

## Evaluasi Kinerja Model Long Short-Term Memory Pada Prediksi Produksi Bawang Merah Kabupaten Brebes

<sup>1</sup>Ainan Zaky Nurrofiq, <sup>2</sup>Nur Ariesanto Ramdhan, <sup>3</sup>Otong Saeful Bachri  
<sup>1,2,3</sup>Teknik Informatika, Universitas Muhadi Setiabudi, Brebes

E-mail: <sup>1</sup>[chasperzaky@gmail.com](mailto:chasperzaky@gmail.com), <sup>2</sup>[ariesantoramdhan@gmail.com](mailto:ariesantoramdhan@gmail.com),  
<sup>3</sup>[otongsaifulbahriumus@gmail.com](mailto:otongsaifulbahriumus@gmail.com)

### ABSTRAK

Produksi bawang merah merupakan komoditas hortikultura penting yang berkontribusi terhadap ketahanan pangan dan perekonomian daerah, khususnya di Kabupaten Brebes sebagai salah satu sentra produksi terbesar di Indonesia. Fluktuasi produksi akibat perubahan kondisi cuaca dan faktor pertanian menyebabkan perlunya metode prediksi yang akurat. Penelitian ini bertujuan mengevaluasi kinerja metode *Long Short-Term Memory (LSTM)* dalam memprediksi produksi bawang merah di Kabupaten Brebes. Data yang digunakan meliputi produksi, luas panen, curah hujan, dan suhu periode 2016–2025 sebanyak 1.288 data. Tahapan penelitian meliputi *preprocessing* data, *feature engineering*, normalisasi menggunakan *Min-Max Scaling*, pembentukan model *LSTM*, dan evaluasi menggunakan *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, serta *koefisien determinasi (R<sup>2</sup>)*. Model *LSTM* dibangun dengan dua lapisan *LSTM* berukuran 64 dan 32 neuron, *dropout* 0,2, serta satu *dense layer* berukuran 16 neuron. Hasil pengujian menunjukkan nilai *MAE* sebesar 9.478,55 kuintal, *RMSE* sebesar 15.763,86 kuintal, dan *R<sup>2</sup>* sebesar 0,8990. Hasil tersebut menunjukkan bahwa model mampu menjelaskan 89,90% variasi data produksi sehingga metode *LSTM* memiliki kinerja yang sangat baik dan berpotensi digunakan sebagai pendukung pengambilan keputusan dalam perencanaan produksi bawang merah di Kabupaten Brebes.

**Kata kunci :** *Long Short-Term Memory, prediksi produksi, bawang merah, time series, deep learning.*

### ABSTRACT

Shallot production is one of the most important horticultural commodities that contributes to national food security and regional economic development, particularly in Brebes Regency, one of the largest shallot-producing areas in Indonesia. Production fluctuations caused by weather conditions and agricultural factors necessitate the use of accurate forecasting methods. This study aims to evaluate the performance of the *Long Short-Term Memory (LSTM)* method in predicting shallot production in Brebes Regency. The dataset consisted of production, harvested area, rainfall, and temperature data from 2016 to 2025, comprising 1,288 records. The research process included data preprocessing, feature engineering, *Min-Max Scaling* normalization, *LSTM* model development, and performance evaluation using *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, and the coefficient of determination (*R<sup>2</sup>*). The proposed model employed two *LSTM* layers with 64 and 32 neurons, a dropout rate of 0.2, and a dense layer with 16 neurons. Experimental results showed that the model achieved an *MAE* of 9,478.55 quintals, an *RMSE* of 15,763.86 quintals, and an *R<sup>2</sup>* value of 0.8990. These results indicate that the model can explain 89.90% of the variation in shallot production data. Therefore, the *LSTM* model demonstrates excellent predictive performance and has the potential to support decision-making in agricultural production planning in Brebes Regency.

**Keyword :** *Long Short-Term Memory, production prediction, shallot, time series, deep learning.*

## 1. PENDAHULUAN

Bawang merah (*Allium ascalonicum* L.) merupakan salah satu komoditas hortikultura strategis yang memiliki peran penting dalam memenuhi kebutuhan pangan masyarakat Indonesia. Selain digunakan sebagai bahan utama dalam berbagai produk pangan, bawang merah juga menjadi sumber pendapatan bagi petani dan berkontribusi terhadap perekonomian daerah. Kabupaten Brebes dikenal sebagai sentra produksi bawang merah terbesar di Indonesia yang menghasilkan produksi dalam jumlah besar setiap tahunnya. Namun, produksi bawang merah sering mengalami fluktuasi akibat berbagai faktor yang memengaruhi pertumbuhan dan produktivitas tanaman. Kondisi tersebut menyebabkan perlunya upaya untuk memprediksi produksi bawang merah secara lebih akurat guna mendukung perencanaan produksi dan pengambilan keputusan di sektor pertanian (Jabed & Murad, 2024).

