

PREDIKSI PENJUALAN PRODUK ELEKTRONIK SHOPEE MENGGUNAKAN REGRESI LINIER BERBASIS PYTHON

Sindi Amelia¹, Nining Rahaninsih², Irfan Ali³, Denni Pratama⁴, Umi Hayati⁵

Program Studi Teknik Informatika¹⁵
Program Studi Komputerisasi Akuntansi²⁴
Program Studi Komputerisasi Akuntansi³

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
snyamli913@gmail.com

(*) Corresponding Author : snyamli913@gmail.com

Published : 30 Mei 2026

Abstract— The growing digital transformation in the e-commerce sector has increased the need for reliable sales prediction to support strategic decision-making, particularly in inventory management, pricing, and promotional allocation. This study aims to develop a sales prediction model for electronic products on the Shopee marketplace using the linear regression algorithm implemented in Python. The dataset consists of 80 weekly observations with variables such as number of orders, number of buyers, conversion rate, and sales value. The research stages include data acquisition and preprocessing, exploratory data analysis, splitting the data into training and testing sets (80:20), testing linearity and residual normality assumptions, building a simple linear regression model, and evaluating its performance using the coefficient of determination (R^2) and Mean Squared Error (MSE). The experimental results show that the number of orders is the main predictor, explaining approximately 79.8% of the variance in sales value, with a positive coefficient of around IDR 118,114.02 per order. Although the residual normality test indicates a violation of the normality assumption, the model's predictive performance is still considered adequate for practical purposes and can serve as an efficient and interpretable baseline model. These findings suggest that linear regression is suitable as a starting point for developing sales prediction systems for electronic product sellers on Shopee, before exploring more complex algorithms in future research.

Keywords: e-commerce, Shopee, electronic products, sales prediction, linear regression, Python

Abstrak— Digitalisasi di sektor e-commerce menuntut kemampuan prediksi penjualan yang andal untuk mendukung keputusan strategis, khususnya dalam pengelolaan stok, penetapan harga, dan alokasi promosi. Penelitian ini bertujuan mengembangkan model prediksi penjualan produk elektronik di marketplace Shopee menggunakan algoritma regresi linier berbasis Python. Dataset yang digunakan terdiri atas 80 observasi mingguan dengan variabel jumlah pesanan, jumlah pembeli, tingkat konversi, dan nilai penjualan. Tahapan penelitian meliputi akuisisi dan pra-pemrosesan data, eksplorasi data, pembagian data menjadi latih dan uji (80:20), uji asumsi linieritas dan normalitas residual, pembangunan model regresi linier sederhana, serta evaluasi kinerja model menggunakan koefisien determinasi (R^2) dan Mean Squared Error (MSE). Hasil eksperimen menunjukkan bahwa variabel jumlah pesanan merupakan prediktor utama yang mampu menjelaskan sekitar 79,8% variasi nilai penjualan, dengan koefisien positif sekitar Rp118.114,02 per pesanan. Meskipun uji normalitas residual mengindikasikan pelanggaran asumsi normalitas, performa prediktif model masih dinilai memadai untuk keperluan praktis dan dapat dijadikan model dasar yang efisien dan mudah diinterpretasikan. Temuan ini menunjukkan bahwa regresi linier layak dimanfaatkan sebagai titik awal pengembangan sistem prediksi penjualan bagi pelaku usaha produk elektronik di Shopee, sebelum dieksplorasi algoritma yang lebih kompleks pada penelitian lanjutan.

Kata Kunci : e-commerce, Shopee, produk elektronik, prediksi penjualan, regresi linier, Python

INTRODUCTION

Transformasi digital dalam sektor e-commerce telah mengubah cara perusahaan mengelola operasi

bisnis, mulai dari rantai pasok hingga strategi pemasaran. Dalam konteks ini, kemampuan memprediksi penjualan secara akurat menjadi elemen penting untuk menghindari risiko kelebihan stok (overstock) maupun kekurangan stok

(stockout), sekaligus mengoptimalkan penetapan harga dan alokasi anggaran promosi [1] Pada marketplace seperti Shopee, pola transaksi yang dinamis dan fluktuatif membuat prediksi penjualan tidak lagi sekadar fungsi pendukung, tetapi menjadi komponen strategis bagi keberlangsungan dan daya saing bisnis.

