai variabel bebas. Hasilnya memberikan informasi estimasi yang dapat digunakan untuk perencanaan di masa mendatang. [4]

Studi ini dilakukan di Program Studi Manajemen Fakultas Ekonomi dan Bisnis, Universitas Bina Bangsa, dengan data selama lima tahun dari tahun 2015 hingga 2019. Model regresi linear sederhana yang dibentuk dapat digunakan untuk memprediksi jumlah penerimaan mahasiswa baru di masa mendatang dengan tingkat error sebesar 3,444%, atau tingkat akurasi 96,556%. [6]

Masalah yang diangkat dalam penelitian ini adalah harga beras di Indonesia yang terus naik sejak Agustus 2023, yang berdampak signifikan pada tingkat inflasi, stabilitas ekonomi, dan tingkat kemiskinan. Oleh karena itu, diperlukan metode yang akurat untuk memprediksi harga beras guna membantu pemerintah dalam pengambilan keputusan. Metode penelitian yang digunakan adalah regresi linear dalam kerangka SEMMA (Sample, Explore, Modify, Model, Assess). Data yang digunakan adalah data harga beras dari tahun 2022-2023 yang diperoleh dari PHIPS (Pusat Informasi Harga Pangan Strategis). Data ini dibagi menjadi 70% untuk pelatihan dan 30% untuk pengujian, dengan akurasi model diukur menggunakan RMSE. Penelitian ini memberikan model yang dapat digunakan untuk memprediksi harga beras dengan akurat, yang bermanfaat bagi pemerintah dalam mengontrol harga beras dan membuat kebijakan yang tepat. Selain itu, metode yang digunakan dalam penelitian ini dapat diaplikasikan pada prediksi harga komoditas lainnya, sehingga memberikan manfaat lebih luas dalam bidang prediksi harga komoditas. [15]

3. METODE PENELITIAN



Gambar 1. Kerangka kerja KDD

Penelitian ini menggunakan kerangka kerja KDD (*Knowledge Discovery in Databases*) untuk analisis data dan pengembangan model. KDD adalah serangkaian langkah yang digunakan untuk mengekstraksi pengetahuan atau informasi berharga dari data yang besar, kompleks, dan sering kali tidak terstruktur.

menggambarkan alur penelitian berdasarkan kerangka kerja KDD, yang terdiri dari beberapa tahapan utama. Setiap tahapan akan dijelaskan secara rinci untuk memberikan pemahaman yang komprehensif mengenai proses penelitian, mulai dari perencanaan hingga pelaksanaan dan analisis hasil. Pemahaman terhadap tahapan-tahapan ini akan membantu menilai metodologi yang digunakan serta validitas dari kesimpulan yang dihasilkan.

2.1 Metode pengumpulan data, instrumen penelitian, dan metode pengujian

Penelitian ini menggunakan teknik observasi langsung untuk mengumpulkan data penjualan di PT Panjunan. Data yang dikumpulkan meliputi volume penjualan, nama toko, dan waktu transaksi selama periode 2022-2023. Observasi dilakukan secara sistematis untuk memastikan data yang diperoleh akurat dan representatif. Informasi ini digunakan sebagai dasar dalam analisis dan pembuatan model prediksi penjualan.

Tahapan penelitian

a. Data Selection

Data penjualan harian diubah menjadi rata-rata penjualan tahunan menggunakan Microsoft Excel untuk menyederhanakan analisis dan mengidentifikasi pola jangka panjang.

b. Data Preprocessing

Data penjualan tahunan dibersihkan dengan menghapus nilai kosong dan *outlier* menggunakan *Python* dan pustaka *pandas* untuk meningkatkan kualitas data yang akan dianalisis.

c. Data Transformation

Data yang sudah dibersihkan disiapkan untuk analisis lebih lanjut dengan membagi dataset menjadi data pelatihan dan pengujian, serta membangun model regresi linear untuk memprediksi penjualan.

d. Data Mining

Model regresi linear dibangun menggunakan data Tahun2022 untuk memprediksi penjualan Tahun2023, kemudian dilatih dan diuji dengan membagi data menjadi data pelatihan dan pengujian.

e. Interpretation and Evaluation

Hasil prediksi model dievaluasi dengan membandingkan nilai prediksi dengan nilai asli, serta menggunakan metrik seperti R^2 , MSE, RMSE, dan MAE untuk mengukur akurasi model.

