



PREDIKSI HARGA CABAI MUSIMAN MENGGUNAKAN MODEL LSTM DI JAWA TENGAH

Gunawan Gunawan^{*1}, Wresti Andriani², Nur Tulus Ujianto¹

¹Program Studi Sistem Informasi, Universitas Pancasakti Indonesia

²Program Studi Informatika, Universitas Bima Sakapenta, Indonesia

Abstrak: Penelitian ini mengembangkan model prediksi harga cabai musiman di Jawa Tengah menggunakan algoritma Long Short-Term Memory (LSTM) yang mengintegrasikan data harga harian dan data cuaca, termasuk curah hujan, suhu udara, dan kelembapan relatif. Model ini dirancang untuk memprediksi fluktuasi harga cabai yang dipengaruhi oleh faktor musiman dan kondisi cuaca, yang sering menyebabkan ketidakstabilan harga. Hasil eksperimen menunjukkan bahwa model LSTM berhasil menghasilkan prediksi dengan nilai RMSE 512,83, MAE 387,49, dan R^2 0,861, yang mengindikasikan kemampuan model dalam menangkap pola harga yang dipengaruhi oleh faktor eksternal. Keunggulan utama LSTM dibandingkan dengan model lain seperti Support Vector Regression (SVR) dan Random Forest terletak pada kemampuannya untuk menangkap korelasi temporal dan pola musiman dalam data deret waktu. Implikasi praktis dari penelitian ini meliputi penggunaan model untuk membantu petani dalam menentukan waktu tanam dan panen yang optimal serta bagi pemerintah daerah dalam mengatur distribusi dan pengendalian harga cabai untuk mengurangi dampak inflasi pangan dan meningkatkan ketahanan pangan. Penelitian ini membuka peluang untuk penelitian lanjutan yang dapat mengembangkan model yang lebih kompleks dan mengakomodasi faktor eksternal lainnya.

Kata kunci: Cabai, cuaca, fluktuasi harga, LSTM, model prediksi

I. PENDAHULUAN

Pergerakan harga cabai di Indonesia, khususnya di wilayah Jawa Tengah, merupakan salah satu komoditas yang sangat fluktuatif sehingga berdampak langsung terhadap stabilitas ekonomi serta inflasi bahan pangan (Wahyuni et al., 2024). Ketidakpastian harga cabai tidak hanya merugikan konsumen, tetapi juga petani yang kesulitan menentukan waktu tanam dan panen yang optimal. Oleh karena itu, diperlukan suatu sistem prediksi

harga yang mampu memberikan estimasi secara akurat untuk membantu pengambilan keputusan di berbagai pihak (Arisandi et al., 2025).

Pada penelitian terdahulu, beberapa metode telah digunakan untuk memprediksi harga komoditas pertanian, mulai dari metode statistik klasik seperti Regresi Linier untuk mengetahui hasil prediksi harga beras menggunakan RMSE, hasilnya 337.996 +/- 0.000 (Veri Arinal, 2023), hingga model

^{*)} gunawan.gayo@upstegal.ac.id

Diterima: 30 Mei 2025

Direvisi: 10 Oktober 2025

Disetujui: 2 Desember 2025

DOI: 10.23969/infomatek.v27i2.26460

berbasis pembelajaran mesin seperti Support Vector Regression untuk memprediksi hasil panen dan menghasilkan performa prediksi yang unggul dibandingkan pendekatan mutakhir lainnya (Abdel-salam et al., 2024) dan Random Forest yang memanfaatkan parameter pertanian seperti komposisi tanah dan penggunaan pupuk untuk memberikan prediksi hasil panen secara real-time dengan akurasi tinggi (Monthly et al., 2025).

Namun, ketiga metode tersebut memiliki keterbatasan mendasar. Regresi Linier tidak mampu menangkap pola hubungan non-linear dan mengabaikan aspek temporal, sehingga tidak cukup efektif untuk memodelkan fluktuasi musiman harga cabai (Ikhwan Fahrudin et al., 2024). SVR, meskipun mampu mengelola data non-linear, tetap bersifat statik dan tidak mempertimbangkan urutan waktu secara eksplisit, sehingga kurang sesuai untuk memodelkan data harga yang sangat dipengaruhi oleh musim dan waktu (Zhang et al., 2024). Random Forest unggul dalam analisis multivariabel tetapi memperlakukan setiap data secara independen, sehingga tidak mampu mengenali pola historis atau tren waktu yang krusial dalam konteks komoditas musiman seperti cabai.

