JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 5, No. 3, November 2026, pp. 255-268
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/175

Simulation of Paving Block Sales Volume Using K-Nearest Neighbor Regression at CV Surya Perkasa

Rima Anjani[1], Nuk Ghurroh Setyoningrum[2]
[1,2] Fakultas Sains dan Teknologi, Universitas Cipasung
Jl. Borolong, 46466, Kabupaten Tasikmalaya, Jawa Barat, Indonesia
e-mail: rimaanjaniuncip@gmail.com , nuke@uncip.ac.id

ARTICLE INFO
Article History:

Received: 04-06-2026
Revised: 06-07-2026
Accepted: 20-09-2026
Published: 28-10-2026

ABSTRACT
The advancement of information technology encourages businesses to utilize historical data for decision-making. CV Surya Perkasa, a paving block manufacturer, faces challenges in analyzing sales fluctuations due to the limited utilization of historical transaction data. This study aims to implement K-Nearest Neighbor (KNN) Regression as an initial simulation for estimating paving block sales volumes. The research employed a quantitative data mining approach using Python on Google Colab. The dataset comprised 500 sales transactions from 2022 to 2025, aggregated into monthly sales data by product type. Data preprocessing included cleaning, temporal aggregation, feature transformation, One-Hot Encoding, normalization, and historical feature construction using LAG_1 and LAG_12. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R^2. The optimized model achieved an MAE of 331.36, MSE of 160,281.61, RMSE of 400.35, MAPE of 36.33%, and R^2 of 0.56. These results indicate that KNN Regression partially captures historical sales patterns but remains limited by demand fluctuations and historical features. Therefore, further improvements are required before operational forecasting implementation.

Keywords: *Data Mining, K-Nearest Neighbor Regression, Paving block, Sales Volume, Prediction Simulation*

Abstrak

Perkembangan teknologi informasi mendorong pemanfaatan data historis untuk mendukung pengambilan keputusan bisnis. CV Surya Perkasa, perusahaan manufaktur *paving block*, menghadapi kendala dalam menganalisis fluktuasi penjualan akibat belum optimalnya pemanfaatan data transaksi historis. Penelitian ini bertujuan menerapkan algoritma *K-Nearest Neighbor* (KNN) Regression sebagai simulasi awal estimasi volume penjualan *paving block*. Penelitian menggunakan pendekatan kuantitatif berbasis *data mining* dengan Python pada Google Colab. Dataset mencakup 500 transaksi penjualan periode 2022–2025 yang diagregasi menjadi data penjualan bulanan berdasarkan jenis produk. Tahapan prapemrosesan meliputi pembersihan data, agregasi waktu, transformasi fitur, *One-Hot Encoding*, normalisasi, serta pembentukan fitur historis LAG_1 dan LAG_12. Kinerja model dievaluasi menggunakan *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan R^2. Model yang dioptimalkan menghasilkan MAE sebesar 331,36, MSE sebesar 160.281,61, RMSE sebesar 400,35, MAPE sebesar 36,33%, dan R^2 sebesar 0,56. Hasil menunjukkan bahwa model mampu menangkap sebagian pola historis penjualan, tetapi masih terbatas akibat fluktuasi permintaan dan keterbatasan fitur historis. Oleh karena itu, pengembangan lebih lanjut diperlukan sebelum model digunakan untuk peramalan operasional.

Kata Kunci: *Data Mining, K-Nearest Neighbor Regression, Paving block,* Simulasi Estimasi, Volume Penjualan

DOI: https://doi.org/10.51903/18143727

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

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong perubahan dalam pengelolaan data pada berbagai sektor bisnis. Data transaksi yang sebelumnya hanya digunakan sebagai catatan operasional kini dapat dimanfaatkan sebagai sumber informasi untuk memahami pola historis dan mendukung proses pengambilan keputusan [1]. Salah satu pemanfaatan data tersebut adalah analisis pola penjualan untuk membantu perusahaan memahami perubahan permintaan produk berdasarkan data historis yang tersedia [2]. CV Surya Perkasa menghadapi tantangan dalam mengelola fluktuasi permintaan dan variasi produk paving block. Meskipun data transaksi penjualan terus bertambah, data tersebut belum dimanfaatkan secara optimal untuk menganalisis pola penjualan musiman. Akibatnya, proses evaluasi dan perencanaan penjualan perusahaan masih bergantung pada analisis historis manual.

Pemanfaatan metode *data mining* dapat menjadi salah satu pendekatan untuk menggali pola dari data historis penjualan. Data mining memungkinkan proses identifikasi hubungan antarvariabel dan pola tersembunyi dalam kumpulan data yang besar melalui pendekatan komputasi [3]. Salah satu metode yang dapat digunakan adalah *K-Nearest Neighbor* (KNN) *Regression*, yaitu algoritma berbasis kemiripan data yang melakukan estimasi berdasarkan kedekatan karakteristik antara data baru dengan data historis. Meskipun demikian, performa KNN *Regression* sangat dipengaruhi oleh karakteristik dataset, jumlah observasi, pemilihan parameter K, serta fitur yang digunakan dalam proses pemodelan [4]. Untuk efisiensi, penelitian ini mengimplementasikan algoritma KNN menggunakan platform *Google Colab*. Dukungan library *Python* seperti *Scikit-Learn*, *Pandas*, dan *Matplotlib* memungkinkan pemrosesan dataset besar dan perhitungan jarak *Euclidean* dilakukan jauh lebih cepat dan akurat dibandingkan metode manual.

Algoritma *K-Nearest Neighbor* (KNN) *Regression* merupakan metode *machine learning* berbasis kemiripan data yang melakukan estimasi berdasarkan kedekatan pola antara data baru dan historis. Performanya dipengaruhi oleh karakteristik data, pemilihan fitur, dan parameter K, sehingga efektif membantu mengidentifikasi pola penjualan berdasarkan kemiripan produk dan waktu [4]. Sebagai material konstruksi, permintaan *paving block* dipengaruhi oleh aktivitas pembangunan, kondisi pasar, dan faktor musiman yang fluktuatif. Fluktuasi ini menjadi tantangan tersendiri bagi perusahaan dalam mengevaluasi pola penjualan serta menyusun strategi produksi yang akurat berdasarkan data historis [5].

Berdasarkan data BPS (2025), sektor konstruksi memiliki kontribusi penting terhadap perekonomian nasional dengan kontribusi sebesar 9,48% terhadap PDB. Namun, dinamika pertumbuhan sektor konstruksi menunjukkan adanya perubahan aktivitas pembangunan yang dapat memengaruhi kebutuhan terhadap material konstruksi, termasuk paving block. Kondisi tersebut menunjukkan bahwa perusahaan penyedia material konstruksi perlu memiliki pendekatan berbasis data untuk memahami perubahan pola permintaan dan mendukung perencanaan persediaan secara lebih efektif [6]. Salah satu perusahaan yang menghadapi tantangan tersebut adalah CV Surya Perkasa yang bergerak pada produksi paving block. Meskipun perusahaan memiliki data transaksi penjualan historis, data tersebut belum dimanfaatkan secara optimal untuk mengidentifikasi pola penjualan dan memperkirakan perubahan volume permintaan. Penerapan metode KNN Regression dalam penelitian ini digunakan untuk menghasilkan simulasi estimasi volume penjualan paving block berdasarkan pola historis transaksi sebagai informasi pendukung dalam memahami karakteristik permintaan produk.