Produk elektronik merupakan salah satu kategori dengan tingkat persaingan tinggi dan siklus hidup relatif cepat. Perubahan tren teknologi, variasi harga, serta pengaruh promosi massal (seperti kampanye 11.11 atau 12.12) menyebabkan pola permintaan menjadi sulit diperkirakan hanya dengan intuisi. Oleh karena itu, pemanfaatan pendekatan data-driven melalui model statistik dan machine learning diperlukan untuk mengungkap pola historis yang tersembunyi dalam data penjualan, seperti tren musiman, pengaruh promosi, dan respon pelanggan terhadap perubahan strategi pemasaran ([2])

Sejumlah penelitian terdahulu menunjukkan bahwa pemodelan prediksi penjualan di ranah ritel dan e-commerce telah banyak dilakukan, namun sebagian besar berfokus pada konteks negara maju, kategori produk tertentu, atau platform selain Shopee [3] Selain itu, masih terdapat keterbatasan dalam pemanfaatan model yang sederhana namun mudah diimplementasikan oleh pelaku usaha kecil dan menengah (UMKM). Di Indonesia, UMKM sering kali menghadapi kendala sumber daya teknis, sehingga model yang transparan, ringan, dan tidak membutuhkan infrastruktur komputasi besar menjadi lebih relevan dibandingkan model kompleks yang sulit diinterpretasikan.

Dalam konteks tersebut, regresi linier menempati posisi penting sebagai salah satu metode dasar yang banyak digunakan untuk pemodelan hubungan antara variabel bebas dan terikat. Metode ini memiliki keunggulan pada sisi interpretabilitas, kemudahan implementasi, serta tersedianya perangkat diagnostik seperti koefisien determinasi (R^2) dan analisis residual [4], [5]. Bagi pelaku bisnis, kemudahan dalam memahami bagaimana perubahan jumlah pesanan memengaruhi nilai penjualan menjadi lebih penting daripada sekadar mengejar akurasi prediksi tertinggi tanpa penjelasan yang jelas.

Di Indonesia, pengembangan e-commerce juga tidak lepas dari peran UMKM yang memanfaatkan marketplace sebagai kanal pemasaran utama. Berbagai studi menyoroti bahwa adopsi teknologi analitik dan sistem prediksi masih relatif terbatas,

baik karena keterbatasan literasi data maupun kurangnya contoh implementasi yang kontekstual [6] Dengan demikian, pengembangan model prediksi penjualan produk elektronik di Shopee berbasis regresi linier dan diimplementasikan menggunakan Python menjadi relevan, baik dari sisi ilmiah maupun praktis.

Berdasarkan uraian tersebut, permasalahan utama dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana merancang model regresi linier untuk memprediksi penjualan produk elektronik pada platform Shopee berdasarkan data historis penjualan?
2. Bagaimana proses implementasi dan eksperimen model prediksi menggunakan bahasa pemrograman Python pada dataset penjualan produk elektronik di Shopee?
3. Sejauh mana kinerja model regresi linier yang dibangun dalam menjelaskan variansi penjualan dan menghasilkan prediksi yang cukup akurat untuk mendukung pengambilan keputusan bisnis?