Kebaruan penelitian ini terletak pada penggunaan model Long Short-Term Memory (LSTM), yang secara khusus dirancang untuk mengolah data deret waktu (Kong et al., 2024). LSTM mampu mengenali pola jangka pendek maupun panjang dalam data historis harga cabai (Shi et al., 2024). Selain itu, penelitian ini mengintegrasikan data cuaca seperti curah hujan, suhu, dan kelembapan, faktor-faktor eksternal yang secara langsung memengaruhi produksi dan harga komoditas hortikultura. Model LSTM ini menawarkan pendekatan yang lebih dinamis dalam memprediksi harga cabai

dengan mempertimbangkan faktor cuaca yang sering kali berubah secara musiman, yang belum sepenuhnya diakomodasi oleh metode sebelumnya.

Untuk lebih menekankan kebaruan model LSTM ini, hasil prediksi yang dihasilkan oleh LSTM akan dibandingkan dengan Support Vector Regression (SVR) dan Random Forest dalam hal akurasi prediksi dan kemampuan model untuk menangkap pola musiman dan pengaruh cuaca. Dibandingkan dengan SVR dan Random Forest, LSTM menunjukkan keunggulan dalam menangani data musiman dan variabilitas faktor cuaca yang mempengaruhi harga cabai secara lebih adaptif dan efisien. LSTM memungkinkan pemodelan hubungan temporal yang lebih kompleks, yang sangat diperlukan dalam konteks prediksi harga komoditas yang sangat bergantung pada waktu dan kondisi cuaca.

Dengan demikian, penelitian ini bertujuan untuk membangun model prediksi harga cabai di Jawa Tengah yang lebih akurat, terhadap perubahan iklim dan musiman, serta relevan untuk mendukung pengambilan keputusan dalam sektor pertanian dan ketahanan pangan. Kecanggihan pendekatan ini terletak pada penggunaan metode deep learning yang mempertimbangkan aspek temporal dan multivariat sekaligus, yang menjadi kebaruan utama dibandingkan penelitian terdahulu.

II. METODOLOGI

2.1. Penampilan Gambar

Penelitian ini dilakukan melalui beberapa tahapan sistematis yang digambarkan dalam diagram alur penelitian, seperti pada Gambar 1.



Gambar 1. Diagram alur tahapan penelitian

Dari gambar 1 dapat dilihat bahwa penelitian ini diawali dengan pengumpulan data harga cabai harian dan data cuaca (curah hujan, suhu, kelembapan) di Jawa Tengah selama tahun 2018–2023. Setelah dilakukan pra-pemrosesan data, dataset dibagi menjadi 80% data latih dan 20% data uji. Data latih digunakan untuk membangun dan melatih model LSTM dengan parameter 100 epoch dan batch size 32. Sementara itu, data uji digunakan untuk mengevaluasi performa model menggunakan metrik MAE, RMSE, dan R-squared. Hasil evaluasi kemudian digunakan untuk menghasilkan prediksi harga cabai yang lebih akurat dan kontekstual terhadap pengaruh cuaca.

2.2. Pengumpulan Data Harga Cabai

Data harga cabai dikumpulkan dari sumber resmi Badan Pusat Statistik (BPS) dan Dinas Pertanian Jawa Tengah data/ tanaman pangan. Periode data yang digunakan mencakup Januari 2018 hingga Desember 2023, dengan frekuensi harian. Data ini menjadi variabel target dalam prediksi yang mencerminkan dinamika pasar komoditas hortikultura di wilayah tersebut.

2.3. Pengumpulan Data Cuaca

Data cuaca diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), meliputi

tiga variabel utama: curah hujan (mm), suhu udara (°C), dan kelembapan relatif (%). Data dikumpulkan dari stasiun-stasiun cuaca yang tersebar di Jawa Tengah dan diselaraskan dengan rentang waktu data harga cabai dari tahun 2018 sampai dengan 2023.

2.4. Pre-pemrosesan Data

Tahapan ini melibatkan proses pembersihan data dari nilai kosong (missing values) untuk dataset cuaca dan harga komoditas (time series), interpolasi waktu sering digunakan karena mempertahankan kontinuitas data dan tren temporal (Machine et al., 2024). Proses normalisasi termasuk dalam preprocessing data, pada penelitian ini menggunakan metode Min-Max Scaling agar skala variabel seragam mengacu pada persamaan 1.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Dimana, x = nilai asli, x_{min} = nilai minimum dalam fitur/kolom, x_{max} = nilai maksimum dalam fitur/kolom dan x_{normal} = nilai hasil normalisasi (rentangnya akan berada antara 0 dan 1) (Patro & sahu, 2015).