Produksi bawang merah dipengaruhi oleh berbagai faktor, baik faktor internal maupun eksternal. Faktor eksternal yang sering menjadi perhatian adalah kondisi cuaca, seperti curah hujan dan suhu udara. Curah hujan yang terlalu tinggi dapat menyebabkan genangan air dan meningkatkan risiko serangan penyakit tanaman, sedangkan suhu yang tidak sesuai dapat menghambat proses pertumbuhan dan pembentukan umbi. Selain itu, luas panen juga memiliki hubungan langsung dengan jumlah produksi yang dihasilkan. Oleh karena itu, diperlukan metode prediksi yang mampu memanfaatkan data historis produksi dan variabel cuaca untuk menghasilkan estimasi produksi yang lebih baik (Shook et al., 2021).

Perkembangan teknologi informasi telah mendorong pemanfaatan kecerdasan buatan (*Artificial Intelligence*) dalam berbagai bidang, termasuk pertanian. Salah satu teknologi yang banyak digunakan adalah *Machine Learning* yang mampu mempelajari pola dari data historis untuk menghasilkan prediksi secara otomatis. Penggunaan *Machine Learning* dan *Deep Learning* pada berbagai sektor menunjukkan peningkatan kemampuan dalam menyelesaikan

permasalahan prediksi yang kompleks dibandingkan pendekatan statistik konvensional (Janiesch et al., 2021)

Dalam bidang pertanian, teknologi *Deep Learning* telah digunakan untuk berbagai kebutuhan, seperti prediksi hasil panen, klasifikasi tanaman, identifikasi penyakit tanaman, serta analisis kondisi lingkungan. Kemampuan *Deep Learning* dalam mempelajari hubungan nonlinier antarvariabel menjadikannya sebagai salah satu pendekatan yang banyak digunakan dalam pengembangan sistem pertanian modern berbasis data (Rakhmatulin, 2021). Sejalan dengan perkembangan tersebut, konsep *smart agriculture* mulai diterapkan dengan memanfaatkan data cuaca, kondisi lingkungan, dan teknologi kecerdasan buatan untuk meningkatkan efisiensi serta produktivitas sektor pertanian (Ali et al., 2024).

Salah satu metode *Deep Learning* yang banyak digunakan dalam prediksi data deret waktu (*time series*) adalah *Long Short-Term Memory (LSTM)*. Metode ini merupakan pengembangan dari *Recurrent Neural Network (RNN)* yang dirancang untuk mengatasi permasalahan *vanishing gradient* dan mampu mempelajari hubungan jangka panjang pada data sekuensial. Kemampuan tersebut menjadikan *LSTM* banyak digunakan pada berbagai penelitian prediksi karena dapat menangkap pola temporal yang terdapat dalam data historis secara lebih efektif (Mukhlis et al., 2021).

Beberapa penelitian terdahulu telah dilakukan untuk memprediksi produksi bawang merah menggunakan berbagai metode. Maulana et al. (2023) menerapkan metode *Regresi Linier Sederhana* pada data produksi bawang merah Kabupaten Brebes dan memperoleh nilai *koefisien determinasi (R<sup>2</sup>)* sebesar 0,98. Sementara itu, Sunariadi et al. (2022) menggunakan metode *SARIMA* untuk memprediksi produksi bawang merah di Kabupaten Nganjuk dan memperoleh nilai *MAPE* sebesar 2,01%. Meskipun kedua metode tersebut mampu menghasilkan prediksi yang baik, pendekatan tersebut masih memiliki keterbatasan dalam menangkap

hubungan *nonlinier* dan ketergantungan jangka panjang pada data runtun waktu. Oleh karena itu, penelitian ini menggunakan metode *Long Short-Term Memory (LSTM)* yang memiliki kemampuan mempelajari pola temporal yang kompleks pada data produksi pertanian.

Penerapan metode *LSTM* tidak hanya digunakan dalam bidang pertanian, tetapi juga pada berbagai kasus prediksi lainnya. Penelitian yang dilakukan oleh Herawati dan Prastiti (2024) menunjukkan bahwa metode *LSTM* mampu menghasilkan prediksi yang baik pada data *time series* jumlah usaha mikro, kecil, dan menengah (UMKM). Hasil tersebut menunjukkan bahwa *LSTM* memiliki kemampuan yang baik dalam mempelajari pola historis dan menghasilkan prediksi yang stabil pada berbagai jenis data.