Sejalan dengan rumusan masalah di atas, tujuan penelitian ini adalah:

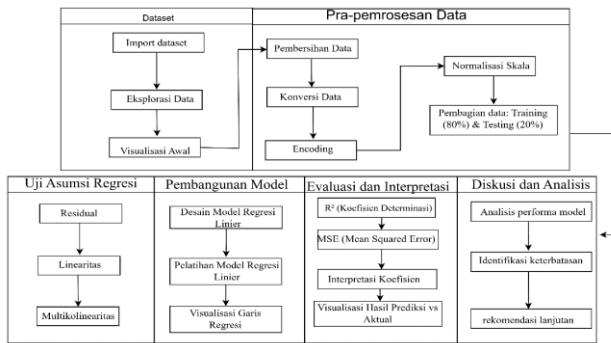
1. Merancang model regresi linier untuk memodelkan hubungan antara variabel penjelas (misalnya jumlah pesanan, jumlah pembeli, dan tingkat konversi) dengan nilai penjualan produk elektronik di Shopee.
2. Mengimplementasikan model tersebut menggunakan Python dan library data science seperti pandas, scikit-learn, dan matplotlib dengan alur eksperimen yang terstruktur.
3. Mengevaluasi performa model dengan metrik seperti koefisien determinasi (R^2) dan Mean Squared Error (MSE), serta menganalisis kelayakan model sebagai alat bantu prediksi penjualan.

Kontribusi penelitian ini mencakup dua aspek. Pertama, secara ilmiah, penelitian ini menghadirkan studi kasus penerapan regresi linier pada data penjualan produk elektronik di marketplace Shopee dengan tahapan yang replikatif dan dapat dijadikan acuan penelitian lanjutan. Kedua, secara praktis, hasil penelitian diharapkan dapat membantu pelaku usaha, khususnya penjual produk elektronik di Shopee, untuk memanfaatkan model prediktif sederhana sebagai dasar perencanaan stok,

penentuan target penjualan, dan pengelolaan strategi promosi.

MATERIALS AND METHODS

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian prediktif. Fokus utama penelitian adalah membangun model regresi linier yang mampu memprediksi nilai penjualan produk elektronik berdasarkan variabel historis yang tersedia dalam dataset. Model dikembangkan dan diuji menggunakan data penjualan mingguan dari marketplace Shopee, dengan tujuan menghasilkan model dasar yang sederhana, transparan, dan dapat direplikasi.



Gambar 1 Tahapan Alur Penelitian

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data historis penjualan produk elektronik pada marketplace Shopee dengan satuan observasi mingguan. Dataset terdiri atas 80 observasi yang merepresentasikan kinerja penjualan selama periode tertentu. Variabel yang digunakan meliputi:

Tabel 1. Data Set

Time Series	Tahun	Penjualan (Rp)	Pesanan	Pembeli	Konversi (%)
1 jan - 7 jan	2024	208.000	3	3	2.34%
8 jan- 14 jan	2024	158.000	2	2	1.38%
5 feb- 11 feb	2024	376000	3	3	1.69%
12 feb - 18 feb	2024	243.000	3	3	2.50%

1. Penjualan (Rp): nilai penjualan per minggu (variabel dependen/target).
2. Pesanan: jumlah pesanan yang berhasil tercatat.
3. Pembeli: jumlah akun pembeli yang melakukan transaksi.
4. Tingkat Konversi: perbandingan antara jumlah pengunjung dan pembeli, dinyatakan dalam persentase.
5. (Variabel waktu seperti tahun atau minggu dapat digunakan sebagai konteks tambahan dalam eksplorasi data.)

Secara konseptual, variabel Pesanan, Pembeli, dan Tingkat Konversi berperan sebagai variabel penjelas yang diharapkan memengaruhi nilai Penjualan. Ringkasan definisi dan jenis variabel dapat ditampilkan dalam tabel deskriptif pada Praproses Data

Data diperoleh melalui proses pengambilan (scraping) dan/atau ekspor laporan penjualan dari Shopee dalam format spreadsheet (misalnya .xlsx). Berkas tersebut kemudian diimpor ke lingkungan Python menggunakan library *pandas*:

```
df = pd.read_excel("/content/penjualan_shopee.xlsx")
df.head()
```