2. TINJAUAN PUSTAKA

2.1. Literatur Review

Penelitian mengenai estimasi dan estimasi penjualan menggunakan algoritma *K-Nearest Neighbor* (KNN) menunjukkan bahwa metode berbasis kemiripan jarak mampu digunakan untuk mengenali pola berdasarkan karakteristik data historis. Sejumlah studi, seperti Bakriansyah et al. [7], Gautama dan Arijanto [8], serta Munthe et al. [1], telah membuktikan efisiensi KNN untuk mengklasifikasikan produk terlaris guna menunjang pengendalian stok. Di sektor material konstruksi, pendekatan analisis data juga banyak dieksplorasi, mulai dari penggunaan algoritma *Apriori* untuk mencari pola asosiasi bahan bangunan hingga metode hibrida (*hybrid*) C4.5 dan KNN menggunakan *RapidMiner*. Khusus pada komoditas *Paving block* Kurniawan et al. [9], telah menerapkan Regresi Linear Berganda berbasis web untuk memproyeksikan harga jual pasar. Meskipun algoritma KNN telah banyak digunakan untuk estimasi penjualan, beberapa penelitian sebelumnya masih berfokus pada karakteristik produk tanpa mempertimbangkan fitur historis berbasis waktu secara eksplisit.

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

Penambahan fitur historis (*lag*) dalam penelitian ini diterapkan untuk membantu model menangkap perubahan volume penjualan berdasarkan periode sebelumnya secara lebih optimal. *Research gap* penelitian ini meliputi: mengembangkan pendekatan KNN Regression dengan menambahkan fitur waktu dan fitur historis berupa LAG_1 dan LAG_12 untuk memperoleh representasi pola penjualan berdasarkan histori data. Penelitian ini menawarkan kebaruan melalui penerapan algoritma *K-Nearest Neighbor Regression* sebagai simulasi awal estimasi volume penjualan *paving block* di CV Surya Perkasa berdasarkan fitur waktu dan karakteristik produk. Hasil kajian ini diharapkan dapat memberikan gambaran awal penerapan *data mining* untuk mendukung analisis penjualan serta menjadi acuan pengembangan model estimasi yang lebih kompleks pada masa mendatang.

2.2. *Data mining*

Data mining mengekstrak informasi dan pola tersembunyi dari sekumpulan data melalui pendekatan statistik, matematika, dan komputasi untuk mendukung pengambilan keputusan bisnis. Dalam sektor komersial, penerapan data mining pada data transaksi membantu perusahaan memahami pola penjualan historis guna mengantisipasi fluktuasi permintaan, di mana kualitas pemodelan sangat bergantung pada tahapan pra-pemrosesan serta karakteristik variabel yang digunakan [10].

2.3. Fitur historis pada data penjualan

Karakteristik data penjualan sangat dipengaruhi oleh waktu, sehingga fitur *lag* seperti *LAG_1* (periode sebelumnya) dan *LAG_12* (musiman tahun lalu) penting untuk merepresentasikan pola temporal dalam model *machine learning*. Namun, penggunaannya memerlukan data historis yang memadai agar tidak mengurangi ukuran sampel efektif akibat eliminasi baris data awal[11].

2.4. Algoritma *K-Nearest Neighbor (KNN) Regression*

K-Nearest Neighbor (KNN) adalah algoritma *machine learning* berbasis *instance-based learning* yang melakukan estimasi berdasarkan tingkat kemiripan data baru dengan sejumlah data historis terdekat (K). Dalam kasus regresi, nilai estimasi dihitung dari rata-rata nilai target tetangga terdekat [12]. Perhitungan kedekatan data dalam algoritma KNN umumnya menggunakan *Euclidean Distance*. Penentuan nilai parameter K sangat memengaruhi performa model: nilai yang terlalu kecil membuat model sensitif terhadap *noise*, sedangkan nilai yang terlalu besar menghilangkan pola lokal data. Oleh karena itu, optimasi dilakukan melalui pengujian berbagai konfigurasi untuk meminimalkan tingkat kesalahan [13]. Meskipun efektif mengenali pola historis, performa KNN sangat bergantung pada volume data, distribusinya, pemilihan fitur, serta proses normalisasi. Oleh karena itu, penerapan metode ini pada data penjualan memerlukan transformasi fitur yang tepat agar karakteristik temporal dan struktural data dapat direpresentasikan secara optimal [14]. diaplikasikan untuk simulasi peramalan *volume* penjualan *paving block* berdasarkan data historis yang tersedia di CV Surya Perkasa.

2.5. Evaluasi Model *(Mean Absolute Error/ Mean Squared Error)*

Evaluasi performa model regresi menggunakan metrik seperti MAE (rata-rata selisih absolut), RMSE (penalti kesalahan tinggi melalui akar kuadrat MSE), MAPE (persentase rata-rata kesalahan relatif), dan R^2 (proporsi variasi target yang dijelaskan model) untuk mengukur akurasi serta generalisasi estimasi [15]. Penggunaan beberapa metrik evaluasi secara bersamaan diperlukan karena setiap metrik memiliki karakteristik interpretasi yang berbeda. Dengan demikian, performa model dapat dianalisis secara lebih menyeluruh dan tidak hanya bergantung pada satu ukuran kesalahan [16].

2.6. *Google Colaboratory*

Google Colaboratory adalah *platform* belajar berbasis *cloud* yang populer untuk pembelajaran Kecerdasan Buatan (AI) dan pemrograman *Python* dengan metode praktik langsung. Fasilitas *notebook interaktif* memungkinkan pengguna untuk mengintegrasikan kode, teks, dan visual dalam satu ruang kerja yang terpadu [17]. *Google Colab* sangat digemari dalam data mining karena menyediakan komputasi GPU/TPU gratis, integrasi mudah dengan Google Drive, dan *library Python* yang lengkap, sehingga membebaskan peneliti dari keterbatasan spesifikasi perangkat keras lokalnya [18]. *Google Colab* mendukung penuh analisis dan pemodelan *machine learning* lewat integrasi *library* seperti *NumPy*, *Pandas*, dan *Scikit-Learn*. Platform ini memungkinkan eksekusi algoritma estimasi serta evaluasi model (MAE dan MSE) secara presisi tanpa membebani perangkat keras lokal [19]. Keunggulan *Google Colab* terletak pada komputasi berbasis *cloud* dengan dukungan GPU/TPU gratis, *library Python* yang lengkap, integrasi Google Drive tanpa instalasi lokal, serta fitur kolaborasi *real-time* [20].

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

2.7. *Paving block* dan Manajemen Persediaan

Paving block merupakan pelapis permukaan jalan modular yang mengandalkan sistem *interlocking* untuk mendistribusikan beban secara merata. Karakteristik fisik ini menjadikannya produk dengan penjualan yang ber*fluktuasi* sesuai dengan tren pembangunan infrastruktur di suatu wilayah [21]. Disajikan dalam gambar 1.



Gambar 1 Jenis Jenis *Paving block* . *(Sumber: indonusa-conblock.com)[22].*

Gambar 1 menunjukkan dimensi fisik dan estimasi kebutuhan kuantitas *paving block* per m² berdasarkan SNI 03-0691-1996. Variasi bentuk geometri memengaruhi jumlah keping per m² karena perbedaan luas permukaan (misal: Bata 44-50 buah, Hexagon 27 buah, Tahu 100 buah). Sementara itu, ketebalan (6-10 cm) tidak memengaruhi jumlah keping, melainkan menentukan klasifikasi kekuatan beban kendaraan yang sanggup ditahan (Mutu A hingga D).