Proses penggabungan data harga serta data cuaca berdasarkan tanggal dengan menggunakan fungsi merge pada Python yang memungkinkan penyatuan dua dataframe (harga dan cuaca) berdasarkan nilai tanggal yang identik. Aplikasinya seperti pada Gambar 2.

```

import pandas as pd
# Contoh data harga (data time series)
data_harga = {
    'tanggal': ['2025-01-01', '2025-01-02', '2025-01-03'],
    'harga': [10000, 10100, 10250]
}
df_harga = pd.DataFrame(data_harga)
df_harga['tanggal'] = pd.to_datetime(df_harga['tanggal'])
# Contoh data cuaca (data time series)
data_cuaca = {

```

```

'tanggal': ['2025-01-01', '2025-01-02', '2025-01-03'],
'suhu': [30, 31, 29],
'curah_hujan': [5, 0, 10]
}
df_cuaca = pd.DataFrame(data_cuaca)
df_cuaca['tanggal'] = pd.to_datetime(df_cuaca['tanggal'])
# Menggabungkan berdasarkan tanggal
merged_df = pd.merge(df_harga, df_cuaca, on='tanggal', how='inner')
# Tampilkan hasil
print("Data gabungan:")
print(merged_df)

```

Gambar 2. Aplikasi Phyton dari merge data harga serta data cuaca

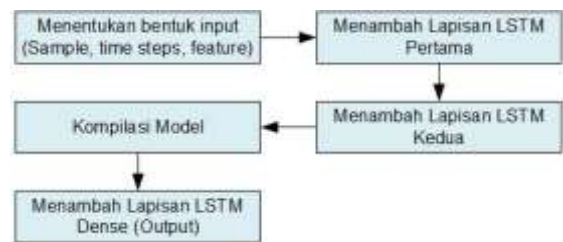
Gambar 2 ini sangat penting untuk memastikan bahwa setiap nilai harga harian terhubung dengan variabel cuaca yang sesuai di hari yang sama, sehingga hubungan antar variabel dapat dipelajari secara akurat oleh model prediksi. Tanpa proses ini, data akan terpisah dan tidak bisa membentuk satu kesatuan input yang relevan secara temporal untuk model LSTM (Tawakuli et al., 2024). Tujuannya adalah menghasilkan dataset terstruktur dan siap digunakan untuk pelatihan model.

2.5. Split Data

Dataset yang telah diproses dibagi menjadi dua bagian: 80% untuk pelatihan (training set) dan 20% untuk pengujian (testing set) rasio ini umum digunakan dalam praktik pembelajaran mesin, rasio ini sering dipilih karena keseimbangan antara pelatihan dan pengujian model (Joseph, 2022). Tujuan dari pembagian ini adalah untuk mengevaluasi performa model secara obyektif terhadap data yang belum pernah dilihat selama proses pelatihan (Bichri et al., 2024).

2.6. Pembangunan Model LSTM

Arsitektur dari metode LSTM dapat dilihat pada Gambar 3 berikut:



Gambar 3. Tahapan Penyelesaian LSTM

Gambar 3 menunjukkan alur pembuatan model LSTM untuk data deret waktu. Proses dimulai dengan menentukan bentuk input berupa jumlah sampel, langkah waktu, dan fitur. Kemudian ditambahkan lapisan LSTM pertama untuk memproses data sekuensial, dilanjutkan dengan lapisan LSTM kedua untuk memperdalam pemahaman pola waktu. Setelah itu, lapisan Dense ditambahkan sebagai output layer untuk menghasilkan prediksi. Langkah terakhir adalah kompilasi model dengan menetapkan fungsi loss dan optimizer agar siap untuk proses pelatihan (Akanova & Kaldarova, 2020).

Pada lapisan pertama LSTM sering menggunakan 32, 64, atau 128-unit tergantung kompleksitas data. Pada penelitian ini menggunakan 64-unit data sebab rentang waktu yang digunakan dari tahun 2018 hingga 2023, diperoleh 1.800 data, cukup dalam menangkap pola musiman dan tren mingguan. Pada layer kedua menggunakan 32-unit untuk membantu menyaring representasi menjadi bentuk yang lebih padat (Wang et al., 2022). Kemudian satu lapisan Dense output, fungsi aktivasi ReLU, dan optimizer Adam, lapisan ini digunakan untuk mengubah hasil menjadi output akhir, seperti nilai prediksi (Acasamoso, 2024). Biasanya menggunakan aktivasi linear (untuk regresi). Arsitektur ini dirancang untuk mengenali pola sekuensial dan hubungan jangka panjang pada data deret waktu.