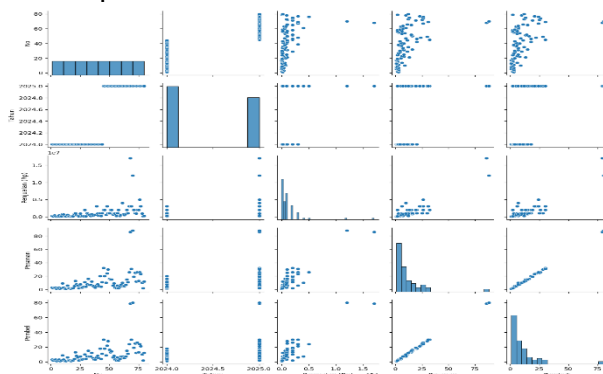
Langkah pra-pemrosesan dilakukan untuk memastikan kualitas dan konsistensi data sebelum pemodelan, meliputi:

- a) Pemeriksaan kelengkapan data Menggunakan fungsi `df.info()` dan `df.isna().sum()` untuk memastikan tidak terdapat nilai hilang (missing values). Apabila terdapat nilai hilang, dapat dilakukan penghapusan baris/kolom tertentu atau imputasi sesuai konteks.
- b) Pemeriksaan statistik deskriptif Fungsi `df.describe()` digunakan untuk menilai rentang, rata-rata, dan sebaran nilai masing-masing variabel. Langkah ini membantu mengidentifikasi potensi outlier yang terlalu ekstrem.
- c) Transformasi tipe data Kolom tingkat konversi yang semula dalam format persentase string (misalnya "2,50%") dikonversi ke nilai numerik bertipe float (misalnya 0.025). Hal ini diperlukan agar variabel tersebut dapat digunakan dalam perhitungan matematis dan pemodelan regresi.
- d) Seleksi fitur Dari beberapa variabel yang tersedia, penelitian ini memfokuskan pemodelan awal pada hubungan antara variabel jumlah pesanan dan nilai penjualan, sehingga model yang dikembangkan berbentuk regresi linier sederhana. Namun, pendekatan ini tetap dapat diperluas menjadi regresi linier berganda dengan menambahkan

variabel lain seperti pembeli dan tingkat konversi.

Eksplorasi Data (EDA)

Exploratory Data Analysis (EDA) dilakukan untuk memahami pola dasar dan hubungan antar variabel. Salah satu bentuk visualisasi yang digunakan adalah *pairplot* yang menampilkan distribusi masing-masing variabel serta relasi dua variabel secara bersamaan.



Gambar .2.2 Explorasi Data Analys

Dari visualisasi tersebut terlihat bahwa:

1. Variabel Pesanan dan Pembeli memiliki korelasi sangat kuat dengan pola mendekati garis lurus, yang menunjukkan bahwa hampir setiap pesanan berasal dari pembeli unik (rasio mendekati 1:1).
2. Terdapat korelasi positif antara Pesanan dan Penjualan, meskipun penyebaran titik relatif lebih lebar, menandakan bahwa peningkatan jumlah pesanan secara umum diikuti peningkatan nilai penjualan, namun tidak selalu proporsional karena dipengaruhi nilai per pesanan.
3. Distribusi penjualan cenderung miring ke kanan (right-skewed), mencerminkan banyak periode dengan penjualan relatif rendah dan beberapa periode dengan penjualan sangat tinggi.

Pembangunan Model Regresi Linier

Model regresi linier dibangun dengan menjadikan variabel Penjualan (Y) sebagai variabel dependen dan variabel Pesanan (X) sebagai variabel independen utama. Secara matematis, model dapat ditulis sebagai:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

di mana Y adalah nilai penjualan, X adalah jumlah pesanan, β_0 adalah intercept, β_1 adalah koefisien regresi, dan ϵ adalah galat (error). Implementasi

model menggunakan library *scikit-learn* dengan tahapan:

1. Menentukan fitur dan target:

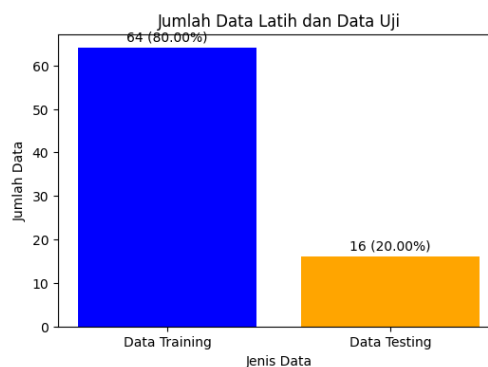
```
X = df[['Pesanan']]
y = df['Penjualan']
```

2. Membagi data menjadi data latih dan data uji dengan rasio 80:20:

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

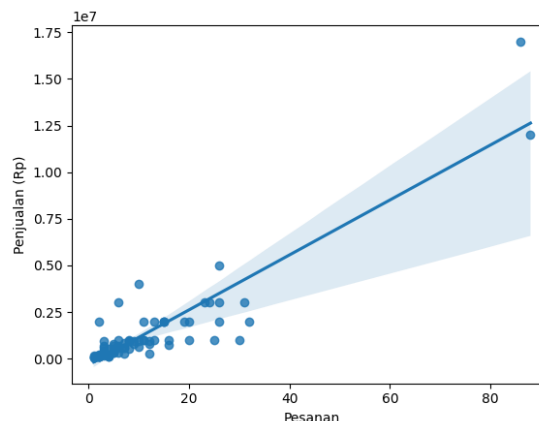
3. Melatih model:

```
model = LinearRegression()
model.fit(X_train, y_train)
```



Gambar 3 Data Train dan Data Test

Uji Asumsi Klasik (Linieritas dan Normalitas Residual) Untuk memastikan kelayakan regresi linier, dilakukan pemeriksaan asumsi dasar: Asumsi linieritas dievaluasi melalui grafik sebar antara jumlah pesanan dan nilai penjualan menggunakan regplot. Hasilnya menunjukkan pola titik yang mengikuti garis lurus, sehingga hubungan keduanya dapat dianggap mendekati linier.



Gambar 4 Uji Linieritas

Asumsi normalitas residual diperiksa menggunakan uji Shapiro-Wilk. Hasil pengujian menunjukkan nilai $p < 0,05$, yang mengindikasikan bahwa residual tidak berdistribusi normal secara statistik. Meskipun demikian, dalam konteks data riil dan ukuran sampel yang terbatas, regresi linier tetap cukup robust sehingga model masih dapat dimanfaatkan dengan catatan bahwa pelanggaran asumsi ini perlu dilaporkan dan dipertimbangkan dalam interpretasi hasil.

RESULTS AND DISCUSSION**Deskripsi Dataset dan Hasil EDA**

Dataset yang digunakan terdiri atas 80 observasi mingguan dengan variasi nilai penjualan, jumlah pesanan, jumlah pembeli, dan tingkat konversi. Secara umum, nilai penjualan menunjukkan fluktuasi yang cukup tajam antara minggu, dengan beberapa periode puncak penjualan dan beberapa periode penurunan signifikan. Dalam contoh potongan data awal, terlihat bahwa penjualan tertinggi dalam deret awal mencapai sekitar Rp 376.000, sementara penjualan terendah berada pada kisaran Rp 55.000, dengan tingkat konversi yang juga berfluktuasi.

Hasil EDA menunjukkan bahwa:

1. Korelasi antara Pesanan dan Pembeli sangat tinggi, sehingga keduanya pada dasarnya membawa informasi yang mirip.
2. Korelasi positif antara Pesanan dan Penjualan menguatkan hipotesis bahwa jumlah pesanan merupakan prediktor yang baik bagi nilai penjualan.
3. Tingkat konversi memperlihatkan variasi yang tidak selalu sejalan dengan penjualan, mengindikasikan adanya faktor lain seperti strategi promosi, tampilan etalase, atau faktor eksternal lain yang tidak terekam dalam dataset.

