

Penerapan Algoritma Regresi Linear Berganda untuk Prediksi Produksi Rokok pada Pabrik Rokok CARI

¹Santowi Azis, ²Otong Saeful Bachri, ³Bambang Irawan
^{1,2,3}Teknik Informatika, Universitas Muhadi Setiabudi, Brebes

E-mail: ¹azissantowi@gmail.com, ²otongsaifulbahriumus@gmail.com,
³bambangumus@gmail.com

ABSTRAK

Perencanaan produksi merupakan salah satu aspek penting dalam industri manufaktur karena berperan dalam menentukan jumlah produksi yang sesuai dengan kebutuhan pasar dan kondisi operasional perusahaan. Di Pabrik Rokok CARI, proses penentuan jumlah produksi masih dilakukan berdasarkan kondisi stok yang tersedia tanpa memanfaatkan data historis secara optimal, sehingga berpotensi menimbulkan ketidaksesuaian antara jumlah produksi dan kebutuhan aktual. Penelitian ini bertujuan menerapkan metode Regresi Linear Berganda untuk memprediksi produksi rokok berdasarkan data historis penjualan, stok, dan jam kerja. Penelitian menggunakan 352 data harian yang dibagi menjadi 281 data *training* (80%) dan 71 data *testing* (20%). Hasil pemodelan menghasilkan persamaan regresi $Y = -24,0255 + (0,3838 \times X_1) + (0,0015 \times X_2) + (78,0807 \times X_3)$, dengan X_1 merupakan penjualan, X_2 merupakan stok, dan X_3 merupakan jam kerja. Evaluasi model menunjukkan nilai koefisien determinasi (R^2) sebesar 87,51%, *Mean Squared Error* (MSE) sebesar 10.688,72, *Root Mean Squared Error* (RMSE) sebesar 103,39, dan *Mean Absolute Error* (MAE) sebesar 73,22. Hasil tersebut menunjukkan bahwa metode Regresi Linear Berganda mampu memberikan tingkat akurasi yang baik dalam memprediksi produksi rokok sehingga dapat dimanfaatkan sebagai dasar pendukung perencanaan produksi. Model prediksi kemudian diimplementasikan dalam aplikasi berbasis web untuk mempermudah proses pengolahan data historis dan penyajian hasil prediksi kepada pihak manajemen.

Kata kunci : Regresi Linear Berganda, Prediksi Produksi, Perencanaan Produksi, Data Historis, Industri Rokok.

ABSTRACT

Production planning is an essential aspect of the manufacturing industry as it determines the appropriate production quantity to meet market demand and operational conditions. At Pabrik Rokok CARI, production planning is still primarily based on available inventory without optimal utilization of historical data, which may lead to discrepancies between production output and actual demand. This study aims to apply the Multiple Linear Regression method to predict cigarette production using historical data on sales, inventory, and working hours. The dataset consisted of 352 daily records, which were divided into 281 training data (80%) and 71 testing data (20%). The resulting regression model was $Y = -24.0255 + (0.3838 \times X_1) + (0.0015 \times X_2) + (78.0807 \times X_3)$, where X_1 represents sales, X_2 represents inventory, and X_3 represents working hours. Model evaluation produced a coefficient of determination (R^2) of 87.51%, a Mean Squared Error (MSE) of 10,688.72, a Root Mean Squared Error (RMSE) of 103.39, and a Mean Absolute Error (MAE) of 73.22. These results indicate that the Multiple Linear Regression method provides good predictive performance for estimating cigarette production and can serve as a reliable basis for supporting production planning. The prediction model was further implemented in a web-based application to facilitate historical data processing and provide production prediction results for management decision support.

Keyword : Multiple Linear Regression, Production Prediction, Production Planning, Historical Data, Cigarette Industry.