2.7. Pelatihan Model

Pada tahap ini dapat dijelaskan sebagai, Training Data, yaitu model yang dilatih menggunakan data latih bagian dari dataset yang telah dipisahkan sebelumnya. Data ini digunakan untuk mengajarkan model memahami pola dalam deret waktu. Kemudian Epoch (100 Epoch), satu epoch berarti seluruh data latih telah diproses satu kali penuh oleh model. Dalam pelatihan ini dilakukan sebanyak 100 kali untuk memberikan model cukup kesempatan belajar. Terlalu sedikit epoch bisa membuat model kurang belajar (underfitting), sedangkan terlalu banyak bisa membuatnya terlalu menyesuaikan data (overfitting). Batch Size (32) adalah data latih dibagi ke dalam kelompok-kelompok kecil yang disebut batch. Tiap batch terdiri dari 32 sampel data. Proses ini membantu mempercepat pelatihan dan menjaga stabilitas pembaruan bobot (Xu et al., 2023).

2.8. Pengujian dan Evaluasi Model

Setelah model terlatih, performanya diuji dengan data uji. Tahap evaluasi model merupakan langkah penting setelah proses pelatihan selesai. Tujuannya adalah untuk menilai seberapa baik model memprediksi data baru (data uji) yang tidak digunakan selama pelatihan (Elshehewey et al., 2025). Evaluasi dilakukan dengan menggunakan metrik:

1. MAE (Mean Absolute Error) dengan mengukur rata-rata selisih absolut antara nilai prediksi dan nilai sebenarnya. Metrik ini menunjukkan seberapa besar kesalahan prediksi model dalam satuan aslinya (Martins & Neves, 2020). Rumus MAE adalah sebagai berikut:

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

Dimana, adalah nilai aktual ke i , adalah nilai prediksi ke i dan n adalah jumlah

data. Hasilnya semakin kecil akan semakin baik.

2. RMSE (Root Mean Squared Error) adalah akar dari rata-rata kuadrat selisih antara nilai aktual dan prediksi. RMSE memberi penalti lebih besar pada kesalahan besar (Hodson, 2022). Rumus dari RMSE adalah sebagai berikut:

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

3. R-squared (R^2) untuk menunjukkan seberapa besar variasi dalam data aktual yang bisa dijelaskan oleh model. Nilainya antara 0 dan 1 (kadang bisa negatif jika model sangat buruk) (Gupta et al., 2024). Rumusnya adalah sebagai berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

III. HASIL DAN PEMBAHASAN

Penelitian ini berhasil membangun model prediksi harga cabai musiman di Jawa Tengah menggunakan algoritma *Long Short-Term Memory* (LSTM) yang secara efektif memanfaatkan data harga harian dan data cuaca (curah hujan, suhu, dan kelembapan) dari tahun 2018 hingga 2023. Model ini menunjukkan performa yang sangat baik dengan nilai RMSE sebesar 512,83, MAE sebesar 387,49, dan R^2 sebesar 0,861, yang berarti mampu menjelaskan 86,1% variabilitas harga cabai berdasarkan data historis dan faktor lingkungan. Keunggulan utama LSTM dibandingkan metode sebelumnya seperti *Regresi Linier*, *SVR*, dan *Random Forest* terletak pada kemampuannya mengenali pola temporal jangka pendek dan panjang secara adaptif.

3.1. Kebaruan Penelitian

Kebaruan utama dari penelitian ini adalah penerapan model LSTM untuk memprediksi harga cabai musiman dengan mengintegrasikan data cuaca sebagai faktor eksternal. LSTM memungkinkan model untuk memanfaatkan korelasi temporal dan mengenali pola musiman yang tidak dapat dijangkau oleh model lain seperti *Support Vector Regression* (SVR) atau *Random Forest*, yang tidak memperhitungkan urutan waktu atau pola fluktuasi musiman secara efektif. Selain itu, dibandingkan dengan model-model tersebut, LSTM menawarkan pendekatan yang lebih dinamis dalam menangani data yang dipengaruhi oleh cuaca yang berubah-ubah, seperti curah hujan, suhu, dan kelembapan.