3. METODOLOGI PENELITIAN

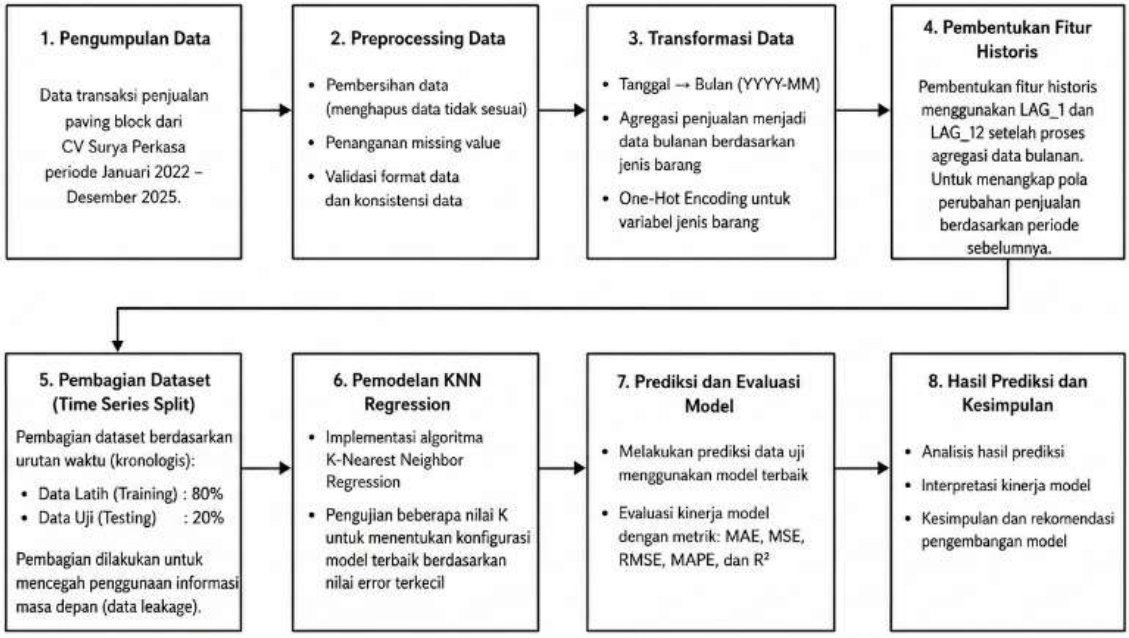
3.1. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui tiga metode utama, yaitu wawancara langsung dengan pemilik dan staf administrasi CV Surya Perkasa untuk memperoleh data operasional serta faktor luar yang memengaruhi penjualan, dokumentasi berupa permintaan data historis transaksi penjualan *paving block* periode 2022–2025 sebagai dataset utama algoritma *K-Nearest Neighbor Regression*, serta studi literatur dari jurnal akademik dan buku referensi guna memperkuat dasar ilmiah metodologi data mining yang digunakan.

3.2. Tahapan Penelitian

Rangkaian tahapan penelitian disistematisasikan sedemikian rupa demi memastikan pelaksanaan analisis yang memenuhi kaidah ilmiah serta berorientasi pada target yang telah ditetapkan. Prosedur kerja penelitian ini disajikan dalam gambar 2.

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



Gambar 2 Tahapan penelitian

Gambar 2 menggambarkan tahapan penelitian estimasi volume penjualan *paving block* menggunakan algoritma *K-Nearest Neighbor* (KNN) *Regression*. Tahapan penelitian dimulai dari proses pengumpulan data transaksi, *preprocessing data*, transformasi data, pembentukan fitur historis, normalisasi data, pembagian dataset secara kronologis menjadi data latih dan data uji, pemodelan KNN Regression, serta evaluasi kinerja model menggunakan metrik MAE, MSE, RMSE, MAPE, dan R^2. Tahapan penelitian diakhiri dengan analisis hasil estimasi dan penarikan kesimpulan.

3.3. Analisis Data

Tahap analisis data dalam penelitian ini dilakukan melalui beberapa tahapan, yaitu *preprocessing data*, transformasi data, pembentukan fitur historis, pembagian dataset, pemodelan *K-Nearest Neighbor Regression*, dan evaluasi model. Data transaksi yang diperoleh terlebih dahulu melalui proses pembersihan dan penyesuaian format untuk memastikan kualitas data. Selanjutnya, data ditransformasikan menjadi data berbasis waktu dan dilakukan pembentukan fitur historis untuk merepresentasikan pola penjualan periode sebelumnya [23]. Dataset kemudian dibagi secara kronologis menggunakan pendekatan *time series split* dengan 80% data awal sebagai data pelatihan dan 20% data terakhir sebagai data pengujian. Berdasarkan pembagian tersebut, data pelatihan menggunakan periode Oktober 2022-September 2025, sedangkan data pengujian menggunakan periode Oktober-Desember 2025. Implementasi algoritma *K-Nearest Neighbor* (KNN) *Regression* dilakukan dengan menghitung kedekatan antar data menggunakan jarak *Euclidean*. Tahapan pemodelan meliputi preprocessing data, pembentukan fitur historis, transformasi data kategorikal, normalisasi, pembagian data secara kronologis, penentuan parameter K optimal, serta evaluasi model menggunakan MAE, MSE, RMSE, MAPE, dan R^2 untuk menghasilkan simulasi estimasi volume penjualan [24]. Rumus jarak *Euclidean* yang diaplikasikan didefinisikan sebagai rumus (1):

$$d = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \quad (1)$$

Dimana:
x_i= data training
y_i= data testing
n= jumlah atribut

Data mentah dari CV Surya Perkasa (mencakup nomor dan tanggal transaksi, lokasi, jenis barang, harga satuan, total harga, *volume*, dan keterangan) melalui tahap *preprocessing* sebelum pemodelan KNN *Regression*. Tahapan ini meliputi pembersihan data, penanganan nilai kosong (*missing value*), transformasi dan agregasi transaksi menjadi periode bulanan, Variabel kategorikal berupa jenis barang dikonversi menggunakan *One-Hot Encoding*. Metode ini dipilih karena tidak memberikan hubungan ordinal antar kategori produk sehingga tidak menghasilkan jarak numerik semu pada proses perhitungan KNN. disajikan pada Tabel 1.

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

Tabel 1. Data mentah

No	Tanggal	Lokasi	Barang	Harga Satuan	Total Harga	*Volume*	Keterangan
1	01/01/2022	Smpn 3 Sukajaya Majalengka	Paving Cacing 8cm	90000	52200000	580	Terkirim
2	04/01/2022	Kampus Kencana Kadungora Kab. Bandung	Paving Trihex 6cm	80000	152000000	1900	Terkirim
3	07/01/2022	Sdn 7 Karangjaya Cirebon	Paving Bata 6cm	80000	172400000	2155	Terpasang
4	10/01/2022	Gedung Serbaguna Nusantara Cigugur Bogor	Paving Bata 8cm	90000	98460000	1094	Terpasang
5	13/01/2022	Tk 3 Mekarjaya Indramayu	Paving Bata 6cm	80000	160320000	2004	Terpasang
…	…	…	…			…	…
500	30/12/2025	Perumahan Makmur Sukasari Bandung	Paving Topi Uskup 6cm	80000	31760000	397	Terkirim

Setelah tahap *preprocessing* dilakukan, penelitian menentukan atribut yang digunakan sebagai fitur masukan (*input*) pada model *K-Nearest Neighbor Regression*. Variabel target yang diestimasi adalah volume penjualan (*volume*), sedangkan fitur prediktor yang digunakan terdiri dari variabel waktu, karakteristik produk, harga satuan, serta fitur historis penjualan berupa LAG_1 dan LAG_12. Atribut total harga yang tersedia pada data transaksi awal tidak digunakan sebagai fitur prediktor karena memiliki hubungan matematis langsung dengan volume dan harga satuan (*Total Harga = Volume × Harga Satuan*), sehingga berpotensi menyebabkan *target leakage* pada model.