Beberapa penelitian terdahulu telah menunjukkan keberhasilan penggunaan jaringan saraf berulang (*Recurrent Neural Network*) dan *LSTM* dalam bidang prediksi. Aziz Kustiyo dan Mukhlis (2022) menerapkan metode RNN untuk peramalan produksi tebu nasional dan memperoleh hasil yang menunjukkan kemampuan model dalam mempelajari pola historis produksi. Selain itu, Pravesh et al. (2024) mengembangkan model prediksi hasil panen berbasis *Deep Learning* dengan mempertimbangkan pengaruh perubahan iklim dan menunjukkan peningkatan performa dibandingkan pendekatan konvensional. (Pravesh et al., 2024)

Berdasarkan berbagai penelitian tersebut, metode *LSTM* memiliki potensi yang baik untuk diterapkan pada prediksi produksi bawang merah. Namun, penelitian yang secara khusus mengevaluasi kinerja model *LSTM* dalam memprediksi produksi bawang merah di Kabupaten Brebes masih relatif terbatas. Oleh karena itu, penelitian ini dilakukan untuk mengevaluasi kinerja metode *Long Short-Term Memory (LSTM)* dalam memprediksi produksi bawang merah di Kabupaten Brebes menggunakan data produksi, luas panen, curah hujan, dan suhu periode 2016–2025. Evaluasi dilakukan menggunakan metrik *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan koefisien determinasi ( $R^2$ ) untuk mengetahui

kemampuan model dalam menghasilkan prediksi yang akurat.

## 2. LANDASAN TEORI

### 2.1 Produksi Bawang Merah

Bawang merah (*Allium ascalonicum L.*) merupakan salah satu komoditas hortikultura yang memiliki nilai ekonomi tinggi dan berperan penting dalam memenuhi kebutuhan pangan masyarakat Indonesia. Produksi bawang merah dipengaruhi oleh berbagai faktor, seperti luas panen, kondisi iklim, kualitas benih, kesuburan tanah, serta teknik budidaya yang digunakan. Perubahan kondisi lingkungan dan iklim dapat menyebabkan fluktuasi produksi sehingga diperlukan upaya prediksi yang akurat untuk mendukung pengambilan keputusan di sektor pertanian (Jabed & Murad, 2024).

### 2.2 Data Time Series

Data *time series* merupakan data yang disusun berdasarkan urutan waktu tertentu sehingga setiap pengamatan memiliki keterkaitan dengan pengamatan sebelumnya. Karakteristik utama data *time series* adalah adanya pola tren, musiman (*seasonality*), siklus, dan fluktuasi yang terjadi dari waktu ke waktu. Dalam bidang pertanian, data produksi tanaman termasuk ke dalam data *time series* karena jumlah produksi pada suatu periode dipengaruhi oleh kondisi pada periode sebelumnya. Oleh karena itu, analisis *time series* banyak digunakan untuk melakukan prediksi terhadap hasil produksi pada periode mendatang (Plaat, 2023).

### 2.3 Machine Learning

*Machine Learning* merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan komputer mempelajari pola dari data dan menghasilkan prediksi tanpa diprogram secara eksplisit. Algoritma *machine learning* memanfaatkan data historis untuk membangun model yang dapat digunakan dalam proses prediksi maupun pengambilan keputusan. Saat ini, *machine learning* telah banyak diterapkan pada berbagai bidang, termasuk kesehatan, ekonomi, pendidikan, dan pertanian (Janiesch et al., 2021).

### 2.4 Deep Learning

*Deep Learning* merupakan bagian dari *machine learning* yang menggunakan jaringan saraf tiruan (*Artificial Neural Network*)

dengan banyak lapisan tersembunyi (*hidden layer*) untuk mempelajari pola yang kompleks dalam data. Metode ini memiliki kemampuan yang baik dalam mengolah data berukuran besar dan menemukan hubungan nonlinier antarvariabel. *Deep learning* mampu melakukan ekstraksi fitur secara otomatis sehingga dapat meningkatkan performa model pada berbagai permasalahan prediksi dibandingkan metode konvensional. (Zhang et al., 2023)

Selain itu, perkembangan teori dan matematika pada *deep learning* menunjukkan bahwa jaringan saraf modern mampu memodelkan hubungan yang kompleks serta menangani data berdimensi tinggi dengan tingkat akurasi yang baik. Kemampuan tersebut menjadikan *deep learning* sebagai salah satu pendekatan yang banyak digunakan dalam penelitian prediksi berbasis data (*data-driven prediction*) (Jentzen et al., 2025).

## 2.5 Long Short-Term Memory

*Long Short-Term Memory (LSTM)* merupakan salah satu pengembangan dari *Recurrent Neural Network (RNN)* yang dirancang untuk mengatasi permasalahan *vanishing gradient* pada proses pelatihan jaringan saraf. LSTM memiliki kemampuan menyimpan informasi dalam jangka panjang melalui mekanisme memori yang terdiri atas *forget gate*, *input gate*, dan *output gate*. Dengan mekanisme tersebut, model mampu mempelajari hubungan temporal pada data *time series* secara lebih efektif dibandingkan *RNN* konvensional (Mukhlis et al., 2021).

*LSTM* banyak digunakan pada berbagai penelitian prediksi karena mampu menangani data sekuensial dan mempertahankan informasi penting dari periode sebelumnya. Penelitian Herawati dan Prastiti (2024) menunjukkan bahwa metode *LSTM* mampu menghasilkan prediksi yang baik pada data *time series* sehingga menjadi salah satu metode yang banyak digunakan dalam berbagai bidang penelitian.