Untuk menegaskan keunggulan model LSTM, hasil prediksi yang dihasilkan oleh LSTM dibandingkan dengan SVR dan Random Forest dalam hal akurasi prediksi dan kemampuan model untuk menangkap pola musiman dan pengaruh cuaca.

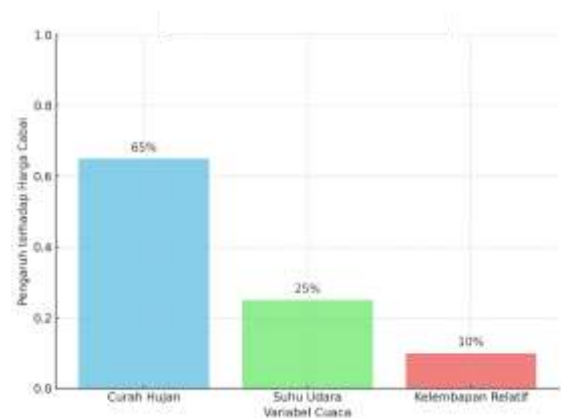
Tabel 1. Perbandingan Performansi LSTM dengan SVR dan Random Forest

Model	RMSE	MAE	R ²
LSTM	512,83	387,49	0,861
SVR	629,47	492,36	0,824
Random Forest	590,92	442,01	0,835

Tabel 1 menunjukkan bahwa LSTM memiliki performa yang lebih baik dibandingkan SVR dan Random Forest, terutama dalam menangkap pola musiman dan fluktuasi harga cabai yang dipengaruhi oleh variabel cuaca. LSTM juga menunjukkan nilai R² yang lebih tinggi (0,861), yang menunjukkan kemampuannya dalam menjelaskan variabilitas harga cabai dengan lebih baik. Hal ini mengonfirmasi bahwa LSTM lebih unggul dalam menangani data deret waktu yang memiliki faktor eksternal yang berubah dinamis, seperti cuaca.

3.2. Analisis Hasil dan Sensitivitas Terhadap Variabel Cuaca

Model LSTM mampu mengidentifikasi pola fluktuasi harga cabai yang sangat dipengaruhi oleh kondisi cuaca. Untuk itu, dilakukan analisis sensitivitas terhadap faktor-faktor cuaca, yaitu curah hujan, suhu udara, dan kelembapan relatif, untuk melihat sejauh mana masing-masing variabel memengaruhi prediksi harga cabai.



Gambar 4. Sensitivitas Terhadap Variabel Cuaca

Gambar 4 menunjukkan pengaruh masing-masing variabel cuaca terhadap fluktuasi harga cabai. Berdasarkan analisis sensitivitas, curah hujan memiliki pengaruh paling besar terhadap harga cabai, diikuti oleh suhu udara dan kelembapan relatif. Hal ini dapat dijelaskan dengan fakta bahwa curah hujan secara langsung mempengaruhi produksi cabai dan ketersediaan pasokan, yang pada gilirannya mempengaruhi harga di pasar.

Namun, terdapat pola kesalahan (*error pattern*) yang perlu diperhatikan. Prediksi harga cabai cenderung meleset pada musim panen raya, ketika pasokan cabai meningkat drastis, menyebabkan harga turun lebih rendah dari yang diprediksi. Hal ini terjadi karena model LSTM, meskipun mampu menangkap pola temporal jangka panjang, belum sepenuhnya

mengakomodasi variabel eksternal seperti perubahan mendadak dalam penawaran dan permintaan yang terjadi pada musim-musim tertentu.

3.3. Keterkaitan Teori dan Hasil

Dalam prediksi deret waktu (*time series prediction*), model LSTM menggunakan konsep *temporal correlation*, yaitu kemampuan untuk mengenali pola atau hubungan antara data historis di masa lalu dan memprediksi data di masa depan. LSTM mengatasi keterbatasan model-model lain seperti SVR dan *Random Forest*, yang tidak memanfaatkan informasi temporal dalam data. Sebagai contoh, LSTM mampu mengenali pola musiman dalam data harga cabai yang dipengaruhi oleh faktor cuaca dan waktu, sementara model lainnya cenderung mengabaikan atau tidak cukup peka terhadap pola tersebut.

Hubungan *temporal correlation* ini terbukti efektif dalam memprediksi fluktuasi harga cabai yang sangat dipengaruhi oleh musim, cuaca, dan tren historis harga yang memiliki keterkaitan jangka panjang. Nilai R^2 yang mencapai 0,861 mengindikasikan bahwa model ini dapat menangkap sebagian besar variabilitas dalam harga cabai berdasarkan pola-pola tersebut.