Kinerja model dievaluasi menggunakan metrik regresi yaitu MAE, MSE, RMSE, MAPE, dan R^2 untuk mengetahui tingkat kesalahan estimasi serta kemampuan model dalam menjelaskan variasi volume penjualan. Kinerja model dievaluasi menggunakan beberapa metrik regresi, yaitu *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2). Penggunaan beberapa metrik tersebut bertujuan untuk mengetahui tingkat kesalahan estimasi serta kemampuan model dalam menjelaskan variasi data volume penjualan. Rumus *Mean Absolute Error:*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

Dimana:
n = rata-rata kesalahan *absolut* antara nilai aktual dan nilai estimasi
i = jumlah data
y_i = indeks data
$\hat{y}_i$ = nilai aktual
$| y_i - \hat{y}_i |$ = selisih absolut antara nilai aktual dan estimasi [25].

Rumus *Mean Squared Error*:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (3)$$

Dimana:
n = rata-rata kesalahan absolut antara nilai aktual dan nilai estimasi
i = jumlah data
y_i = indeks data
$\hat{y}_i$ = nilai aktual

$(y_i - \hat{y}_i)^2$ = selisih antara nilai aktual dan estimasi yang dikuadratkan [26].

Rumus Root Mean Squared *Error*:

$$RMSE = \sqrt{MSE} \quad (4)$$

Dimana akar nilai rata-rata kesalahan kuadrat

Rumus Mean Absolute Perecentage *Error* :

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \quad (5)$$

Dimana :
n = Jumlah total observasi atau banyak data yang di evaluasi

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

y_i = Nilai aktual *volume* penjualan pada periode ke- i
$\hat{y}_i$ = Nilai estimasi *volume* penjualan algoritma KNN pada periode ke- i
$\| =$ Simbol nilai mutlak (absolut), mengabaikan tanda negatif pada selisih antara nilai estimasi dan nilai aktual

MAPE memiliki beberapa kategori evaluasi, disajikan pada tabel 2.

Tabel 2. Kategori MAPE

Nilai	Kategori
<10%	Sangat baik
10% - 20%	Baik
20% - 50%	Cukup
50%	Kurang baik

Rumus R^2 Koefisien Determinasi (ukuran proporsi variansi data) :

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (6)$$

Dimana :
n = Jumlah total observasi atau banyak data yang di evaluasi
y_i = Nilai aktual *volume* penjualan pada periode ke- i
$\hat{y}_i$ = Nilai estimasi *volume* penjualan algoritma KNN pada periode ke- i
$\hat{y}$ = Nilai rata-rata dari keseluruhan data aktual *volume* penjualan
$\sum_{i=1}^{n}$ = Simbol sigma, menunjukkan penjumlahan perhitungan dari data pertama (i = 1) hingga data ke- n

Nilai R^2 memiliki ketegori, disajikan pada tabel 3.

Tabel 3. *Kategori* R^2

Nilai	Kategori
0.90 – 1.00	Sangat kuat (Sangat baik)
0.70 – 0.89	Kuat (baik)
0.50 – 0.69	Sedang
<0.50	Lemah

4. HASIL DAN PEMBAHASAN

4.1. Transformasi Data Dan Pembentukan Fitur Historis

Data transaksi dilakukan agregasi berdasarkan periode bulanan dan jenis barang untuk memperoleh volume penjualan serta harga satuan pada setiap periode. Selanjutnya, dilakukan pembentukan fitur historis berupa LAG_1 dan LAG_12 serta representasi waktu menggunakan Time Index. Fitur LAG_1 digunakan untuk merepresentasikan volume penjualan periode sebelumnya, sedangkan LAG_12 digunakan untuk menangkap pola penjualan pada periode yang sama di tahun sebelumnya. Selain itu, dilakukan normalisasi terhadap variabel numerik agar setiap fitur memiliki rentang nilai yang sebanding sebelum proses pemodelan KNN Regression. Hasil transformasi data disajikan pada Tabel 4.

Tabel 4. Transformasi Data dan Pembentukan Fitur Historis

No	Tahun	Bulan	Barang	Volume	Harga Satuan	Time Index	LAG_1	LAG_12
102	2023	4	Paving Bata 6cm	1540	80000.0	24280	842.0	6361.0
119	2023	6	Paving Grass Block 6cm	228	80000.0	24282	500.0	832.0
122	2023	6	Paving Trihex 6cm	3510	80000.0	24282	3089.0	1900.0
124	2023	7	Paving Bata 6cm	1720	80000.0	24283	1540.0	1471.0
125	2023	7	Paving Grass Block 6cm	491	80000.0	24283	228.0	1383.0

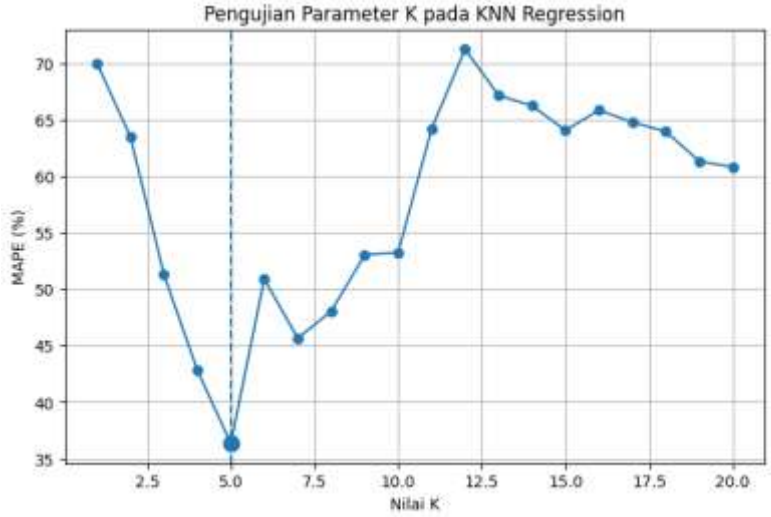
Variabel total harga tidak digunakan sebagai fitur masukan model karena memiliki hubungan matematis langsung dengan variabel target dan harga satuan (*Total Harga = Volume × Harga Satuan*). Penggunaan variabel tersebut dapat menyebabkan informasi target masuk secara tidak langsung ke dalam fitur prediktor sehingga berpotensi menimbulkan *target leakage* pada proses pemodelan.

4.2. Pemodelan *K-nearest neighbor (KNN) Regression*

Tahap Implementasi algoritma *K-Nearest Neighbor* (KNN) *Regression* dilakukan dengan menghitung kedekatan antar data menggunakan jarak Euclidean. Pemodelan dilakukan melalui proses normalisasi fitur, pengujian beberapa nilai parameter K untuk memperoleh konfigurasi model terbaik, serta evaluasi performa model menggunakan MAE, MSE, RMSE, MAPE, dan R^2. Pengujian dilakukan dengan mencoba beberapa

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

nilai parameter K untuk memperoleh konfigurasi model terbaik berdasarkan nilai kesalahan terkecil. Berdasarkan hasil pengujian, diperoleh nilai K=5 sebagai parameter terbaik dengan nilai MAPE terendah sebesar 36,33%. Hasil pengujian parameter K disajikan pada Gambar 3.



Gambar 3 Hasil parameter k terbaik

Gambar 3 menyajikan hasil pengujian beberapa nilai K berdasarkan nilai MAPE. Nilai K terbaik dipilih berdasarkan tingkat kesalahan terkecil yang diperoleh pada proses pengujian model. Berdasarkan hasil pengujian tersebut, nilai K=5 menghasilkan nilai MAPE terendah sehingga digunakan sebagai parameter akhir model KNN *Regression*.