Hasil Pembangunan Model Regresi Linier

Setelah model dilatih menggunakan data latih, diperoleh persamaan regresi linier sederhana antara jumlah pesanan dan nilai penjualan. Koefisien regresi yang diperoleh menunjukkan bahwa setiap tambahan satu pesanan berasosiasi dengan kenaikan nilai penjualan rata-rata sekitar Rp118.114,02. Nilai koefisien yang positif ini

konsisten dengan ekspektasi bahwa meningkatnya jumlah pesanan akan meningkatkan nilai penjualan.

Secara formal, model dapat dinyatakan sebagai:

$$Y = \beta_0 + 118.114,02 \times X$$

dengan merupakan nilai penjualan prediksi dan X jumlah pesanan per minggu. Nilai intercept β_0 menggambarkan nilai penjualan ketika jumlah pesanan bernilai nol, meskipun secara praktis nilai ini lebih bersifat parameter model daripada interpretasi bisnis langsung.

Hasil uji linieritas memperkuat bahwa hubungan antara pesanan dan penjualan bersifat linier, sebagaimana terlihat pada pola sebaran titik yang mengikuti garis regresi

Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan data uji (20% dari total observasi). Proses evaluasi mencakup:

1. Prediksi pada data uji

Nilai prediksi diperoleh menggunakan perintah:

```
y_pred = model.predict(X_test)
```

2. Perhitungan metrik R² dan MSE

Kinerja model diukur menggunakan koefisien determinasi (R²) dan Mean Squared Error (MSE):

```
print("R2 Score:", r2_score(y_test, y_pred))
```

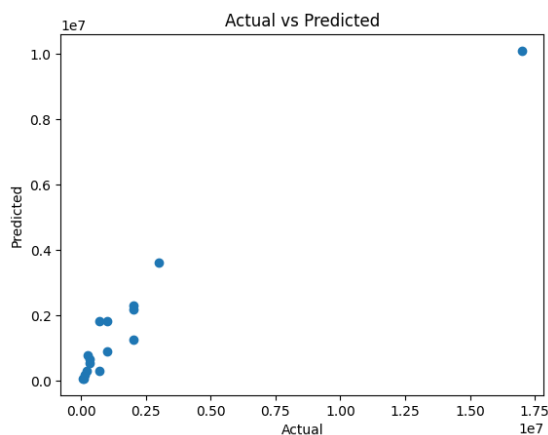
```
print("MSE:", mean_squared_error(y_test, y_pred))
```

Hasil menunjukkan nilai R² sekitar **0,798**, yang berarti sekitar **79,8% variansi nilai penjualan** pada data uji dapat dijelaskan oleh variabel jumlah pesanan dalam model. Nilai ini mengindikasikan bahwa model memiliki kemampuan yang cukup baik dalam menjelaskan hubungan antara pesanan dan penjualan dalam konteks dataset yang digunakan. Nilai MSE menggambarkan rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi; semakin kecil nilai MSE, semakin baik akurasi model. Dalam penelitian ini, nilai MSE dinilai masih dalam kisaran yang wajar jika dilihat dari skala satuan rupiah yang digunakan, sehingga kesalahan

prediksi yang terjadi masih dapat diterima secara praktis.

3. Visualisasi hasil prediksi

Hubungan antara nilai aktual dan nilai prediksi divisualisasikan melalui *scatter plot*, dengan sumbu X merepresentasikan nilai aktual dan sumbu Y



Gambar 5 predicted

Sebaran titik yang relatif dekat dengan garis diagonal imajiner menunjukkan bahwa prediksi model cukup sejalan dengan nilai aktual, meskipun terdapat beberapa titik yang menyimpang pada area penjualan tinggi. Pola ini lazim terjadi ketika data memiliki beberapa observasi ekstrem (outlier) yang sulit diprediksi secara tepat dengan model linier sederhana.