4. HASIL DAN PEMBAHASAN

1. Data Selection

Tahap pertama adalah *data selection*, yang bertujuan untuk mengorganisasikan dan menyeleksi data transaksi yang relevan guna menyederhanakan analisis. Dalam penelitian ini, data penjualan harian diubah menjadi rata-rata penjualan tahunan menggunakan Microsoft Excel, sehingga mempermudah identifikasi pola penjualan jangka panjang. Transformasi dilakukan dengan memanfaatkan fitur *PivotTable*, di mana data diatur ulang dengan kolom *Year* sebagai *Columns*, *NAMA_BARANG* sebagai *Rows*, dan *sales* sebagai *Values*. Fungsi agregasi diubah menjadi *Average* untuk menghitung rata-rata penjualan tahunan yang representatif [14]. Langkah ini penting untuk memastikan data yang digunakan lebih bersih dan terstruktur, sehingga meminimalkan potensi kesalahan analisis [1]

1	Average of TOTAL_PENJUALAN	Column Labels	
2	Row Labels	2022	2023
3	MULTI ECO FACIAL 900 GRAM 16BGx900G	9,76	45
4	MULTI ECO NAPKIN COCKTAIL 100S 1X40	200	102,5714286
5	MULTI ECO TOILET 150S X 10R NC 1X16	30,44827586	22,83333333
6	MULTI ECO TOWEL INTERFOLD 150S 1X20	77,20588235	57,85714286
7	MULTI FACIAL 900 GRAM 16BGx900G	5,818181818	5,5
8	MULTI FACIAL SOFTPACK 40X200S	6	
9	MULTI NAPKIN LUNCH 100S MU 10X6 MNO	70,8	8
10	MULTI TOILET 180S X 10R 1X10 PB09	22,16393443	19,66666667
11	MULTI TOILET 220S X 10R 1X10 PB10	8	23
12	PB-24 M TOILET BUBBLE 50BAG		82
100	TISSUE TESSA TOILET 110G 1X100 TOT-	57,39575972	50,93006993
101	TISSUE YOA FACIAL RF 650G 1X16 KF-8		3
102	TISSUE YOA FACIAL SP 250S 1X40 KF-7	218,7948718	81,89473684
103	TMUR 0012-108 TESSA MAKE UP REMOVER		25
104	TWWG 0012-072 TESSA WET WIPES GENER		3
105	YOA PREMIUM 1 ROL		4
106	YOA PREMIUM 300S		46,5
107	YOA PREMIUM TRAVEL PAK	4	
108	YOA TISSU 2 ROL		20
109	YOA TISSU 6 ROL		21

Gambar 2. Hasil data selection

2. Data Preprocessing

Pada tahap *data preprocessing*, data dibersihkan dari nilai kosong (*missing values*) dan *outlier*. Data penjualan tahunan diimpor menggunakan *Python* dan *library pandas*, yang memungkinkan pengolahan data dalam bentuk tabel secara efisien. Fungsi *df.info()* digunakan untuk memeriksa struktur data, mengidentifikasi nilai kosong pada kolom Tahun2022 dan Tahun2023.

```
import pandas as pd
df = pd.read_excel('E:/TUGAS/skripsi/datasetskripsi.xlsx')
print(df)
df.info()
```