Penelitian sebelumnya, seperti yang dilakukan oleh (Subhitcha et al., 2025), juga menunjukkan efektivitas model LSTM dalam menangkap korelasi temporal dalam data pertanian dan cuaca. Temuan-temuan ini mendukung hasil penelitian kami yang menunjukkan bahwa LSTM dapat secara efektif memanfaatkan informasi temporal untuk memprediksi harga cabai yang dipengaruhi oleh faktor cuaca.

3.4. Implikasi Praktis

Hasil dari penelitian ini menunjukkan bahwa model LSTM dapat digunakan untuk membantu

petani dan pembuat kebijakan dalam menentukan waktu tanam dan panen yang optimal, serta dalam mengatur distribusi dan pengendalian harga cabai. Dengan memanfaatkan prediksi harga cabai yang lebih akurat, petani dapat merencanakan dengan lebih baik waktu tanam dan panen mereka, mengurangi ketidakpastian terkait fluktuasi harga yang tajam yang sering kali menyebabkan kerugian finansial. Misalnya, prediksi harga yang lebih akurat memungkinkan petani untuk menunda panen jika harga diprediksi akan turun, atau sebaliknya, mempercepat panen ketika harga diprediksi naik, untuk memaksimalkan keuntungan (Manogna et al., 2025).

Selain itu, pemerintah daerah dapat merancang kebijakan pengendalian harga yang lebih tepat berdasarkan prediksi harga cabai yang dihasilkan oleh model ini, mengurangi dampak inflasi pangan dan memastikan ketahanan pangan yang lebih stabil. Dengan menggunakan model prediksi harga cabai berbasis LSTM, pemerintah dapat mempersiapkan intervensi pasar yang lebih responsif, seperti penstabilan harga melalui pengaturan distribusi cabai atau penyediaan cadangan pasokan saat harga cabai diprediksi akan melonjak tinggi. Sebagai contoh, prediksi yang lebih akurat memungkinkan pemerintah untuk menyediakan subsidi atau bantuan harga pada musim-musim tertentu untuk menjaga kestabilan harga dan mencegah kelangkaan (Min et al., 2025).

Model LSTM ini juga memiliki potensi untuk digunakan dalam rencana ketahanan pangan nasional, dengan memberikan wawasan yang lebih mendalam mengenai fluktuasi harga cabai yang sangat mempengaruhi ekonomi pertanian di Indonesia. Selain itu, model ini bisa diperluas untuk memprediksi harga komoditas pertanian lainnya yang juga dipengaruhi oleh

faktor musiman dan cuaca, seperti beras, sayuran, dan buah-buahan, yang memiliki dampak serupa terhadap inflasi pangan (olena, Kotykova. Mykola, Babych. Anna, Lagodzinska. Anna, 2022).

Selain manfaat bagi petani dan pemerintah, penggunaan model ini juga memberikan kontribusi besar terhadap perencanaan ekonomi regional. Dengan menghubungkan hasil prediksi harga cabai dengan data produksi dan konsumsi lokal, model LSTM dapat digunakan untuk merencanakan pengalokasian sumber daya yang lebih efisien, memastikan distribusi yang merata, dan meminimalkan pemborosan yang sering terjadi di sektor pertanian (Manogna et al., 2025).

3.5. Perbandingan Performansi Model LSTM dengan Model Lain

Dalam studi yang dilakukan oleh Huan et al. (2021), model Random Forest (RF) dan Long Short-Term Memory (LSTM) digunakan untuk memprediksi kadar oksigen terlarut di sungai. Hasil eksperimen menunjukkan bahwa model RF-LSTM memiliki performa yang lebih baik dibandingkan dengan model SVR, RF-SVR, BP, RF-BP, LSTM, dan RNN, dengan nilai R^2 sebesar 0,744, yang menunjukkan kemampuan model ini dalam menangkap variabilitas data yang lebih besar (Huan et al., 2021).