RESEARCH RESULT

Hasil penelitian menunjukkan bahwa regresi linier sederhana dengan variabel jumlah pesanan sebagai prediktor utama mampu memodelkan penjualan produk elektronik di Shopee dengan tingkat penjelasan variansi yang cukup tinggi ($R^2 \approx 0,798$). Hal ini sejalan dengan temuan sebelumnya bahwa variabel terkait volume transaksi merupakan indikator kuat bagi kinerja penjualan di sektor ritel dan e-commerce (Karl, 2024a; Ahaggach et al., 2024). Bagi pelaku usaha, informasi bahwa setiap tambahan satu pesanan berasosiasi dengan peningkatan penjualan rata-rata sekitar Rp118.114 memberikan gambaran kuantitatif yang mudah dipahami dan dapat dijadikan dasar perencanaan target penjualan.

Dari sisi metodologis, penggunaan regresi linier menawarkan sejumlah kelebihan. Model ini relatif ringan secara komputasi, mudah

diimplementasikan dengan Python, dan hasilnya dapat dijelaskan kepada pemangku kepentingan non-teknis. Hal ini menjadikannya pilihan yang relevan bagi UMKM yang ingin mulai mengadopsi analitik prediktif tanpa harus berinvestasi pada infrastruktur dan keahlian machine learning yang kompleks [6], [7] Namun demikian, penelitian ini juga memiliki beberapa keterbatasan. Pertama, model yang digunakan masih bersifat sederhana dengan hanya satu variabel penjelas utama, yaitu jumlah pesanan. Padahal, dalam praktik, penjualan produk elektronik sangat mungkin dipengaruhi faktor lain seperti harga, diskon, rating produk, jumlah ulasan, maupun faktor eksternal seperti kampanye promosi massal dan kondisi ekonomi makro (Brüggemann et al., 2025). Kedua, dataset yang digunakan memiliki ukuran terbatas (80 observasi mingguan), sehingga generalisasi model ke periode atau konteks lain perlu dilakukan dengan hati-hati.

Keterbatasan tersebut membuka peluang penelitian lanjutan, misalnya dengan mengembangkan model regresi linier berganda yang memasukkan variabel tambahan, atau dengan membandingkan kinerja regresi linier dengan algoritma lain seperti Random Forest, Gradient Boosting, atau XGBoost. Di samping itu, penggunaan teknik Explainable AI seperti SHAP atau LIME dapat membantu menjelaskan kontribusi masing-masing fitur dalam prediksi, sehingga meningkatkan kepercayaan pengguna terhadap model.

CONCLUSION

Penelitian ini bertujuan mengembangkan model prediksi penjualan produk elektronik di marketplace Shopee menggunakan algoritma regresi linier berbasis Python. Berdasarkan hasil eksperimen dan analisis, beberapa kesimpulan utama dapat dirangkum sebagai berikut.

1. model regresi linier sederhana dengan variabel jumlah pesanan sebagai prediktor utama mampu memodelkan hubungan dengan nilai penjualan secara cukup baik. Nilai koefisien regresi yang positif (sekitar Rp118.114,02 per pesanan) menunjukkan bahwa peningkatan jumlah pesanan berasosiasi langsung dengan peningkatan nilai penjualan.
2. hasil evaluasi model dengan menggunakan data uji menunjukkan nilai koefisien determinasi (R^2) sekitar 0,798. Artinya, sekitar 79,8% variansi penjualan pada data uji dapat dijelaskan oleh variasi jumlah

pesanan. Meskipun uji normalitas residual menunjukkan ketidaksesuaian dengan distribusi normal secara sempurna, performa prediktif model masih dapat dianggap memadai untuk keperluan praktis.

3. Ketiga, dari sisi implementasi, penggunaan Python dan library data science seperti pandas dan scikit-learn memungkinkan proses pemodelan dilakukan secara sistematis, mulai dari akuisisi data, eksplorasi, pra-pemrosesan, pelatihan model, hingga evaluasi kinerja. Alur ini dapat direplikasi oleh pelaku usaha maupun peneliti lain yang ingin mengembangkan model serupa.