4.3. Hasil Estimasi Dan Evaluasi Kinerja Model

Tahap akhir penelitian dilakukan dengan menggunakan model *K-Nearest Neighbor Regression* terbaik untuk menghasilkan estimasi volume penjualan pada data pengujian. Model yang digunakan merupakan model dengan parameter K=5 berdasarkan hasil pengujian parameter sebelumnya. Evaluasi kinerja model dilakukan menggunakan beberapa metrik regresi, yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2). Hasil evaluasi menunjukkan bahwa model memperoleh nilai MAE sebesar 331,36, yang menunjukkan rata-rata selisih antara nilai aktual dan hasil estimasi volume penjualan. Nilai RMSE yang diperoleh sebesar 400,35 menunjukkan tingkat penyimpangan estimasi dengan memberikan penalti lebih besar terhadap kesalahan yang tinggi.

Selain itu, nilai *Mean Squared Error* (MSE) yang diperoleh sebesar 160281.61 menunjukkan besarnya rata-rata kuadrat kesalahan antara nilai aktual dan hasil estimasi model. Nilai MSE digunakan sebagai indikator tambahan untuk melihat tingkat penyimpangan estimasi, terutama terhadap kesalahan dengan nilai yang besar. Sementara itu, nilai MAPE sebesar 36,33% menunjukkan bahwa rata-rata persentase kesalahan estimasi model masih berada pada tingkat tertentu, sehingga hasil estimasi perlu diinterpretasikan sebagai simulasi berdasarkan pola historis yang tersedia. Nilai koefisien determinasi (R^2) sebesar 0,56 menunjukkan bahwa model mampu menjelaskan sekitar 56% variasi data volume penjualan berdasarkan fitur yang digunakan dalam penelitian. Nilai tersebut menunjukkan bahwa model telah mampu menangkap sebagian pola historis penjualan, namun masih terdapat faktor lain di luar variabel penelitian yang dapat memengaruhi perubahan volume penjualan.

Berdasarkan hasil evaluasi tersebut, model *K-Nearest Neighbor Regression* dapat digunakan untuk memberikan gambaran awal mengenai pola estimasi volume penjualan berdasarkan data historis transaksi. Namun, hasil estimasi belum dapat digunakan sebagai sistem peramalan operasional secara langsung karena performa model masih dipengaruhi oleh jumlah data, karakteristik pola permintaan, dan variabel lain yang belum dimasukkan dalam model. Hasil penelitian ini masih berada pada tahap simulasi awal berbasis data historis. Pengembangan lebih lanjut dapat dilakukan dengan menambahkan fitur berbasis waktu yang lebih kompleks dan memperluas periode pengamatan agar model mampu menangkap pola permintaan yang lebih beragam. Untuk memberikan gambaran yang komprehensif mengenai performa akhir dari model *K-Nearest Neighbor* (KNN) *Regression*, seluruh hasil komputasi dari kelima metrik evaluasi yang telah diuji dirangkum ke dalam sebuah matriks perbandingan. Rangkuman matriks hasil evaluasi disajikan pada Tabel 5.

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

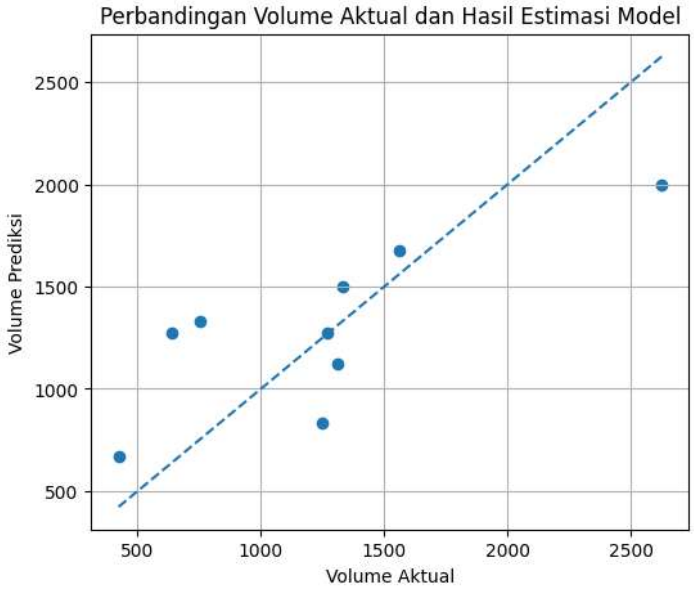
Tabel 5. Matriks evaluasi model

Metric	Value
MAE	331.36
MSE	160281.61
RMSE	400.35
MAPE	36.33%
R^2	0.56

Tabel 5 menyajikan matriks evaluasi kinerja model *K-Nearest Neighbor Regression* dengan parameter terbaik K=5. Hasil evaluasi menunjukkan nilai MAE sebesar 331,36, MSE sebesar 160.281,61, RMSE sebesar 400,35, MAPE sebesar 36,33%, dan R^2 sebesar 0,56. Nilai R^2 tersebut menunjukkan bahwa model mampu menjelaskan sebagian variasi volume penjualan berdasarkan fitur yang digunakan. Sementara itu, nilai error menunjukkan masih terdapat perbedaan antara hasil estimasi model dan nilai aktual, sehingga hasil estimasi perlu diinterpretasikan sebagai simulasi berdasarkan pola historis penjualan. Model ini dapat digunakan sebagai pendekatan awal dalam mengevaluasi pola penjualan, namun penerapannya dalam pengambilan keputusan operasional masih perlu mempertimbangkan faktor lain yang memengaruhi perubahan permintaan.

4.4. Perbandingan Nilai Aktual dan Hasil Estimasi Model

Evaluasi akhir dilakukan dengan membandingkan nilai aktual volume penjualan dengan hasil estimasi model pada data pengujian. Perbandingan ini bertujuan untuk melihat kemampuan model *K-Nearest Neighbor Regression* dalam mengikuti pola perubahan volume penjualan berdasarkan data historis yang tersedia. Visualisasi perbandingan nilai aktual dan hasil estimasi disajikan pada Gambar 4.



Gambar 4 Perbandingan Volume Aktual dan Hasil Estimasi Model

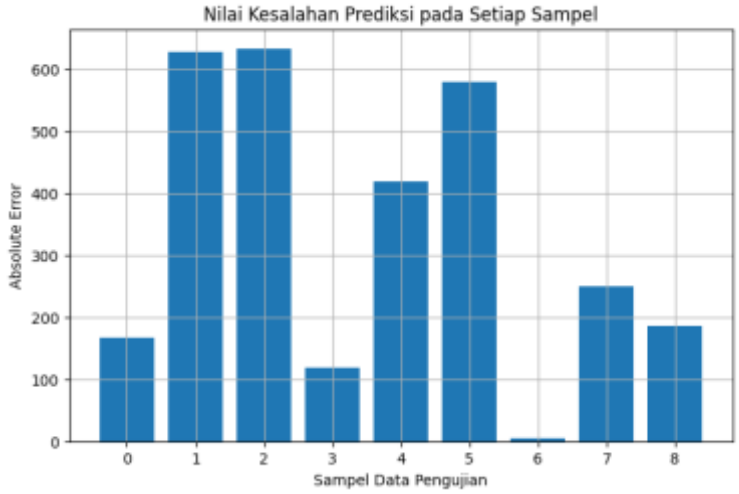
Gambar 4 menunjukkan perbandingan antara volume aktual dan hasil estimasi model pada data pengujian yang diperoleh dari pembagian dataset secara kronologis. Garis diagonal pada grafik menunjukkan kondisi ideal ketika nilai estimasi memiliki nilai yang sama dengan nilai aktual. Sebagian titik estimasi berada di sekitar garis diagonal, yang menunjukkan bahwa model mampu menangkap sebagian pola perubahan volume penjualan berdasarkan fitur historis yang digunakan. Namun, masih terdapat beberapa titik dengan penyimpangan cukup jauh dari garis ideal yang menunjukkan adanya perbedaan antara nilai aktual dan hasil estimasi. Kondisi tersebut sejalan dengan nilai R^2 sebesar 0,56 yang menunjukkan bahwa model mampu menjelaskan sebagian variasi volume penjualan. Perbedaan antara nilai aktual dan hasil estimasi menunjukkan bahwa masih terdapat variasi pola penjualan yang belum sepenuhnya direpresentasikan oleh fitur historis dan karakteristik produk yang digunakan dalam penelitian.