## 2.6 Feature Engineering

*Feature engineering* merupakan proses pembentukan variabel baru dari data asli untuk meningkatkan kemampuan model dalam mempelajari pola yang terdapat pada data.

Pada penelitian ini digunakan beberapa fitur tambahan seperti *lag feature*, *rolling mean*, fitur musiman, serta variabel cuaca dan luas panen. Penggunaan fitur-fitur tersebut bertujuan untuk memberikan informasi historis kepada model sehingga hubungan temporal pada data dapat dipelajari dengan lebih baik.

Penerapan teknik *feature engineering* terbukti mampu meningkatkan performa model berbasis LSTM pada berbagai penelitian prediksi. Pemanfaatan informasi historis melalui fitur tambahan dapat membantu model menghasilkan prediksi yang lebih akurat dibandingkan hanya menggunakan data asli (Amin et al., 2025).

## 2.7 Normalisasi Min-Max

Normalisasi merupakan proses transformasi data ke dalam rentang tertentu agar seluruh variabel memiliki skala yang seragam. Pada penelitian ini digunakan metode Min-Max Scaling karena sesuai untuk model berbasis jaringan saraf tiruan.

## 2.8 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.

## 2.9 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.

## 2.10 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)

## 2.11 Penelitian Terdahulu

Penelitian terkait prediksi produksi bawang merah telah dilakukan oleh berbagai peneliti dengan menggunakan metode yang berbeda. Penelitian oleh Maulana et al. (2023) menggunakan metode *Regresi Linier* Sederhana untuk memprediksi hasil produksi panen bawang merah di Kabupaten Brebes. Dataset yang digunakan berasal dari Dinas Pertanian dan Ketahanan Pangan Kabupaten Brebes. Hasil penelitian menunjukkan nilai *MAE* sebesar 29.051, *RMSE* sebesar 45.533, dan koefisien determinasi ( $R^2$ ) sebesar 0,98. Penelitian tersebut menunjukkan bahwa metode regresi linier mampu digunakan untuk memprediksi produksi bawang merah, namun masih memiliki keterbatasan dalam menangkap pola nonlinier pada data produksi pertanian. (Alfin Maulana, Martanto, 2023)

Penelitian lain dilakukan oleh Sunariadi et al. (2022) yang menerapkan metode *Seasonal Autoregressive Integrated Moving Average (SARIMA)* untuk memprediksi produksi bawang merah di Kabupaten Nganjuk. Penelitian menggunakan data runtun waktu produksi bawang merah periode 2017–2020 yang diperoleh dari Badan Pusat Statistik Kabupaten Nganjuk. Hasil penelitian menunjukkan bahwa model SARIMA (3,0,2)(2,1,2)<sub>12</sub> merupakan model terbaik dengan nilai *MAPE* sebesar 2,01%, sehingga mampu memberikan hasil prediksi yang sangat akurat terhadap pola musiman produksi bawang merah. (Sunariadi et al., 2022)

Penelitian lain dilakukan oleh Syabrian et al. (2025) yang menerapkan metode *Weighted Moving Average (WMA)* untuk meramalkan produksi bawang merah di Provinsi Jawa Tengah menggunakan data produksi periode 2013–2023. Hasil penelitian menunjukkan bahwa metode WMA mampu menghasilkan peramalan produksi hingga tahun 2028 dengan nilai *Mean Absolute Error (MAE)* sebesar 18.193,6 ton. Penelitian tersebut menunjukkan bahwa metode WMA dapat digunakan untuk melakukan peramalan produksi bawang merah, namun masih bergantung pada pola data historis dan belum mampu mempelajari hubungan nonlinier maupun pengaruh variabel eksternal secara mendalam. (Syabrian et al., 2025)

Selanjutnya, Jati et al. (2024) menerapkan metode *Fuzzy Mamdani* untuk memprediksi produksi bawang merah di Kota Yogyakarta dengan memanfaatkan variabel curah hujan, suhu, dan kelembapan. Hasil penelitian menunjukkan bahwa metode tersebut menghasilkan nilai *RMSE* sebesar 14,11%. Meskipun mampu memberikan hasil prediksi yang cukup baik, penelitian tersebut menyarankan perlunya pengembangan metode yang memiliki kemampuan lebih tinggi dalam mempelajari pola data yang kompleks agar tingkat akurasi prediksi dapat ditingkatkan. (Jati et al., 2023)

Berdasarkan beberapa penelitian terdahulu tersebut, berbagai metode seperti *Regresi Linier*, *SARIMA*, *Weighted Moving Average*, dan *Fuzzy Mamdani* telah digunakan untuk memprediksi produksi bawang merah. Namun, metode-metode tersebut memiliki keterbatasan dalam memodelkan hubungan nonlinier serta ketergantungan jangka panjang pada data deret waktu. Oleh karena itu, penelitian ini menerapkan metode *Long Short-Term Memory (LSTM)* dengan memanfaatkan data produksi, luas panen, curah hujan, suhu, serta fitur historis hasil *feature engineering*. Pendekatan tersebut diharapkan mampu menghasilkan prediksi yang lebih akurat karena dapat mempelajari pola temporal dan hubungan antarvariabel secara lebih efektif.