1. PENDAHULUAN

perencanaan produksi merupakan salah satu aspek penting dalam industri manufaktur karena berperan dalam menentukan jumlah produksi yang sesuai dengan kebutuhan pasar serta kondisi operasional perusahaan. perencanaan produksi yang tepat dapat membantu perusahaan menjaga keseimbangan antara jumlah produk yang dihasilkan dengan tingkat permintaan sehingga risiko kelebihan maupun kekurangan produksi dapat diminimalkan. oleh karena itu, perusahaan memerlukan suatu pendekatan yang mampu menghasilkan prediksi produksi secara lebih objektif berdasarkan data historis (Athira & Herlina, 2022).

pabrik rokok cari merupakan perusahaan manufaktur yang bergerak di bidang produksi rokok dan secara rutin melakukan pencatatan data produksi, penjualan, stok, serta aktivitas operasional harian. berdasarkan hasil observasi, penentuan jumlah produksi masih dilakukan dengan mempertimbangkan kondisi stok yang tersedia dan pengalaman dalam proses produksi. pendekatan tersebut belum memanfaatkan data historis secara optimal sehingga peluang untuk memperoleh informasi prediksi yang lebih akurat belum dapat dimanfaatkan secara maksimal.

data historis yang dimiliki perusahaan sebenarnya mengandung informasi yang dapat digunakan untuk membangun model prediksi produksi. variabel seperti penjualan, stok, dan jam kerja diduga memiliki hubungan terhadap jumlah produksi yang dihasilkan setiap harinya. dengan memanfaatkan hubungan antarvariabel tersebut, perusahaan dapat memperoleh gambaran mengenai jumlah produksi yang diperlukan pada periode berikutnya sehingga proses perencanaan produksi dapat dilakukan secara lebih terukur.

salah satu metode yang dapat digunakan untuk memodelkan hubungan antara beberapa variabel independen terhadap satu variabel dependen adalah *regresi linear berganda*. metode ini mampu membentuk persamaan matematis berdasarkan data historis sehingga dapat digunakan untuk

memperkirakan nilai produksi pada periode selanjutnya. selain memiliki proses perhitungan yang relatif sederhana, regresi linear berganda juga banyak digunakan dalam penelitian prediksi karena mampu menjelaskan pengaruh masing-masing variabel independen terhadap variabel dependen (Technology et al., 2025).

berdasarkan permasalahan tersebut, penelitian ini menerapkan algoritma regresi linear berganda untuk membentuk model prediksi produksi rokok harian berdasarkan variabel penjualan, stok, dan jam kerja. model yang dihasilkan kemudian dievaluasi menggunakan parameter mean absolute error (mae), mean squared error (mse), root mean squared error (rmse), dan koefisien determinasi (r^2) untuk mengetahui tingkat akurasi model dalam memprediksi jumlah produksi rokok.

Penelitian mengenai prediksi telah banyak dilakukan dengan berbagai metode sesuai karakteristik data yang digunakan. Penelitian oleh Oliver Januardi Ababil dkk. menerapkan metode Regresi Linear untuk memprediksi penjualan *liquid vape*. Hasil penelitian menunjukkan bahwa regresi linear mampu menghasilkan prediksi yang cukup baik berdasarkan data historis penjualan. Namun, penelitian tersebut hanya menggunakan satu variabel bebas dan berfokus pada prediksi penjualan. (Ababil et al., 2022)

Penelitian ini bertujuan untuk membentuk model prediksi produksi rokok harian menggunakan Algoritma Regresi Linear Berganda berdasarkan variabel penjualan, stok, dan jam kerja pada Pabrik Rokok CARI. Selain itu, penelitian ini bertujuan mengevaluasi performa model menggunakan parameter MAE, MSE, RMSE, dan Koefisien Determinasi (R^2) sehingga dapat diketahui tingkat akurasi model dalam memprediksi jumlah produksi rokok.

2. LANDASAN TEORI

2.1 Forecasting

Prediksi (forecasting) merupakan proses memperkirakan suatu kondisi atau nilai yang akan terjadi pada masa mendatang

berdasarkan data historis dan pola yang telah terjadi sebelumnya. Dalam bidang manufaktur, prediksi produksi digunakan sebagai dasar untuk menentukan jumlah produksi yang sesuai dengan kebutuhan pasar sehingga perusahaan dapat mengurangi risiko kelebihan maupun kekurangan produksi. Hasil prediksi yang akurat membantu perusahaan dalam menyusun perencanaan produksi, mengoptimalkan penggunaan sumber daya, serta meningkatkan efisiensi operasional (Adiyatma & Rohman, 2025).