4.5. Analisis Kesalahan Estimasi pada Data Pengujian

Analisis kesalahan estimasi dilakukan untuk melihat perbedaan antara nilai aktual dan hasil estimasi model pada setiap data pengujian. Perhitungan kesalahan ini digunakan untuk mengetahui variasi selisih estimasi

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

secara individual yang tidak terlihat pada nilai evaluasi agregat seperti MAE, RMSE, MAPE, dan R^2. Hasil analisis kesalahan estimasi disajikan pada Gambar 5.

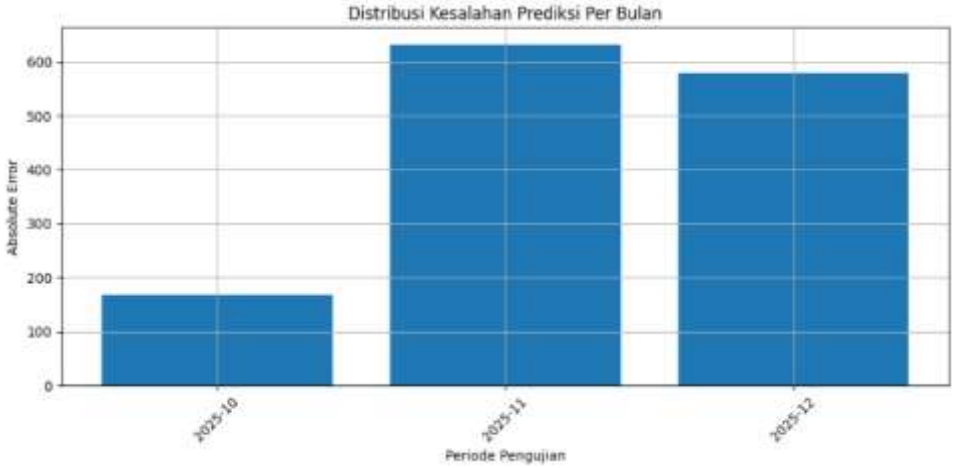


Gambar 5 Hasil kesalahan estimasi persampel

Gambar 5 menunjukkan nilai kesalahan absolut (*absolute error*) pada setiap sampel data pengujian. Nilai tersebut menggambarkan besarnya selisih antara volume aktual dan hasil estimasi model *K-Nearest Neighbor Regression* dengan parameter terbaik K=5. Perbedaan nilai kesalahan pada setiap sampel menunjukkan bahwa kemampuan model dalam melakukan estimasi dapat bervariasi bergantung pada karakteristik data yang diuji. Beberapa sampel memiliki nilai kesalahan lebih tinggi karena terdapat perbedaan pola antara data historis yang dipelajari model dan kondisi aktual pada periode pengujian. Analisis ini digunakan sebagai pendukung evaluasi metrik keseluruhan dalam menilai performa model.

4.6. Distribusi Kesalahan Estimasi Per bulan

Analisis distribusi kesalahan estimasi per bulan dilakukan untuk melihat variasi tingkat kesalahan model pada setiap periode pengujian. Nilai kesalahan dihitung berdasarkan selisih absolut antara volume aktual dan hasil estimasi model *K-Nearest Neighbor Regression* dengan parameter terbaik K=5. Hasil distribusi kesalahan estimasi per bulan disajikan pada Gambar 6.



Gambar 6 Hasil Distribusi Kesalahan Estimasi Per bulan

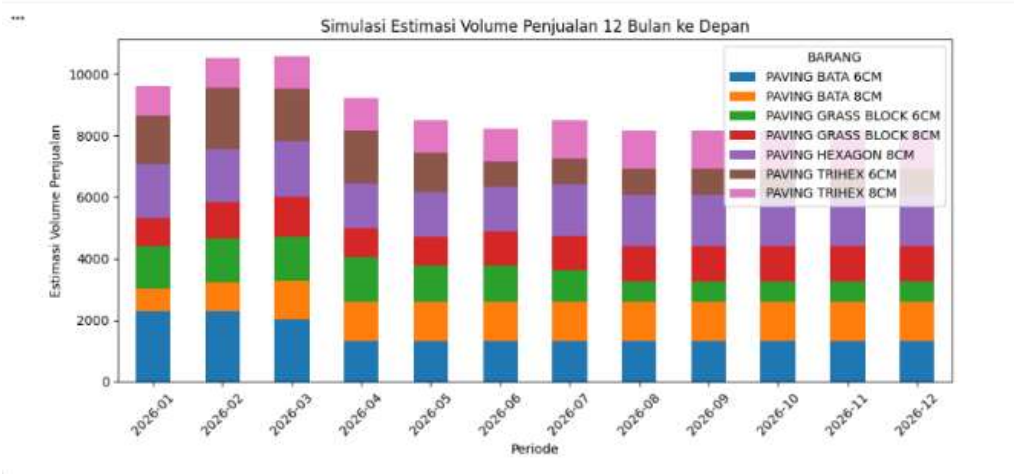
Gambar 6 menunjukkan variasi nilai *absolute error* pada periode pengujian Oktober hingga Desember 2025. Berdasarkan hasil pengamatan, nilai kesalahan estimasi berbeda pada setiap periode, dengan kesalahan tertinggi terjadi pada November 2025, sedangkan kesalahan terendah terjadi pada Oktober 2025. Perbedaan nilai kesalahan tersebut menunjukkan bahwa kemampuan model dalam melakukan estimasi dapat berubah mengikuti karakteristik data pada masing-masing periode. Tingginya nilai kesalahan pada periode tertentu dapat disebabkan oleh adanya perbedaan pola antara data historis yang digunakan dalam proses pelatihan

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

dengan kondisi aktual pada data pengujian. Oleh karena itu, hasil estimasi model perlu dipahami sebagai simulasi berdasarkan pola historis penjualan, bukan sebagai estimasi operasional yang bersifat mutlak.

4.7. Simulasi Estimasi Volume Penjualan Periode Januari-Desember 2026

Sebagai tahap akhir penelitian, dilakukan simulasi estimasi volume penjualan selama 12 bulan ke depan menggunakan model *K-Nearest Neighbor Regression* dengan parameter terbaik K=5. Simulasi ini dilakukan untuk memberikan gambaran pola perubahan volume penjualan berdasarkan karakteristik data historis yang telah dipelajari oleh model. Hasil Simulasi estimasi volume penjualan periode januari-desember 2026 disajikan pada Gambar 7.



Gambar 7 Simulasi Estimasi Volume Penjualan Periode Januari-Desember 2026

Gambar 7 menunjukkan hasil simulasi estimasi volume penjualan beberapa jenis *paving block* pada periode Januari hingga Desember 2026 berdasarkan pola historis yang dipelajari oleh model *K-Nearest Neighbor Regression*. Hasil simulasi menunjukkan adanya variasi estimasi volume antarperiode dan antarjenis produk. Perubahan nilai estimasi pada setiap periode menggambarkan respons model terhadap pola historis penjualan yang direpresentasikan melalui fitur waktu dan fitur lag. Namun, hasil simulasi ini masih memiliki keterbatasan karena hanya menggunakan informasi historis tanpa mempertimbangkan faktor eksternal seperti perubahan permintaan pasar, kondisi proyek, dan faktor musiman lainnya.