## 3. METODOLOGI

### 3.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *Deep Learning* menggunakan algoritma *Long Short-Term Memory (LSTM)*. Metode ini dipilih karena memiliki kemampuan dalam mempelajari pola historis pada data *time series* dan mampu menangkap hubungan nonlinier antarvariabel yang memengaruhi produksi bawang merah. Evaluasi model dilakukan menggunakan metrik *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan koefisien determinasi ( $R^2$ ) untuk mengetahui tingkat akurasi prediksi yang dihasilkan.

### 3.2 Data Penelitian

Data yang digunakan dalam penelitian merupakan data sekunder yang terdiri atas data produksi bawang merah, luas panen, curah hujan, dan suhu. Dataset mencakup periode Januari 2016 sampai Desember 2025

yang berasal dari 12 kecamatan di Kabupaten Brebes. Setelah dilakukan proses pembersihan data (*data cleaning*) dan penghapusan nilai kosong (*missing value*), diperoleh sebanyak 1.288 data yang digunakan dalam penelitian.

### 3.3 Tahapan Penelitian

Penelitian dilakukan melalui beberapa tahapan yang meliputi pengumpulan data, praproses data, pembentukan fitur (*feature engineering*), normalisasi, pembangunan model *LSTM*, pelatihan model, pengujian model, dan evaluasi hasil prediksi.

### 3.4 Data Preprocessing

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses pelatihan model. Tahapan ini meliputi:

#### 1. Konversi Tanggal

Variabel tanggal diubah ke format *datetime* untuk memudahkan proses ekstraksi informasi waktu.

#### 2. Pengurutan Data

Data diurutkan berdasarkan kecamatan dan tanggal agar urutan data sesuai dengan karakteristik *time series*.

#### 3. Penanganan Missing Value

Nilai kosong yang muncul akibat proses pembentukan fitur historis dihapus menggunakan metode *dropna()*.

#### 4. One Hot Encoding

Variabel kecamatan diubah menjadi data numerik menggunakan teknik *One Hot Encoding* agar dapat digunakan sebagai input model.

### 3.5 Feature Engineering

Untuk meningkatkan kemampuan model dalam mempelajari pola produksi bawang merah, dilakukan pembentukan beberapa fitur tambahan.

#### 1. Fitur Musiman

Fitur musiman digunakan untuk menangkap pola periodik tahunan yang terjadi pada data produksi.

Table 1. Fitur Musiman

Fitur	Keterangan
bulan	Bulan pengamatan
tahun	Tahun pengamatan
quarter	Kuartal

bulan_sin	Transformasi sinus bulan
bulan_cos	Transformasi cosinus bulan

#### 2. Lag Feature

Lag feature digunakan untuk memberikan informasi produksi pada periode sebelumnya.

Table 2. Lag Feature

Fitur
lag_1
lag_3
lag_6
lag_12

#### 3. Lag Variabel Pendukung

Table 3. Lag Variabel Pendukung

Fitur
panen_lag1
hujan_lag1
suhu_lag1

#### 4. Rolling Mean

Fitur rolling digunakan untuk menangkap tren produksi rata-rata beberapa periode sebelumnya.

Table 4. Rolling Mean

Fitur
rolling_3
rolling_6
rolling_12

Penggunaan fitur-fitur tersebut bertujuan untuk membantu model memahami pola temporal yang terdapat pada data produksi bawang merah (Mukhlis et al., 2021).

### 3.6 Pembagian Data

Dataset dibagi menjadi dua kelompok berdasarkan periode waktu.

Table 5. Pembagian Data

Dataset	Periode
Training	2016–2024
Testing	2025

Data training digunakan untuk melatih model *LSTM*, sedangkan data testing digunakan untuk mengevaluasi kemampuan model dalam memprediksi data yang belum pernah dilihat sebelumnya.

### 3.7 Normalisasi Data

Normalisasi dilakukan menggunakan metode Min-Max Scaling untuk mengubah rentang data ke interval 0–1. Metode ini dipilih karena dapat meningkatkan stabilitas proses pelatihan jaringan saraf dan mempercepat konvergensi model (Kustiyo et al., 2022)

### 3.8 Arsitektur Model LSTM

Model yang digunakan dalam penelitian ini merupakan model *Long Short-Term Memory (LSTM)* bertingkat (*stacked LSTM*).

Table 6. Arsitektur Model

Layer	Jumlah Neuron
LSTM Layer 1	64
Dropout	0,2
LSTM Layer 2	32
Dropout	0,2
Dense Layer	16
Output Layer	1

Fungsi aktivasi yang digunakan pada *Dense Layer* adalah ReLU (*Rectified Linear Unit*), sedangkan optimizer yang digunakan adalah Adam.

### 3.9 Evaluasi Model

Kinerja model dievaluasi menggunakan tiga metrik utama.