Prediksi produksi umumnya memanfaatkan data historis yang memiliki hubungan dengan jumlah produksi, seperti data penjualan, persediaan, maupun faktor operasional lainnya. Dengan memanfaatkan informasi tersebut, perusahaan dapat menghasilkan estimasi jumlah produksi yang lebih objektif dibandingkan dengan penentuan produksi berdasarkan perkiraan semata.

2.2 Regresi Linear Berganda

Regresi Linear Berganda merupakan salah satu metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan dua atau lebih variabel independen. Metode ini bertujuan membentuk suatu model matematis yang dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan perubahan nilai variabel independennya. Dalam penelitian ini, variabel dependen berupa jumlah produksi rokok, sedangkan variabel independen terdiri atas penjualan, stok, dan jam kerja (Syahfitri et al., 2025).

2.3 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui tingkat akurasi hasil prediksi yang dihasilkan oleh model Regresi Linear Berganda. Pada penelitian ini digunakan empat parameter evaluasi, yaitu:

1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) digunakan untuk mengukur rata-rata kesalahan absolut antara nilai aktual dan nilai prediksi. Semakin kecil nilai *MAE*, semakin baik performa model dalam melakukan prediksi.

2. Mean Squared Errors (MSE)

Mean Squared Error (MSE) merupakan metode evaluasi yang digunakan untuk mengukur rata-rata kuadrat selisih antara nilai aktual dan nilai hasil prediksi. Nilai *MSE* yang semakin kecil menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang semakin rendah. Persamaan *MSE* ditunjukkan pada Persamaan

3. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) merupakan metrik evaluasi yang digunakan untuk mengukur besarnya kesalahan prediksi dengan memberikan penalti lebih besar terhadap kesalahan yang tinggi. Nilai *RMSE* yang lebih rendah menunjukkan bahwa hasil prediksi semakin mendekati nilai aktual.

4. Koefisien Determinasi (R^2)

Koefisien determinasi digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data aktual. Nilai R^2 berada pada rentang 0 sampai 1. Semakin mendekati 1 menunjukkan bahwa model memiliki kemampuan yang semakin baik dalam menjelaskan variasi data aktual (Song et al., 2025)

2.4 Penelitian Terdahulu

Penelitian mengenai prediksi telah banyak dilakukan dengan memanfaatkan berbagai metode sesuai karakteristik data dan kebutuhan pengguna. Penelitian oleh Ababil et al. (2022) menerapkan metode Regresi Linear untuk memprediksi penjualan *liquid vape* berbasis website. Penelitian tersebut menggunakan data historis penjualan sebagai dasar pembentukan model regresi. Hasil penelitian menunjukkan bahwa metode Regresi Linear mampu menghasilkan prediksi yang cukup baik dan membantu proses pengambilan keputusan dalam pengelolaan penjualan. Namun, penelitian tersebut hanya menggunakan satu variabel independen sehingga belum mampu menggambarkan pengaruh beberapa faktor terhadap hasil prediksi secara bersamaan (Ababil et al., 2022).

Penelitian lain yang mengombinasikan metode Linear Regression dan Weighted Moving Average (WMA) untuk memprediksi penjualan produk. Penelitian tersebut menunjukkan bahwa kombinasi kedua metode mampu meningkatkan akurasi prediksi dibandingkan penggunaan satu metode saja. Meskipun demikian, penelitian lebih berfokus pada prediksi penjualan sehingga belum mengevaluasi hubungan beberapa variabel operasional terhadap proses produksi (Adiyatma & Rohman, 2025)

Selanjutnya, penelitian oleh Susilo et al. (2024) menerapkan algoritma K-Nearest Neighbor (KNN) dalam prediksi penjualan produk digital. Hasil penelitian menunjukkan bahwa metode KNN mampu memberikan hasil prediksi yang baik pada data dengan karakteristik tertentu. Akan tetapi, metode tersebut sangat dipengaruhi oleh pemilihan nilai parameter k dan memiliki performa yang menurun apabila jumlah data semakin besar (Susilo et al., 2024)

Penelitian lain dilakukan oleh Al-Faruq dan Hindarto (2024) yang menerapkan metode Fuzzy Mamdani untuk memprediksi produksi rokok klobot berdasarkan beberapa variabel masukan. Hasil penelitian menunjukkan bahwa metode Fuzzy Mamdani mampu menghasilkan prediksi produksi yang cukup baik berdasarkan aturan (*rule base*) yang telah ditentukan. Namun, proses penyusunan aturan fuzzy memerlukan pengetahuan pakar sehingga tingkat akurasi model sangat bergantung pada kualitas aturan yang digunakan (Al-faruq & Hindarto, 2024)