Secara keseluruhan, regresi linier dapat diposisikan sebagai model dasar (baseline) yang efisien, transparan, dan mudah diimplementasikan untuk prediksi penjualan produk elektronik pada marketplace Shopee. Model ini dapat menjadi titik awal yang kuat sebelum beralih ke algoritma yang lebih kompleks jika dibutuhkan.

SUGGESTION

Berdasarkan temuan dan keterbatasan penelitian, beberapa saran yang dapat diajukan adalah:

1. Pengayaan variabel Penelitian selanjutnya disarankan menambahkan variabel lain seperti harga, diskon, rating produk, jumlah ulasan, kategori produk, dan indikator promosi (misalnya partisipasi dalam flash sale) agar model dapat menangkap variasi penjualan secara lebih komprehensif.
2. Perluasan cakupan data Dataset dapat diperluas baik dari sisi cakupan waktu (periode yang lebih panjang) maupun jumlah produk dan kategori. Perluasan ini memungkinkan pengujian model pada konteks yang lebih beragam, termasuk perbandingan antara produk elektronik dan kategori lainnya.
3. Perbandingan dengan algoritma lain Untuk memperoleh gambaran lebih lengkap tentang kinerja model, penelitian lanjutan dapat membandingkan regresi linier dengan algoritma lain seperti Random Forest, XGBoost, atau model berbasis jaringan saraf. Perbandingan ini penting untuk menilai trade-off antara akurasi, kompleksitas, dan interpretabilitas.

4. Integrasi dengan dashboard bisnis Hasil model prediksi dapat diintegrasikan ke dalam dashboard analitik yang menampilkan proyeksi penjualan secara visual, sehingga memudahkan pelaku usaha dalam memantau performa dan melakukan penyesuaian strategi secara berkala.
5. Validasi pada data real-time Untuk menguji ketahanan model dalam menghadapi dinamika pasar yang cepat, disarankan dilakukan pengujian pada data real-time atau *near real-time*, sehingga dapat dievaluasi apakah model masih relevan dan akurat ketika dihadapkan pada kondisi aktual yang terus berubah.

data ulasan serta tidak dilibatkannya algoritma pembandingan. Namun, hasil yang diperoleh tetap menunjukkan potensi besar penggunaan pendekatan *machine learning* dalam meningkatkan kualitas layanan dan pengalaman pengguna.

REFERENCE

- [1] D. Karl, "Forecasting e-commerce consumer returns: A systematic literature review," *Manag. Rev. Q.*, vol. 75, no. 3, pp. 1–56, 2024, doi: 10.1007/s11301-024-00436-x.
- [2] H. Ahaggach, L. Abrouk, and E. Lebon, "Systematic mapping study of sales forecasting: Methods, trends, and future directions," *Forecasting*, vol. 6, no. 3, pp. 502–532, 2024, doi: 10.3390/forecast6030028.
- [3] P. Brüggemann, L. F. Martinez, and F. J. Martínez López, "The future of analytics in digital marketing and electronic commerce: how digitalization and sustainability affect consumers, firms, and policymakers," *J. Mark. Anal.*, vol. 13, pp. 571–573, 2025, doi: 10.1057/s41270-025-00422-6.
- [4] S. W. Jarantow, "Introduction to the use of linear and nonlinear regression in life science data modelling," *Curr. Protoc.*, vol. 3, no. 1, p. e80, 2023, doi: 10.1002/cpz1.801.
- [5] N. Roustaei, "Application and interpretation of linear regression analysis," *BMC Med. Res. Methodol.*, vol. 24, p. 29, 2024, doi: 10.1186/s12874-024-01789-2.
- [6] A. P. Lestari, "E-commerce in Indonesia's economic transformation and its influence on global trade," *Int. J. Comput. Law Polit. Sci.*,

vol. 4, no. 10, pp. 10–23, 2024, doi:
10.34010/injucolpos.vxix.xxxx.

[7] S. W. Jarantow, "Introduction to the use of
linear and non-linear regression models," *Curr.
Protoc.*,