1. Mean Absolute Error (MAE)

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

2. Root Mean Square Error (RMSE)

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

3. Koefisien Determinasi ( $R^2$ )

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

Nilai MAE dan RMSE digunakan untuk mengukur tingkat kesalahan prediksi, sedangkan nilai  $R^2$  digunakan untuk mengetahui kemampuan model dalam menjelaskan variasi data aktual.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Deskripsi Dataset

Penelitian ini menggunakan dataset produksi bawang merah Kabupaten Brebes

periode Januari 2016 hingga Desember 2025. Dataset terdiri atas data produksi, luas panen, curah hujan, dan suhu dari 12 kecamatan yang menjadi sentra produksi bawang merah di Kabupaten Brebes. Setelah dilakukan proses *feature engineering* dan penghapusan data yang mengandung nilai kosong (*missing value*), jumlah data yang digunakan dalam penelitian sebanyak 1.288 data.

Data kemudian dibagi menjadi dua kelompok berdasarkan urutan waktu. Data periode 2016–2024 digunakan sebagai data pelatihan (*training data*), sedangkan data tahun 2025 digunakan sebagai data pengujian (*testing data*). Pembagian data berdasarkan urutan waktu dilakukan untuk mempertahankan karakteristik data *time series* sehingga proses evaluasi model lebih mencerminkan kondisi prediksi pada data aktual.

### 4.2 Hasil Preprocessing dan Feature Engineering

Tahap *preprocessing* dilakukan dengan mengubah format tanggal menjadi tipe *datetime*, mengurutkan data berdasarkan kecamatan dan waktu, serta menghapus data yang mengandung nilai kosong. Selanjutnya dilakukan proses *feature engineering* untuk menghasilkan variabel tambahan yang dapat membantu model memahami pola historis produksi bawang merah.

Fitur yang digunakan dalam penelitian terdiri atas variabel produksi pendukung seperti luas panen, curah hujan, suhu, fitur musiman (*bulan*, *quarter*, *bulan\_sin*, *bulan\_cos*), *lag feature* (*lag\_1*, *lag\_3*, *lag\_6*, *lag\_12*), serta *rolling mean* (*rolling\_3*, *rolling\_6*, *rolling\_12*). Selain itu, variabel kecamatan dikonversi menjadi data numerik menggunakan metode *One Hot Encoding*.

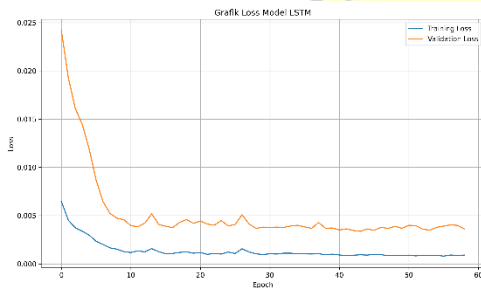
Penggunaan *lag feature* dan *rolling mean* bertujuan untuk memberikan informasi historis kepada model sehingga hubungan temporal pada data produksi dapat dipelajari dengan lebih baik. Menurut Mukhlis et al. (2021), penggunaan fitur historis dapat meningkatkan kemampuan model dalam memprediksi data produksi pertanian yang memiliki pola musiman dan tren jangka panjang.

### 4.3 Hasil Pelatihan Model LSTM

Model *Long Short-Term Memory (LSTM)* dibangun menggunakan dua lapisan

LSTM dengan jumlah neuron 64 dan 32. Untuk mengurangi risiko *overfitting*, digunakan dua lapisan *Dropout* dengan nilai 0,2. Selanjutnya digunakan *Dense Layer* sebanyak 16 neuron dan satu neuron keluaran untuk menghasilkan nilai prediksi produksi bawang merah.

Proses pelatihan dilakukan menggunakan *optimizer Adam* dengan fungsi kerugian (*loss function*) Mean Squared Error (MSE). Pelatihan dilakukan hingga maksimum 100 epoch dengan *batch size* 32 dan *validation split* sebesar 20%. Selain itu, diterapkan metode *Early Stopping* dengan nilai *patience* 15 untuk menghentikan pelatihan secara otomatis ketika nilai *validation loss* tidak mengalami perbaikan.

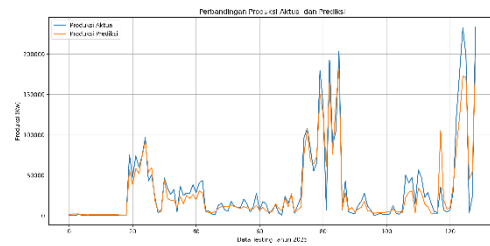


Gambar 1. Grafik Loss

Berdasarkan Gambar 1, terlihat bahwa nilai *training loss* mengalami penurunan yang signifikan pada awal proses pelatihan dan kemudian cenderung stabil hingga akhir pelatihan. Nilai *validation loss* juga menunjukkan pola yang relatif stabil tanpa adanya peningkatan yang drastis. Kondisi tersebut menunjukkan bahwa model mampu mempelajari pola data dengan baik dan tidak mengalami *overfitting* yang signifikan.