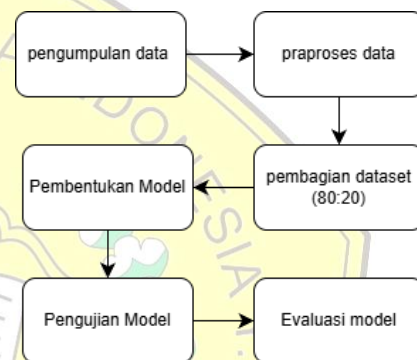
Berdasarkan beberapa penelitian terdahulu tersebut, metode Regresi Linear telah banyak digunakan dalam bidang prediksi, baik pada penjualan maupun hasil produksi. Namun, sebagian besar penelitian masih menggunakan satu variabel independen atau berfokus pada objek penelitian yang berbeda. Selain itu, penelitian yang secara khusus menerapkan Regresi Linear Berganda menggunakan variabel penjualan, stok, dan jam kerja untuk memprediksi produksi rokok harian serta mengevaluasi performa model menggunakan MAE, MSE, RMSE, dan koefisien determinasi (R^2) masih terbatas. Oleh karena itu, penelitian ini menerapkan Regresi Linear Berganda untuk membangun

model prediksi produksi rokok harian dan mengevaluasi kemampuan model dalam menghasilkan prediksi berdasarkan data historis.

3. METODOLOGI

3.1 Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan data, praproses data, pembagian dataset menjadi data training dan data testing, pembentukan model Regresi Linear Berganda menggunakan data training, serta evaluasi model menggunakan data testing.



Gambar 1. Tahapan penelitian

3.2 Pengumpulan Data

Penelitian ini menggunakan data historis produksi harian yang diperoleh dari Pabrik Rokok CARI sebanyak 352 data. Data tersebut merupakan hasil pencatatan operasional perusahaan yang terdiri dari variabel penjualan, stok, jam kerja, dan produksi. Variabel penjualan, stok, dan jam kerja digunakan sebagai variabel independen, sedangkan produksi digunakan sebagai variabel dependen dalam pembentukan model Regresi Linear Berganda.

Tabel 1. Variabel Penelitian

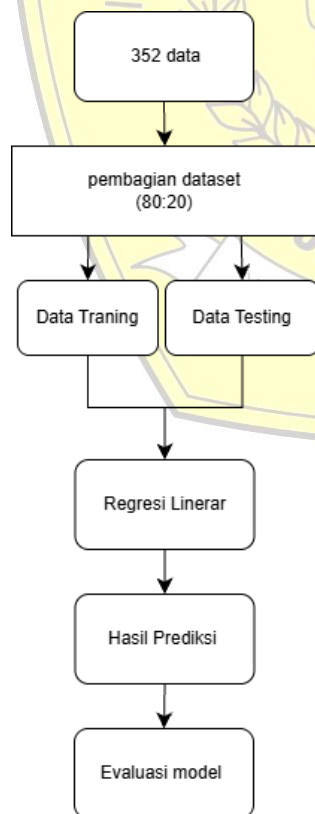
Variabel	Keterangan
Y	Produksi
X ₁	Penjualan
X ₂	Stok
X ₃	Jam Kerja

3.3 Praproses Data

Tahap praproses data dilakukan untuk memastikan data yang digunakan dalam pembentukan model memiliki kualitas yang baik. Pada tahap ini dilakukan pemeriksaan terhadap data historis untuk memastikan tidak terdapat data kosong (*missing value*), data duplikat, maupun data yang tidak sesuai. Selanjutnya, data diurutkan berdasarkan tanggal pencatatan sehingga urutan data tetap sesuai dengan kondisi aktual sebelum dilakukan pembagian menjadi data training dan data testing.

3.4 Pembagian Dataset

Dataset dibagi menggunakan metode Hold-Out Validation, yaitu sebesar 80% sebagai data training dan 20% sebagai data testing. Pembagian dilakukan secara berurutan berdasarkan waktu pencatatan data (*time-based split*) sehingga model dibangun menggunakan data yang lebih lama dan diuji menggunakan data yang lebih baru. Dari total 352 data, sebanyak 281 data digunakan sebagai data training dan 71 data digunakan sebagai data testing.