4.8. Diskusi

Hasil penelitian menunjukkan bahwa algoritma *K-Nearest Neighbor* (KNN) Regression mampu mengidentifikasi sebagian pola historis volume penjualan *paving block* pada CV Surya Perkasa. Berdasarkan pengujian beberapa parameter, nilai K=5 menghasilkan konfigurasi terbaik dengan *Mean Absolute Percentage Error* (MAPE) sebesar 36,33%. Sementara itu, nilai *Mean Absolute Error* (MAE) sebesar 331,36, *Root Mean Squared Error* (RMSE) sebesar 400,35, dan koefisien determinasi (R^2) sebesar 0,56 menunjukkan bahwa model masih memiliki tingkat kesalahan estimasi yang perlu diperhatikan. Berdasarkan kategori evaluasi yang digunakan dalam penelitian, nilai MAPE termasuk kategori cukup, sedangkan R^2 berada pada kategori sedang. Temuan ini menunjukkan bahwa KNN Regression memiliki kemampuan awal dalam merepresentasikan perubahan volume penjualan, tetapi belum sepenuhnya mampu mengikuti fluktuasi permintaan. Hasil tersebut sejalan dengan penelitian Wijaya dkk. [4] yang membahas penerapan KNN untuk estimasi penjualan barang serta penelitian Alif [13] yang menerapkan KNN Regression dalam estimasi penjualan produk. Kedua penelitian tersebut memperkuat relevansi penggunaan pendekatan berbasis kedekatan data historis dalam analisis penjualan.

Penggunaan fitur historis LAG_1 dan LAG_12 menjadi salah satu aspek penting dalam pemodelan karena memungkinkan data penjualan direpresentasikan berdasarkan periode sebelumnya dan periode yang sama pada tahun sebelumnya. Integrasi fitur tersebut dengan variabel waktu, jenis produk, dan harga satuan memberikan informasi tambahan untuk mengidentifikasi kemiripan pola antarperiode. Namun, hasil evaluasi menunjukkan bahwa penggunaan fitur historis belum sepenuhnya mampu merepresentasikan variasi permintaan yang terjadi. Hal ini dapat dipengaruhi oleh karakteristik penjualan material konstruksi yang berkaitan dengan aktivitas proyek, perubahan kebutuhan pasar, serta faktor musiman yang belum dimasukkan secara langsung dalam model. Temuan ini relevan dengan penelitian Salkiawati dkk. [14] mengenai

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

penerapan KNN Regression dalam analisis estimasi penjualan produk. Selain itu, penelitian Putri dkk. [5] yang membahas prediksi permintaan konsumen menggunakan KNN memperlihatkan relevansi pendekatan berbasis data historis untuk memahami pola permintaan. Dengan demikian, penambahan fitur historis merupakan langkah awal yang penting, tetapi masih memerlukan dukungan variabel lain untuk meningkatkan kemampuan estimasi.

Analisis perbandingan nilai aktual dan hasil estimasi memperlihatkan adanya perbedaan tingkat kesalahan pada setiap sampel maupun periode pengujian. Berdasarkan visualisasi hasil penelitian, sebagian nilai estimasi berada mendekati nilai aktual, sedangkan beberapa sampel menunjukkan penyimpangan yang relatif besar. Distribusi kesalahan bulanan juga memperlihatkan bahwa November 2025 memiliki kesalahan estimasi tertinggi, sementara Oktober 2025 menunjukkan kesalahan terendah. Kondisi tersebut mengindikasikan bahwa performa model belum konsisten dalam menghadapi perubahan karakteristik penjualan antarperiode.

Hasil ini memperkuat pentingnya penggunaan beberapa metrik evaluasi secara bersamaan agar kemampuan model tidak hanya dinilai berdasarkan satu indikator. Penelitian Hayya' dkk. [15] dan Ananda serta Abdulloh [16] yang membandingkan KNN dengan metode regresi lainnya turut memberikan dasar bagi perlunya pengujian algoritma alternatif untuk memperoleh gambaran performa yang lebih komprehensif. Oleh karena itu, perbandingan dengan algoritma lain dapat menjadi langkah pengembangan untuk mengetahui kesesuaian metode terhadap karakteristik data penjualan *paving block*.

Secara keseluruhan, kontribusi penelitian ini terletak pada penerapan KNN Regression dengan fitur historis dan pembagian data secara kronologis untuk menghasilkan simulasi estimasi volume penjualan *paving block*. Simulasi periode Januari hingga Desember 2026 memberikan gambaran awal mengenai kemungkinan perubahan permintaan berdasarkan pola data historis yang tersedia. Meskipun demikian, penelitian masih memiliki keterbatasan berupa jumlah data yang relatif terbatas, periode pengujian yang hanya mencakup Oktober hingga Desember 2025, serta belum digunakannya variabel eksternal seperti kondisi proyek konstruksi, perubahan permintaan pasar, dan faktor musiman lainnya. Oleh karena itu, hasil simulasi belum dapat dijadikan dasar tunggal dalam menentukan keputusan produksi maupun persediaan perusahaan. Penelitian selanjutnya disarankan memperluas dataset, menambahkan variabel eksternal, menerapkan validasi berbasis beberapa periode waktu, serta membandingkan KNN Regression dengan algoritma regresi lainnya. Pengembangan tersebut diperlukan untuk memperoleh model dengan kemampuan generalisasi yang lebih baik sebelum diterapkan sebagai sistem peramalan penjualan operasional.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini telah menerapkan algoritma *K-Nearest Neighbor* (KNN) *Regression* untuk melakukan estimasi volume penjualan *paving block* berdasarkan data transaksi historis CV Surya Perkasa. Tahapan penelitian meliputi *preprocessing data*, transformasi data berbasis waktu, pembentukan fitur historis menggunakan LAG_1 dan LAG_12, pembagian dataset secara kronologis, serta pemodelan KNN *Regression*. Berdasarkan hasil pengujian parameter, konfigurasi terbaik diperoleh pada nilai K=5. Evaluasi kinerja model menggunakan metrik regresi menghasilkan nilai MAE sebesar 331,36, MSE sebesar 160.281,61, RMSE sebesar 400,35, MAPE sebesar 36,33%, dan nilai R^2 sebesar 0,56. Hasil tersebut menunjukkan bahwa model mampu menangkap sebagian pola historis volume penjualan, meskipun masih terdapat perbedaan antara hasil estimasi dan nilai aktual pada beberapa periode pengujian. Simulasi estimasi volume penjualan 12 bulan ke depan memberikan gambaran mengenai pola perubahan volume penjualan berdasarkan data historis yang tersedia. Namun, hasil simulasi tersebut masih bersifat estimasi awal dan belum dapat digunakan sebagai dasar keputusan operasional secara langsung karena belum mempertimbangkan faktor eksternal seperti perubahan permintaan pasar, kondisi proyek, maupun faktor lain yang memengaruhi fluktuasi penjualan.

5.2. Saran

Berdasarkan hasil penelitian, pengembangan selanjutnya dapat dilakukan dengan menambahkan variabel pendukung yang berpotensi memengaruhi volume penjualan, seperti faktor musiman, tren permintaan, kondisi proyek, serta faktor eksternal lainnya. Penambahan periode dan jumlah data historis juga diperlukan agar model dapat mengenali pola perubahan penjualan secara lebih baik. Selain itu, penelitian selanjutnya dapat mempertimbangkan penggunaan metode atau kombinasi algoritma lain sebagai pembanding untuk memperoleh model dengan kemampuan generalisasi yang lebih baik. Pengembangan sistem estimasi juga

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

dapat diarahkan pada integrasi data secara berkala sehingga hasil estimasi dapat diperbarui sesuai dengan kondisi penjualan terbaru.

DAFTAR PUSTAKA

[1] A. Munthe and B. Ulya, "Penerapan Data Mining untuk Estimasi Penjualan Roti Terlaris menggunakan Metode K-Nearest Neighbor," Bulletin of Information Technology (BIT), vol. 6, no. 2, pp. 71–78, 2025, doi: 10.47065/bit.v5i2.1783.