	TipeTisu	Tahun2022	Tahun2023
0	MULTI ECO FACIAL 900 GRAM 16BGx900G	9.760000	45.000000
1	MULTI ECO NAPKIN COCKTAIL 100S 1X40	200.000000	102.571429
2	MULTI ECO TOILET 150S X 10R NC 1X16	30.448276	22.833333
3	MULTI ECO TOWEL INTERFOLD 150S 1X20	77.205882	57.857143
4	MULTI FACIAL 900 GRAM 16BGx900G	5.818182	5.500000
...	...	...	...
102	YOA PREMIUM 1 ROL	NaN	4.000000
103	YOA PREMIUM 300S	NaN	46.500000
104	YOA PREMIUM TRAVEL PAK	4.000000	NaN
105	YOA TISSU 2 ROL	NaN	20.000000
106	YOA TISSU 6 ROL	NaN	21.000000

```
[107 rows x 3 columns]
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 107 entries, 0 to 106
Data columns (total 3 columns):
#   Column      Non-Null Count  Dtype
---  -
0   TipeTisu    107 non-null    object
1   Tahun2022   84 non-null     float64
2   Tahun2023   93 non-null     float64
dtypes: float64(2), object(1)
memory usage: 2.6+ KB
```

Gambar 3. Hasil *df.info()*

Kemudian menghapusnya dengan menggunakan fungsi *dropna()*. Pembersihan data juga mencakup penghapusan *outlier*, yang dilakukan dengan menetapkan batas bawah dan batas atas berdasarkan kuantil ke-1% dan ke-99% untuk memastikan

distribusi data yang lebih seragam [2]. Langkah-langkah ini penting untuk meningkatkan kualitas data yang akan digunakan dalam model analisis.

```
lower_bound = df['Tahun2022'].quantile(0.01)
upper_bound = df['Tahun2023'].quantile(0.99)

df = df[(df['Tahun2023'] >= lower_bound) & (df['Tahun2023'] <= upper_bound)]
print(df)

df.dropna(inplace=True)
print(df)
```

Gambar 4. Menghapus missing value dan outlier

3. Data Transformation

Pada tahap *data transformation*, data yang telah diproses kemudian dipersiapkan untuk analisis lebih lanjut. Data penjualan yang sudah dibersihkan selanjutnya digunakan untuk membangun model regresi linear. Pada tahap ini, variabel *Tahun2022* digunakan sebagai fitur (*X*) dan *Tahun2023* sebagai target (*y*), lalu data dibagi menjadi data pelatihan (80%) dan data pengujian (20%) menggunakan *train_test_split* dengan parameter *random_state=42* untuk konsistensi. Model regresi linear dibangun menggunakan *LinearRegression()* dari Scikit-Learn dan dilatih menggunakan data pelatihan, lalu diuji untuk menghasilkan prediksi [15].

```
X = df[['Tahun2022']]
y = df['Tahun2023']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Gambar 5. Membagi data pelatihan dan pengujian

5. Data Mining

Pada tahap *data mining*, dilakukan pembangunan model prediksi menggunakan regresi linear untuk memprediksi data *Tahun2023* berdasarkan data *Tahun2022*. Proses dimulai dengan impor data yang telah diproses dan dibersihkan, serta pemisahan data menjadi dua bagian: fitur (*X*) yang merupakan data penjualan *Tahun2022*, dan target (*y*) yang merupakan data penjualan *Tahun2023*. Data dibagi menjadi dua subset: 80% untuk pelatihan dan 20% untuk pengujian, menggunakan fungsi *train_test_split* dengan parameter *random_state=42* agar hasilnya konsisten. Model regresi linear dibangun menggunakan algoritma *LinearRegression()* dari Scikit-Learn dan kemudian dilatih menggunakan data pelatihan. Setelah model siap, dilakukan pengujian dengan data pengujian untuk menghasilkan prediksi. Model ini kemudian menghasilkan hasil prediksi yang disimpan dalam variabel *y_pred*, yang akan digunakan untuk

mengevaluasi performa model dalam memprediksi data yang tidak terlihat sebelumnya [16].