Gambar 2. Pembagian Dataset

3.5 Pembentukan Model

Pembentukan model dilakukan menggunakan data training untuk memperoleh hubungan antara variabel penjualan (X_1), stok (X_2), dan jam kerja (X_3) terhadap jumlah produksi (Y). Model Regresi Linear Berganda dinyatakan menggunakan Persamaan

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 \quad (1)$$

Nilai konstanta (a) dan koefisien regresi (b_1 , b_2 , dan b_3) dihitung menggunakan metode *Ordinary Least Squares* (OLS) dengan menyusun persamaan normal berdasarkan data training. Persamaan normal yang digunakan adalah sebagai berikut.

$$\sum Y = na + b_1 \sum x_1 + b_2 \sum x_2 + b_3 \sum x_3 \quad (2)$$

$$\sum X_1 Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + b_3 \sum X_1 X_3 \quad (3)$$

$$\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 X_3 \quad (4)$$

$$\sum X_3 Y = a \sum X_3 + b_1 \sum X_1 X_3 + b_2 \sum X_2 X_3 + b_3 \sum X_3^2 \quad (5)$$

Selanjutnya, seluruh persamaan diselesaikan untuk memperoleh nilai konstanta dan koefisien regresi sehingga terbentuk model Regresi Linear Berganda. Model yang diperoleh kemudian digunakan untuk menghitung nilai prediksi pada data testing dan dievaluasi menggunakan parameter MAE, MSE, RMSE, dan koefisien determinasi (R_2).

3.6 Evaluasi Model

evaluasi model dilakukan menggunakan data testing untuk mengetahui kemampuan model dalam memprediksi jumlah produksi rokok. Kinerja model diukur menggunakan empat parameter evaluasi, yaitu:

1. Mean Absolute Error (MAE)

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

2. Mean Square Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

3. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

4. Koefisien Determinasi (R²)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

Nilai MAE, MSE dan RMSE digunakan untuk mengukur tingkat kesalahan prediksi, sedangkan nilai R² digunakan untuk mengetahui kemampuan model dalam menjelaskan variasi data aktual.

4. HASIL DAN PEMBAHASAN

4.1 Pembentukan Model Regresi Linear

Pembentukan model dilakukan menggunakan 281 data training yang diperoleh dari pembagian dataset sebesar 80% dari total 352 data historis produksi. Variabel independen yang digunakan dalam penelitian ini meliputi penjualan (X₁), stok (X₂), dan jam kerja (X₃), sedangkan variabel dependen adalah produksi (Y). Proses pembentukan model dilakukan menggunakan metode Regresi Linear Berganda dengan menghitung nilai konstanta dan koefisien regresi berdasarkan persamaan normal.

Berdasarkan hasil perhitungan menggunakan data training, diperoleh persamaan Regresi Linear Berganda sebagai berikut.

$$Y = -24.0255 + 0.3838X_1 + 0.0015X_2 + 78.0807X_3$$

Tabel 2 Hasil Koefisien Regresi Linear

Parameter	nilai
Konsanta (a)	-24,0255
Penjualan (b ₁)	0,3838
Stok (b ₂)	0,0015
Jam kerja (b ₃)	78,0807

Berdasarkan hasil pelatihan menggunakan 281 data training, diperoleh model Regresi Linear Berganda yang ditunjukkan pada Persamaan (7). Persamaan tersebut selanjutnya digunakan untuk memprediksi jumlah produksi rokok pada 71 data testing dan dievaluasi menggunakan nilai MAE, MSE, RMSE, serta koefisien determinasi (R₂).

4.2 Hasil Prediksi

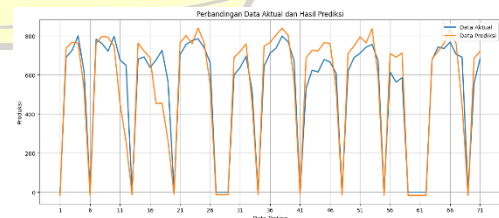
Setelah diperoleh persamaan Regresi Linear Berganda, model digunakan untuk memprediksi jumlah produksi menggunakan 71 data testing yang tidak dilibatkan pada proses pembentukan model. Proses prediksi dilakukan dengan memasukkan nilai variabel penjualan (X₁), stok (X₂), dan jam kerja (X₃) ke dalam persamaan regresi yang telah diperoleh sehingga dihasilkan nilai prediksi produksi untuk setiap data pengujian.