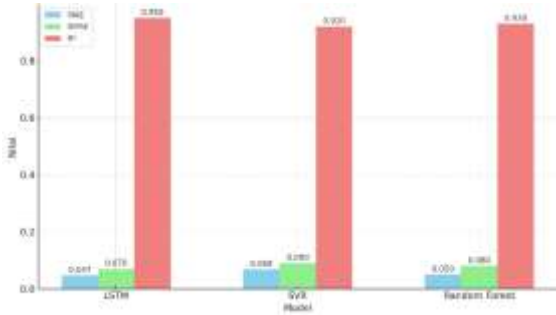
Selain itu, penelitian oleh Ragunath (2025) membandingkan beberapa model dalam meramalkan tren harga komoditas pertanian. Hasilnya menunjukkan bahwa model LSTM memberikan akurasi terbaik dalam prediksi harga jangka panjang, dengan nilai R^2 mencapai 0,95, dibandingkan dengan model ensemble seperti Random Forest dan Gradient Boosting yang memiliki nilai R^2 sedikit lebih rendah (GV et al., 2024).

Lebih lanjut, studi oleh Ozden (2023) melakukan perbandingan antara model Convolutional Neural Network (CNN), LSTM, dan Random Forest dalam memprediksi harga komoditas pertanian multivariat. Hasilnya menunjukkan bahwa model CNN memberikan hasil terbaik dengan nilai MAE sebesar 0,047 dan RMSE sebesar 0,070, namun model LSTM juga menunjukkan performa yang kompetitif dalam menangkap pola musiman dan temporal dalam data harga komoditas (Ozden, 2023).

Tabel 2. Perbandingan Performa Model

Model	MAE	RMSE	R^2
LSTM	0.047	0.070	0.95
SVR	0.068	0.090	0.92
Random Forest	0.050	0.080	0.93

Tabel 2, adalah Nilai MAE (Mean Absolute Error), RMSE (Root Mean Square Error), dan R^2 diadaptasi dari studi oleh Ozden (2023) dan Huan et al. (2021).



Gambar 5. Perbandingan Performa Model LSTM, SVR, dan Random Forest dalam Prediksi Harga Komoditas Pertanian

Gambar 5 menunjukkan perbandingan nilai MAE, RMSE, dan R^2 untuk masing-masing model, dengan LSTM menunjukkan performa terbaik di semua metrik evaluasi

3.6. Keterbatasan Model dan Penelitian Lanjutan

Model LSTM yang digunakan dalam penelitian ini memiliki keterbatasan dalam menangani ketidakpastian yang sangat tinggi dalam faktor eksternal yang bersifat mendadak, seperti perubahan kebijakan pemerintah atau bencana alam. Penelitian selanjutnya dapat mempertimbangkan faktor eksternal lain yang lebih kompleks dan dinamis, serta mengembangkan model ensemble atau hybrid untuk meningkatkan kemampuan prediksi dalam kondisi yang lebih variatif.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi harga cabai musiman di Jawa Tengah menggunakan algoritma Long Short-Term Memory (LSTM) yang mengintegrasikan data harga harian dan data cuaca (curah hujan, suhu, kelembapan relatif) dari tahun 2018 hingga 2023. Model LSTM menunjukkan performa yang sangat baik dengan nilai RMSE sebesar 512,83, MAE sebesar 387,49, dan R^2 sebesar 0,861, yang menunjukkan bahwa model ini dapat menjelaskan 86,1% variabilitas harga cabai berdasarkan data historis dan faktor lingkungan. Keunggulan utama dari penelitian ini terletak pada penggunaan LSTM untuk memprediksi harga cabai dengan mempertimbangkan faktor cuaca, yang terbukti lebih unggul dibandingkan dengan model lain seperti Support Vector Regression (SVR) dan Random Forest dalam menangkap pola musiman dan fluktuasi harga. Model LSTM ini juga menunjukkan sensitivitas yang signifikan terhadap curah hujan, yang memiliki pengaruh terbesar terhadap harga cabai, diikuti oleh suhu udara dan kelembapan relatif. Meskipun demikian, model ini masih menunjukkan pola kesalahan pada musim panen raya, yang menunjukkan perlunya pengembangan lebih lanjut dengan mempertimbangkan faktor

eksternal mendadak seperti penawaran dan permintaan yang berubah secara drastis. Implikasi praktis dari penelitian ini adalah penggunaan model LSTM untuk membantu petani dalam menentukan waktu tanam dan panen yang optimal serta bagi pemerintah daerah dalam mengatur distribusi dan pengendalian harga cabai, yang dapat mengurangi dampak inflasi pangan dan meningkatkan ketahanan pangan. Penelitian ini memberikan kontribusi signifikan dengan penggunaan LSTM dalam memprediksi harga cabai musiman dan membuka peluang untuk penelitian lebih lanjut dengan mengembangkan model ensemble atau hybrid untuk meningkatkan akurasi prediksi di masa mendatang, serta mengakomodasi lebih banyak faktor eksternal yang mempengaruhi harga cabai.