```
model = LinearRegression()

model.fit(X_train, y_train)

y_pred = model.predict(X_test)
```

Gambar 6. Membuat model regresi linear

4. Interpretation and Evaluation

Pada tahap interpretasi dan evaluasi, hasil prediksi model dibandingkan dengan nilai asli untuk menilai akurasi model. Metrik seperti *R-squared* (R^2), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Error* (MAE) digunakan untuk mengevaluasi kinerja model. Hasilnya menunjukkan bahwa model memiliki R^2 yang cukup tinggi, mengindikasikan bahwa model dapat menjelaskan sebagian besar variasi data. Meskipun nilai MSE, RMSE, dan MAE menunjukkan adanya kesalahan, margin kesalahan tersebut masih dalam batas yang dapat diterima, namun perbaikan lebih lanjut diperlukan untuk meningkatkan akurasi prediksi [5][3].

Tabel 1. Hasil Evaluasi

R^2	MSE	RMSE	MAE
0,83	6132,3	78,30	56,76

a. *R-squared* (R^2)

R-squared, atau koefisien determinasi, adalah metrik yang menunjukkan seberapa baik variabel independen menjelaskan variabilitas variabel dependen dalam model regresi. *R-squared* berkisar antara 0 hingga 1. Semakin mendekati 1, semakin baik model menjelaskan data. Nilai $R^2 = 0,83$ (atau 83%) dalam model ini berarti bahwa 83% dari variabilitas dalam variabel dependen dapat dijelaskan oleh model, sedangkan sisanya 17% disebabkan oleh faktor-faktor lain atau error yang tidak dapat dijelaskan oleh model.

b. *Mean Squared Error* (MSE)

Mean Squared Error (MSE) adalah metrik yang menunjukkan rata-rata dari kuadrat kesalahan antara

nilai prediksi dan nilai actual. Kuadrat dari selisih ini berguna untuk menghukum kesalahan besar sehingga model akan lebih sensitif terhadap kesalahan prediksi yang signifikan. Pada model ini, nilai MSE adalah 6132,3, yang menunjukkan jumlah rata-rata kesalahan dalam bentuk kuadrat. MSE memberikan gambaran umum tentang seberapa besar kesalahan model, tetapi nilainya bisa sulit diinterpretasikan secara langsung karena masih dalam bentuk kuadrat dari satuan data.

c. *Root Mean Squared Error (RMSE)*

Root Mean Squared Error (RMSE) adalah akar kuadrat dari MSE. RMSE mengembalikan metrik kesalahan ke satuan yang sama dengan data asli, sehingga lebih mudah diinterpretasikan dibandingkan MSE. Pada model ini, nilai RMSE sebesar 78,30 menunjukkan bahwa, secara rata-rata, prediksi model memiliki kesalahan sebesar 78,34 unit dari nilai aktual.

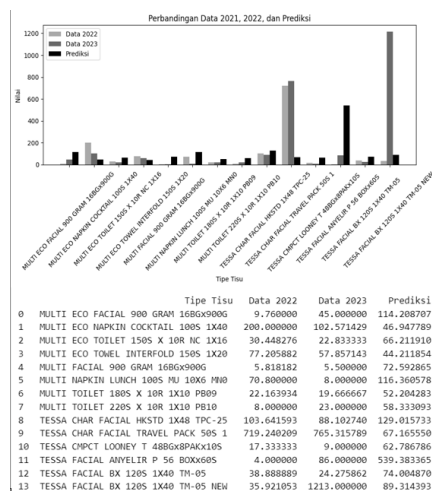
d. *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) adalah rata-rata dari nilai absolut kesalahan antara prediksi Tidak seperti MSE yang menghukum kesalahan besar lebih keras karena menggunakan kuadrat, MAE menghitung rata-rata kesalahan absolut, sehingga sering dianggap lebih "robust" terhadap outlier. Pada model ini, nilai MAE adalah 56,76, yang berarti rata-rata kesalahan antara prediksi dan nilai aktual adalah sekitar 56,76 unit. MAE memberikan gambaran seberapa besar rata-rata kesalahan prediksi tanpa memperhitungkan arah kesalahan, sehingga lebih mudah untuk memahami rata-rata jarak absolut antara prediksi dan nilai aktual.