Tabel 3. Data Testing

X1	X2	X3	Y
561	5671	6	600
701	5710	6	649
701	5658	6	688
870	5645	6	790
670	5564	6	809

Tabel 4. hasil prediksi

Data	Aktual	Prediksi
1	600	668
2	649	722
3	688	722
4	790	787
5	809	710

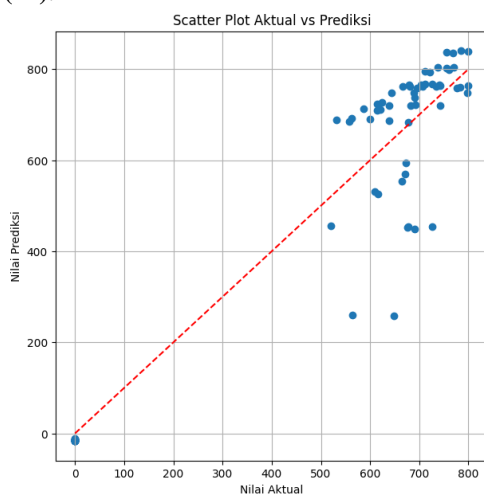


Gambar 3. Grafik Aktual dan prediksi

Gambar 3 menunjukkan perbandingan antara data aktual dan hasil prediksi produksi rokok pada 71 data testing.

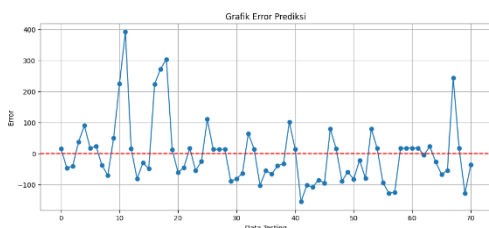
4.3 Evaluasi Model

Evaluasi model dilakukan menggunakan 71 data testing yang tidak dilibatkan dalam proses pembentukan model. Tujuan evaluasi adalah untuk mengetahui kemampuan model Regresi Linear Berganda dalam memprediksi jumlah produksi rokok berdasarkan variabel penjualan, stok, dan jam kerja. Parameter evaluasi yang digunakan meliputi Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan Koefisien Determinasi (R^2).



Gambar 4. Scatter Nilai Aktual Vs Prediksi

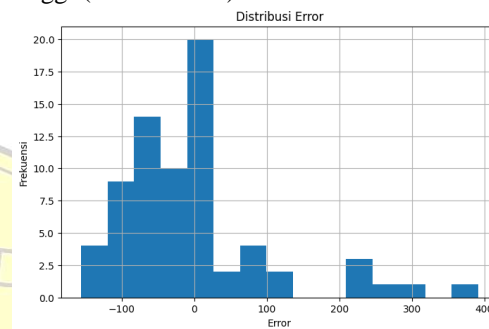
berdasarkan Gambar 4, terlihat bahwa sebagian besar titik data berada di sekitar garis diagonal (garis merah putus-putus), yang menunjukkan bahwa nilai prediksi yang dihasilkan oleh model *Regresi Linear Berganda* memiliki kedekatan dengan nilai aktual. Semakin dekat posisi titik terhadap garis diagonal, semakin kecil selisih antara nilai prediksi dan nilai aktual sehingga menunjukkan tingkat akurasi model yang semakin baik.



Gambar 5. Grafik Error Prediksi

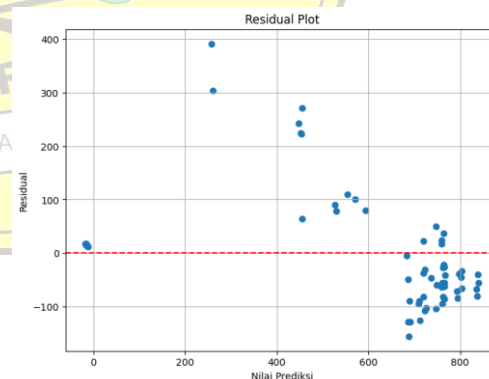
Gambar 5 menunjukkan distribusi error prediksi yang dihasilkan oleh model *Regresi Linear Berganda* pada 71 data testing.