[2] Indra Nasution, Rezkinah Rambe, Khairil Putra, Muhammad Dafa, and Muhammad Syahputra Novelan, "Application of the K-Nearest Neighbor Algorithm in the Data Mining Process to Predict Drug Sales at Pratama Haji Medan-Pancing Clinic," Jurnal Publikasi Ilmu Komputer dan Multimedia, vol. 4, no. 2, pp. 136–147, May 2025, doi: 10.55606/jupikom.v4i2.4081.

[3] Y. Bahtiar, J. Maulindar, and M. E. Yuliana, "Estimasi Kualitas Genteng Mantili Berdasarkan Komposisi Bahan Baku menggunakan Algoritma K-Nearest Neighbor," JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), vol. 9, no. 4, pp. 2030–2041, Nov. 2024, doi: 10.29100/jipi.v9i4.5589.

[4] F. Wijaya, R. Rahmansyah, and N. A. Hasibuan, "Estimasi Penjualan Barang menggunakan Algoritma K-Nearest Neighbor," 2025.

[5] A. Putri N.A, W. Hadikristanto, and Edora, "Predicting Consumer Demand Based on Retail Stock Using the K-Nearest Neighbors Algorithm," International Journal Software Engineering and Computer Science (IJSECS), vol. 4, no. 2, pp. 523–534, Aug. 2024, doi: 10.35870/ijsecs.v4i2.2865.

[6] "Sektor Konstruksi Punya Peran Besar ke Ekonomi, Serap 8,7 Juta Tenaga Kerja." Accessed: May 30, 2026. [Online]. Available: https://finance.detik.com/infrastruktur/d-8068264/sektor-konstruksi-punya-peran-besar-ke-ekonomi-serap-8-7-juta-tenaga-kerja

[7] M. Bakriansyah, A. Subair, and W. Setiawan, "Penerapan Data Mining untuk Memestimasi Penjualan Produk Terlaris UD Timbul Jaya Menggunakan Metode K-Nearest Neighbor," Jurnal Cendekia Ilmiah, vol. 4, no. 2, 2025.

[8] K. Gautama and R. Arijanto, "Penerapan Data Mining untuk Estimasi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor Pada Cv. Nuansa Karuna Lestari," 2024.

[9] A. Kurniawan, D. Mustikasari, and A. T. P. Raharjo, "Implementasi Metode Regresi Linear Berganda untuk Estimasi Harga Penjualan Material Paving block pada CV. Difa Jaya Abadi," Sains Data Jurnal Studi Matematika dan Teknologi, vol. 3, no. 2, pp. 89–96, Nov. 2025, doi: 10.52620/sainsdata.v3i2.281.

[10] N. Surojudin and M. Danny, "Penerapan Data Mining dengan Algoritma C4.5 dan K-Nearest Neighbor untuk Estimasi Penjualan Bahan Bangunan Terlaris," Jurnal Informatika Ekonomi Bisnis, pp. 672–679, Sep. 2025, doi: 10.37034/infeb.v7i3.1241.

[11] A. Elnaeem Balila and A. Bin Shabri, "Comparative Analysis of Machine Learning Algorithms for Predicting Dubai Property Prices," Front. Appl. Math. Stat., vol. 10, 2024, doi: 10.3389/fams.2024.1327376.

[12] S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining untuk Estimasi Penjualan Produk Terlaris menggunakan Metode K-Nearest Neighbor," Building of Informatics, Technology and Science (BITS), vol. 3, no. 4, pp. 639–648, Mar. 2022, doi: 10.47065/bits.v3i4.1408.

[13] A. Alif, "Penerapan K-Nearst Neighbor Regression dalam Memestimasi Penjualan Produk di Berdikari Kopi," Da'watuna: Journal of Communication and Islamic Broadcasting, vol. 4, no. 4, Jun. 2024, doi: 10.47467/dawatuna.v4i4.2078.

[14] R. Salkiawati, H. Lubis, A. M. Syah Al Malik, "Analisis Estimasi Penjualan Produk menggunakan Algoritma K-Nearest Neighbor Regression," Journal of Information System, Informatics and Computing Issue Period, vol. 9, no. 2, pp. 350–262, 2025, doi: 10.52362/jisicom.v9i2.2132.

[15] M. Hayya', F. M. Hana, and A. Prihandono, "Perbandingan Metode K-Nearest Neighbor (KNN) dan Random Forest dalam Memestimasi Harga Rumah," Sainteks, vol. 22, no. 1, pp. 99–108, Apr. 2025, doi: 10.30595/sainteks.v22i1.25998.

[16] D. Yudhistira Ananda and F. Fauzi Abdulloh, "Comparative Analysis of K-Nearest Neighbors

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

Algorithm and Random Forest Regressor for House Price Prediction in Bandung City," 2026. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[17] Nadia Putri Ariyanti, Agung Triayudi, and Ratih Titi Komala Sari, "Analysis of K-NN Algorithm and Linear Regression to Predict House Prices in Jabodetabek," SaNa: Journal of Blockchain, NFTs and Metaverse Technology, vol. 2, no. 1, pp. 65–71, Feb. 2024, doi: 10.58905/sana.v2i1.265.

[18] V. B and R. Gangula, "Exploring the Power and Practical Applications of K-Nearest Neighbours (KNN) in Machine Learning," Journal of Computer Allied Intelligence, vol. 2, no. 1, pp. 8–15, Feb. 2024, doi: 10.69996/jcai.2024002.

[19] Y. Bahtiar, J. Maulindar, and M. E. Yuliana, "Estimasi Kualitas Genteng Mantili Berdasarkan Komposisi Bahan Baku menggunakan Algoritma K-Nearest Neighbor," JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), vol. 9, no. 4, pp. 2030–2041, Nov. 2024, doi: 10.29100/jipi.v9i4.5589.

[20] N. Surojudin and M. Danny, "Penerapan Data Mining dengan Algoritma C4.5 dan K-Nearest Neighbor untuk Estimasi Penjualan Bahan Bangunan Terlaris," Jurnal Informatika Ekonomi Bisnis, pp. 672–679, Sep. 2025, doi: 10.37034/infeb.v7i3.1241.

[21] D. Kurniati, T. Saputro, E. F. Nurhidayatullah, C. D. Saputro, and A. Asyifa, "Kekuatan Tekan Paving Block dengan Memanfaatkan Limbah Las Asetelin," 2021.

[22] "Inilah 6 Model Paving Block yang Menjadi Primadona Di Indonesia!" Accessed: May 30, 2026. [Online]. Available: https://indonusa-conblock.com/model-paving-block-di-indonesia/

[23] D. Nurfauzan and T. Fatimah, "Implementasi Algoritma K-Nearest Neighbors Regression dalam Memestimasi Harga Saham," 2022. [Online]. Available: https://senafti.budiluhur.ac.id/index.php/

[24] Y. Sun, B. Pfahringer, H. M. Gomes, and A. Bifet, "SOKNL: A Novel Way of Integrating K-Nearest Neighbours with Adaptive Random Forest Regression for Data Streams," Data Min. Knowl. Discov., vol. 36, no. 5, pp. 2006–2032, Sep. 2022, doi: 10.1007/s10618-022-00858-9.

[25] V. Ariyani, P. Putri, A. B. Prasetijo, and D. Eridani, "Perbandingan Kinerja Algoritme Naïve Bayes Dan K-Nearest Neighbor (KNN) untuk Estimasi Harga Rumah", doi: 10.14710/transmisi.24.4.162-171.

[26] S. J. Wetzel, "Twin Neural Network Improved K-Nearest Neighbor Regression," Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.00664