#### 4.4 Hasil Prediksi Produksi Bawang Merah

Setelah proses pelatihan selesai, model digunakan untuk memprediksi produksi bawang merah pada data pengujian tahun 2025. Hasil prediksi kemudian dibandingkan dengan data aktual untuk mengetahui kemampuan model dalam mengikuti pola perubahan produksi bawang merah.



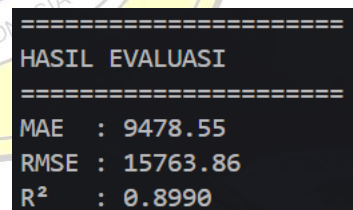
Gambar 2. Grafik Aktual dan Prediksi Produksi

Berdasarkan Gambar 2, terlihat bahwa hasil prediksi memiliki pola yang relatif mengikuti data aktual. Sebagian besar titik prediksi berada dekat dengan pola produksi aktual, baik pada periode produksi rendah maupun periode produksi tinggi. Meskipun masih terdapat beberapa perbedaan pada nilai tertentu, model secara umum mampu menangkap tren produksi bawang merah dengan baik.

Kemampuan model dalam mengikuti pola data menunjukkan bahwa metode LSTM mampu memanfaatkan informasi historis dan variabel cuaca untuk menghasilkan prediksi yang akurat. Hasil ini sejalan dengan penelitian Shook et al. (2021) yang menyatakan bahwa kombinasi data historis dan variabel cuaca dapat meningkatkan performa model prediksi hasil pertanian.

#### 4.5 Evaluasi Kinerja Model

Evaluasi model dilakukan menggunakan tiga metrik, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi ( $R^2$ ).



Gambar 3. Hasil evaluasi

Berdasarkan Gambar 3, model memperoleh nilai MAE sebesar 9.478,55 kuintal. Nilai tersebut menunjukkan bahwa rata-rata selisih absolut antara hasil prediksi dan data aktual sebesar 9.478,55 kuintal. Semakin kecil nilai MAE menunjukkan semakin baik kemampuan model dalam menghasilkan prediksi.

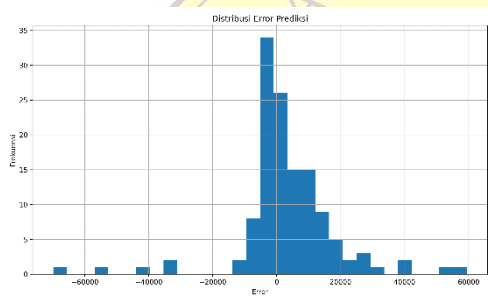
Nilai RMSE sebesar 15.763,86 kuintal menunjukkan bahwa model memiliki tingkat

kesalahan prediksi yang relatif rendah. Nilai RMSE yang lebih besar dibandingkan MAE menunjukkan adanya beberapa data yang memiliki kesalahan prediksi lebih tinggi dibandingkan data lainnya, namun secara keseluruhan kesalahan prediksi masih berada pada tingkat yang dapat diterima.

Sementara itu, nilai koefisien determinasi ( $R^2$ ) sebesar 0,8990 menunjukkan bahwa model mampu menjelaskan 89,90% variasi data produksi bawang merah. Nilai tersebut menunjukkan hubungan yang sangat kuat antara hasil prediksi dan data aktual sehingga model dapat dikategorikan memiliki performa yang sangat baik.

#### 4.6 Analisis Distribusi Error

Distribusi error digunakan untuk mengetahui penyebaran kesalahan prediksi yang dihasilkan model.

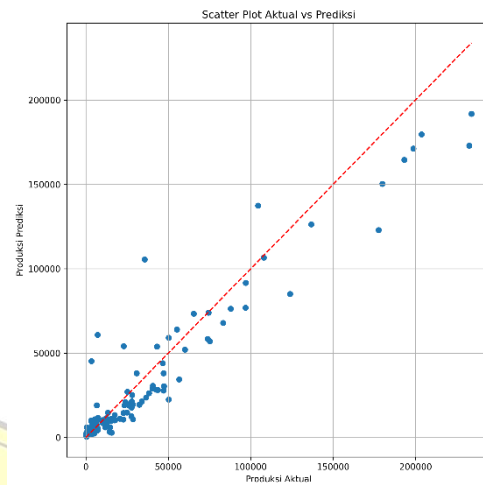


Gambar 4. Grafik Error

Gambar 4 menunjukkan distribusi error hasil prediksi model LSTM pada data pengujian. Sebagian besar nilai error berada di sekitar angka 0, yang menunjukkan bahwa hasil prediksi memiliki selisih yang relatif kecil terhadap data aktual. Meskipun terdapat beberapa error yang cukup besar akibat fluktuasi produksi yang tinggi, jumlahnya relatif sedikit sehingga tidak memengaruhi kinerja model secara keseluruhan. Hal ini didukung oleh nilai MAE sebesar 9.478,55, RMSE sebesar 15.763,86, dan  $R^2$  sebesar 0,8990, yang menunjukkan bahwa model memiliki kemampuan prediksi yang baik terhadap produksi bawang merah Kabupaten Brebes.