Nilai error diperoleh dari selisih antara nilai aktual dan nilai prediksi untuk setiap data pengujian. Nilai error yang berada di atas garis nol menunjukkan bahwa model menghasilkan prediksi yang lebih rendah (*underestimate*) dibandingkan nilai aktual, sedangkan nilai error di bawah garis nol menunjukkan bahwa model menghasilkan prediksi yang lebih tinggi (*overestimate*).



Gambar 6. Distribusi Error

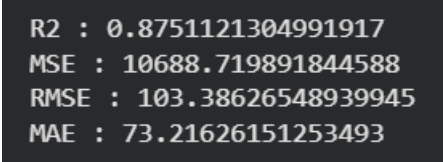
Gambar 6 menunjukkan distribusi error hasil prediksi model *Regresi Linear Berganda* pada data testing. Sebagian besar nilai error terkonsentrasi di sekitar angka nol, yang menunjukkan bahwa hasil prediksi model relatif mendekati nilai aktual. Meskipun terdapat beberapa data dengan nilai error yang cukup besar, frekuensinya relatif sedikit sehingga tidak memberikan pengaruh yang signifikan terhadap performa model secara keseluruhan.



Gambar 7. Residual Plot

Gambar 7 menunjukkan *residual plot* antara nilai prediksi dan residual pada data testing. Sebagian besar residual berada di sekitar garis nol, yang menunjukkan bahwa hasil prediksi model relatif mendekati nilai aktual. Meskipun masih terdapat beberapa residual yang cukup besar pada beberapa data,

secara umum tidak terlihat pola tertentu yang menunjukkan penyimpangan sistematis. Hal ini mengindikasikan bahwa model Regresi Linear Berganda memiliki kemampuan yang baik dalam memprediksi produksi rokok harian.



```

R2 : 0.8751121304991917
MSE : 10688.719891844588
RMSE : 103.38626548939945
MAE : 73.21626151253493
  
```

Gambar 8. Hasil Evaluasi Model

Berdasarkan hasil evaluasi, model memperoleh nilai MAE sebesar 73,22, yang menunjukkan bahwa rata-rata selisih antara nilai prediksi dan nilai aktual adalah sekitar 73 bungkus. Nilai MSE sebesar 10.688,72 dan RMSE sebesar 103,39 menunjukkan bahwa tingkat kesalahan prediksi model masih berada pada batas yang dapat diterima. Selain itu, nilai koefisien determinasi R^2 sebesar 87,51% menunjukkan bahwa variabel penjualan, stok, dan jam kerja mampu menjelaskan 87,51% variasi produksi rokok, sedangkan 12,49% sisanya dipengaruhi oleh faktor lain yang tidak dimasukkan ke dalam model. Hasil tersebut menunjukkan bahwa model Regresi Linear Berganda memiliki kemampuan yang baik dalam memprediksi produksi rokok harian.

4.4 Pembahasan

Berdasarkan hasil pembentukan model, Regresi Linear Berganda mampu menghasilkan persamaan yang menggambarkan hubungan antara variabel penjualan, stok, dan jam kerja terhadap jumlah produksi rokok harian. Persamaan yang diperoleh menunjukkan bahwa ketiga variabel tersebut memiliki pengaruh positif terhadap jumlah produksi, meskipun besarnya pengaruh setiap variabel berbeda. Variabel jam kerja memiliki koefisien terbesar, yaitu 78,0807, yang menunjukkan bahwa peningkatan jam kerja memberikan kontribusi paling besar terhadap peningkatan jumlah produksi. Sementara itu, variabel penjualan juga memberikan pengaruh positif terhadap produksi, sedangkan variabel stok memiliki pengaruh yang relatif kecil dibandingkan dua variabel lainnya.