model regresi linear. Kolom "Data 2022" mewakili input (variabel independen) yang digunakan untuk membuat prediksi, sementara kolom "Data 2023" berisi nilai sebenarnya dari variabel target (variabel dependen) pada tahun tersebut. Kolom "Prediksi" menampilkan hasil prediksi yang dihasilkan oleh model berdasarkan data dari "Data 2023". Dari perbandingan ini, kita dapat melihat seberapa akurat model dalam memprediksi nilai, seperti yang terlihat pada baris pertama di mana nilai asli "Data 2023" adalah 45.000000, tetapi model memprediksi nilai sebesar 114.208707, yang menunjukkan perbedaan signifikan. Namun, pada baris lainnya, prediksi model lebih mendekati nilai asli. Tabel ini membantu dalam mengevaluasi seberapa baik akurasi model, serta memberi gambaran tentang kesalahan prediksi yang mungkin terjadi.

5. KESIMPULAN

Penelitian ini mengevaluasi performa model regresi linear dalam memprediksi penjualan tisu PT Panjuran, dengan hasil menunjukkan kemampuan model menjelaskan 83% variasi data penjualan ($R^2 = 0,83$). Meskipun cukup efektif dalam menangkap pola data, akurasi prediksi masih dapat ditingkatkan, terlihat dari nilai RMSE sebesar 78,3 dan MAE sebesar 56,76 yang mencerminkan margin kesalahan signifikan. Penelitian terkait metode regresi linier untuk prediksi menunjukkan hasil yang bervariasi dalam pengukuran performansi model. Sebagai contoh, penelitian oleh [12] menunjukkan bahwa regresi linier memberikan hasil prediksi yang baik untuk data penjualan harian di restoran Go Chicken Karawang. Hasil ini sejalan dengan temuan [9], yang menemukan prediksi yang akurat untuk penjualan goodie bag di PT Goodiebag Custom Indonesia, meskipun tanpa mempertimbangkan uji asumsi regresi. Sebaliknya, [8] menunjukkan akurasi lebih tinggi dalam memprediksi permohonan paten di Indonesia, yang mungkin dipengaruhi oleh kompleksitas data dan variabel yang digunakan. Sementara itu, [2] mencatat akurasi yang sangat baik untuk prediksi penjualan usaha mikro, kecil, dan menengah, yang menunjukkan bahwa regresi linier lebih cocok untuk dataset kecil dan terstruktur. Hasil-hasil ini mengindikasikan bahwa kinerja regresi linier sangat bergantung pada karakteristik data dan tujuan prediksi, di mana prediksi dengan akurasi tinggi cenderung melibatkan data yang bersifat linier dengan variabel independen yang relevan. Oleh karena itu, pemilihan dan penyesuaian model sesuai dengan karakteristik dataset sangat penting untuk memaksimalkan akurasi prediksi. Untuk meningkatkan akurasi, disarankan menambahkan variabel seperti harga promosi, hari libur, dan tren musiman, serta mempertimbangkan



Gambar 7. Hasil prediksi

Gambar di atas menunjukkan perbandingan antara data asli dan nilai prediksi yang dihasilkan oleh

penggunaan model prediksi yang lebih kompleks, seperti regresi non-linear, Random Forest, atau pembelajaran mesin lainnya, yang lebih efektif dalam menangkap pola data kompleks. Peningkatan ini diharapkan dapat menghasilkan prediksi yang lebih akurat guna mendukung perencanaan strategis perusahaan.

UCAPAN TERIMA KASIH

Penulis. mengucapkan terima kasih kepada Direktorat Pendidikan Tinggi, Departemen Pendidikan dan Kebudayaan Republik Indonesia yang telah memberikan dukungan finansial melalui Beasiswa KIP tahun 2021-2025. Penulis juga mengucapkan terima kasih kepada PT Panjuran yang telah menyediakan data penjualan serta kepada pihak-pihak yang telah memberikan bantuan dan dukungan selama penelitian ini berlangsung.