#### 4.7 Analisis Scatter Plot Aktual dan Prediksi

Scatter plot digunakan untuk melihat hubungan antara data aktual dan data hasil prediksi



Gambar 5. Scatter Plot Aktual dan Prediksi

Pada Gambar 5 terlihat bahwa sebagian besar titik berada di sekitar garis diagonal. Kondisi tersebut menunjukkan bahwa nilai prediksi yang dihasilkan model memiliki kedekatan yang tinggi dengan data aktual. Semakin dekat titik terhadap garis diagonal, semakin kecil kesalahan prediksi yang dihasilkan model.

Sebaran titik yang mengikuti garis diagonal mendukung hasil evaluasi yang menunjukkan nilai  $R^2$  sebesar 0,8990. Dengan demikian, model dapat dikatakan memiliki kemampuan yang sangat baik dalam memprediksi produksi bawang merah di Kabupaten Brebes

### 5. KESIMPULAN

Penelitian ini berhasil menerapkan metode Long Short-Term Memory (LSTM) untuk memprediksi produksi bawang merah di Kabupaten Brebes menggunakan data produksi, luas panen, curah hujan, dan suhu periode 2016–2025. Model menghasilkan nilai MAE sebesar 9.478,55 kuintal, RMSE sebesar 15.763,86 kuintal, dan  $R^2$  sebesar 0,8990 yang menunjukkan kemampuan sangat baik dalam menjelaskan variasi data produksi. Hasil penelitian menunjukkan bahwa metode LSTM efektif digunakan untuk prediksi produksi bawang merah dan berpotensi menjadi alat pendukung pengambilan keputusan pada sektor pertanian.

### 6. DAFTAR PUSTAKA

- Alfin Maulana, Martanto, I. A. (2023). *Prediksi Hasil Produksi Panen Bawang Merah Menggunakan Metode Regresi Linier Sederhana*. 7(4), 2884–2888.
- Ali, T., Rehman, S. U., Ali, S., & Mahmood, K. (2024). *Smart agriculture: Utilizing machine learning and deep learning for drought stress identification in crops*.
- Amin, M. D. I., Hidayat, J. J., & Setyowati, C. (2025). Implementasi Model LSTM untuk Peramalan Curah Hujan di Bekasi dengan Pemanfaatan Data Cuaca BMKG. *Jurnal Teknologi Informasi Digital*, 1(2).
- Jabed, M. A., & Murad, M. A. A. (2024). Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon*.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695.
- Jati, Z. N., Hastono, T., & Andrian, F. (2023). *Prediksi Produksi Bawang Merah di Kota Yogyakarta menggunakan Metode Fuzzy Mamdani*. 2(2).
- Jentzen, A., Kuckuck, B., & von Wurstemberger, P. (2025). *Mathematical Introduction to Deep Learning: Methods, Implementations, and Theory*.
- Kustiyo, A., Mukhlis, & Suharso, A. (2022). Model Recurrent Neural Network untuk Peramalan Produksi Tebu Nasional. *Bina Insani ICT Journal*, 9(1), 1–10.
- Mukhlis, Kustiyo, A., & Suharso, A. (2021). Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory. *Bina Insani ICT Journal*, 8(1), 22–32.
- Plaata, A. (2023). *Deep Reinforcement Learning*. Springer Nature.
- Pravesh, S. Y., Garg, N., Arora, R., Singh, S., & Sankari, S. S. (2024). Predictive Modeling of Crop Yield Using Deep Learning Based Transformer with Climate Change Effects. *International Research Journal of Multidisciplinary Technovation*, 6(6), 223–240.
- Rakhmatulin, I. (2021). *Deep learning, machine vision in agriculture in 2021*.
- Shook, J., Gangopadhyay, T., Wu, L., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2021). Crop yield prediction integrating genotype and weather variables using deep learning. *PLOS ONE*.
- Song, C., Liu, T., & Ning, W. (2025). Wheat Yield Prediction Based on Parallel CNN-LSTM-Attention with Transfer Learning Model. *Agriculture*, 15, 2519.
- Sunariadi, N. M., Intan, P. K., Candra, D., Novitasari, R., & Hariningsih, Y. (2022). *Prediksi Produksi Bawang Merah di Kabupaten Nganjuk Dengan Metode Seasonal Arima ( Sarima )*. 6(1), 49–60. <https://doi.org/10.36526/tr.v>
- Syabrian, I., S, B. R., Maarende, T. Y., Olivia, Y., Purba, S., & Fathoni, M. Y. (2025). *Analisis Peramalan Produksi Bawang Merah di Jawa Tengah Menggunakan Metode Weighted Moving Average*. 9(1), 1377–1382.
- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). *Dive into Deep Learning*. Cambridge University Press.