Hasil evaluasi model menunjukkan nilai MAE sebesar 73,22, MSE sebesar 10.688,72, RMSE sebesar 103,39, dan koefisien determinasi R^2 sebesar 87,51%. Nilai tersebut menunjukkan bahwa model mampu menjelaskan sebagian besar variasi data produksi berdasarkan ketiga variabel yang digunakan. Dengan nilai R^2 sebesar 87,51%, dapat disimpulkan bahwa model memiliki kemampuan yang baik dalam memprediksi produksi rokok harian berdasarkan data historis.

Meskipun demikian, nilai R^2 yang belum mencapai 100% menunjukkan bahwa masih terdapat faktor lain yang memengaruhi jumlah produksi namun belum dimasukkan ke dalam model. Faktor-faktor tersebut dapat berupa jumlah tenaga kerja yang tersedia, kondisi mesin produksi, permintaan pasar, ketersediaan bahan baku, maupun kebijakan operasional perusahaan. Oleh karena itu, penelitian selanjutnya dapat mempertimbangkan penambahan variabel-variabel tersebut agar model prediksi yang dihasilkan memiliki tingkat akurasi yang lebih tinggi.

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, metode *Regresi Linear Berganda* berhasil diterapkan untuk membangun model prediksi produksi rokok harian di Pabrik Rokok CARI menggunakan variabel penjualan, stok, dan jam kerja sebagai variabel independen. Berdasarkan proses pelatihan menggunakan 281 data training, diperoleh persamaan regresi $Y = -24,0255 + 0,3838X_1 + 0,0015X_2 + 78,0807X_3$ yang menunjukkan hubungan antara ketiga variabel tersebut terhadap jumlah produksi. Hasil evaluasi model menggunakan 71 data testing menunjukkan nilai MAE sebesar 73,22, MSE sebesar 10.688,72, RMSE sebesar 103,39, dan koefisien determinasi (R^2) sebesar 87,51%. Nilai evaluasi tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam memprediksi produksi rokok harian berdasarkan data historis. Dengan demikian, metode Regresi Linear Berganda dapat digunakan sebagai salah satu pendekatan dalam memperkirakan jumlah produksi sehingga dapat membantu perusahaan dalam menyusun perencanaan produksi yang lebih terukur dan berbasis data.

6. DAFTAR PUSTAKA

- Ababil, O. J., Wibowo, S. A., & Industri, F. T. (2022). *PENERAPAN METODE REGRESI LINIER DALAM PREDIKSI PENJUALAN LIQUID VAPE DI TOKO VAPOR PANDAAN BERBASIS WEBSITE*. 6(1), 186–195.
- Adiyatma, D. A., & Rohman, M. G. (2025). *Product Sales Prediction System at Starmart Using Linear Regression and Weighted Moving Average Methods*. 4(3), 660–673.
- Al-faruq, R. N., & Hindarto, H. (2024). *Prediksi Produksi Rokok Klobot Menggunakan Metode Logika Fuzzy Mamdani*. 2, 1–14.
- Athira, N., & Herlina, M. (2022). *Identifikasi Faktor yang Mempengaruhi Data Driven Decision pada Pemerintah Desa Menggunakan SEM GSCA*. 145–152.
- Kustiyo, A., Mukhlis, & Suharso, A. (2022). *Model Recurrent Neural Network untuk Peramalan Produksi Tebu Nasional*. *Bina Insani ICT Journal*, 9(1), 1–10.
- Song, C., Liu, T., & Ning, W. (2025). *Wheat Yield Prediction Based on Parallel CNN-LSTM-Attention with Transfer Learning Model*. *Agriculture*, 15, 2519.
- Susilo, B., Ramdhan, N. A., & Bachri, O. S. (2024). *Application of the K-Nearest Neighbor Algorithm for Predicting Digital Product Sales Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Produk Digital*. 4(October), 1466–1476.
- Syahfitri, A. A., Informatika, P. S., Bina, U., Informatika, S., Timur, J., Linear, R., & Timur, J. (2025). *Prediksi Pertumbuhan Penduduk Kota Jakarta Timur Menggunakan Metode Regresi Linear*. 8(4), 2109–2118.
- Technology, N., Rasuli, AmaSerli, R. K., Indah, K., Ningtyas, P., Yozha, S., Kinanti, P., Mandiri, U. N., Bina, U., & Informatika, S. (2025). *Prediksi Stok Produk Susu Pada PT Greenfields Dairy Indonesia Menggunakan Metode Regresi Linear*. 5(1), 23–28.