DAFTAR PUSTAKA

- [1] 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.
- [2] F. H. Hamdanah and D. Fitrihanah, "Analisis Performansi Algoritma Linear Regression dengan Generalized Linear Model untuk Prediksi Penjualan pada Usaha Mikra, Kecil, dan Menengah," *J. Nas. Pendidik. Tek. Inform.,* vol. 10, no. 1, p. 23, 2021, doi: 10.23887/janapati.v10i1.31035.
- [3] S. Lestari, "Analisis Algoritma Regresi Linear Sederhana dalam Memprediksi Tingkat Penjualan Album KPOP," *INSOLOGI J. Sains dan Teknol.,* vol. 2, no. 1, pp. 199–209, 2023, doi: 10.55123/insologi.v2i1.1692.
- [4] F. O. Lusiana, I. Fatma, and A. P. Windarto, "Estimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Simalungun," *J. Informatics Manag. Inf. Technol.,* vol. 1, no. 2, pp. 79–84, 2021, doi: 10.47065/jimat.v1i2.104.
- [5] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," *J. Appl. Informatics Comput.,* vol. 5, no. 1, pp. 95–102, 2021, doi: 10.30871/jaic.v5i1.3184.
- [6] A. A. Muhartini, O. Sahroni, S. D. Rahmawati, T. Febrianti, and I. Mahuda, "Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana," *J. Bayesian J. Ilm. Stat. dan Ekon.,* vol. 1, no. 1, pp. 17–23, 2021, doi: 10.46306/bay.v1i1.2.
- [7] S. Napitupulu and N. A. Siagian, "Prediksi Data Produksi Menggunakan Regresi Linear Sederhana," *JDMIS J. Data Min. Inf. Syst.,* vol. 1, no. 2, pp. 95–105, 2023, doi: 10.54259/jdmis.v1i2.1956.
- [8] D. Novianty, N. D. Palasara, and M. Qomaruddin, "Algoritma Regresi Linear pada Prediksi Permohonan Paten yang Terdaftar di Indonesia," *J. Sist. dan Teknol. Inf.,* vol. 9, no. 2, p. 81, 2021, doi: 10.26418/justin.v9i2.43664.
- [9] M. C. Anjeliyani, Fatia Fatimah, "PEMODELAN REGRESI LINEAR UNTUK MEMPREDIKSI NILAI PENJUALAN DI PT GOODIEBAG CUSTOM INDONESIA TANPA MEMPERTIMBANGKAN UJI ASUMSI," vol. 5, no. September, pp. 1–13, 2024.
- [10] N. Nurzilla, "Prediksi Pertumbuhan Tumor Kanker Payudara Menggunakan Model Regresi Linear Berbasis Machine Learning," *J. Artif. Intell. Appl.,* vol. 1, no. 1, pp. 28–35, 2024.
- [11] M. Panjaitan and S. N. Hutagalung, "Perhitungan Biaya Promosi Dan Volume Penjualan Pada Suatu Perusahaan Dengan Metode Analisis Regresi Linear Dan Korelasi," *Semin. Nas. Sains Teknol. Inf.,* vol. 108, no. 4, p. 103, 2019, [Online]. Available: <http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage%7C103>
- [12] A. M. Dewi, F. N. Azhar, and C. Rozikin, "Prediksi Penjualan Restoran Go Chicken Karawang Menggunakan Metode Linear Regresi," *Socius J. Penelit. Ilmu ...,* vol. 1, no. December, pp. 284–288, 2023, [Online]. Available: <https://ojs.daarulhuda.or.id/index.php/Socius/article/view/93>
- [13] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *J. Sains dan Edukasi Sains,* vol. 6, no. 1, pp. 34–43, 2023, doi: 10.24246/juses.v6i1p34-43.
- [14] R. Hermawan, N. Suarna, I. Ali, and D. Rohman, "OPTIMASI PREDIKSI OMSET PENJUALAN PADA PABRIK OLAHAN TAHU MENGGUNAKAN ALGORITMA REGRESI LINEAR," *J. Inform. dan Tek. Elektro Terap.,* vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5888.
- [15] M. A. Veri Arinal, "Penerapan Regresi Linear Untuk Prediksi Harga Beras Di Indonesia," *J.*

Sains dan Teknol., vol. 5, no. 1, pp. 341–346,
2023.