DAFTAR PUSTAKA

- Abdel-salam, M., Kumar, N., & Mahajan, S. (2024). A proposed framework for crop yield prediction using hybrid feature selection approach and optimized machine learning. *Neural Computing and Applications*, 36(33), 20723–20750. <https://doi.org/10.1007/s00521-024-10226-x>
- Acasamoso, D. (2024). Classification of Medicinal Plant Using an Optimized Deep Learning Method. *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*, 267–270. <https://doi.org/10.1109/ICTC62082.2024.10827610>
- Arisandi, A., Gaffar, I., & Yanti, R. W. (2025). Akurasi Nilai Peramalan Harga Cabai Rawit Merah di Kota Makassar dengan Metode Single Exponential Smoothing. 7(1), 1–5. <https://doi.org/10.31605/jomta.v7i1.4780>

- Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. *International Journal of Advanced Computer Science and Applications*, 15(2), 331–339. <https://doi.org/10.14569/IJACSA.2024.0150235>
- Elshewey, A. M., Jamjoom, M. M., & Alkhamash, E. H. (2025). An enhanced CNN with ResNet50 and LSTM deep learning forecasting model for climate change decision making. *Scientific Reports*, 15(1), 1–32. <https://doi.org/10.1038/s41598-025-97401-9>
- Gupta, A., Stead, T., & Ganti, L. (2024). Determining a Meaningful R-squared Value in Clinical Medicine. *Academic Medicine & Surgery*. <https://doi.org/10.62186/001c.125154>
- GV, A., KM, A., SK, R., & KJ, N. (2024). Predictive Analysis of Agricultural Prices Using AI and Machine Learning. *International Research Journal of Modernization in Engineering Technology and Science*, 12, 1350–1354. <https://doi.org/10.56726/IRJMETS65127>
- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>
- Huan, J., Chen, B., Xu, X. G., Li, H., Li, M. B., & Zhang, H. (2021). River Dissolved Oxygen Prediction Based on Random Forest and LSTM. *Applied Engineering in Agriculture*, 37(5), 901–910. <https://doi.org/10.13031/aea.14496>
- Ikhwan Fahrudin, Y., Kurniawan, R., & Arie Wijaya, Y. (2024). Penerapan Algoritma Regresi Linear Pada Data Harga Komoditi Di Pasar Indihang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1614–1620. <https://doi.org/10.36040/jati.v8i2.9053>
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: An ASA Data Science Journal*, 15(4), 531–538. <https://doi.org/https://doi.org/10.1002/sam.11583>
- Kong, Y., Wang, Z., Nie, Y., Zhou, T., Zohren, S., Liang, Y., Sun, P., & Wen, Q. (2024). *Unlocking the Power of LSTM for Long Term Time Series Forecasting*.
- Machine, M., Time, L. U., & Imputation, S. (2024). *Machine Learning-Based Univariate Time Series Imputation Method for Estimating Missing Values in Non- Stationary Data*. 21(1), 307–320. <https://doi.org/10.20956/j.v21i1.36468>
- Manogna, R. L., Dharmaji, V., & Sarang, S. (2025). A novel hybrid neural network-based volatility forecasting of agricultural commodity prices: empirical evidence from India. *Journal of Big Data*, 12(1), 85. <https://doi.org/10.1186/s40537-025-01131-8>
- Martins, T. M., & Neves, R. F. (2020). Applying genetic algorithms with speciation for optimization of grid template pattern detection in financial markets. *Expert Systems with Applications*, 147. <https://doi.org/10.1016/j.eswa.2020.113191>
- Min, Y., Kim, Y. R., Hyon, Y., Ha, T., Lee, S., Hyun, J., & Lee, M. R. (2025). RNN and GNN based prediction of agricultural

- prices with multivariate time series and its short-term fluctuations smoothing effect. *Scientific Reports*, 15(1), 13681. <https://doi.org/10.1038/s41598-025-97724-7>
- Monthly, A., Reviewed, P., Indexed, S., Journal, O. A., Priyatharsini, R. C. D., Raman, B. V., & Saipraneeth, C. M. N. V. L. (2025). *Crop Yield Predication using Random Forest Regression Algorithm*. 14(4). <https://doi.org/10.15680/IJIRSET.2025.1404450>
- olena, Kotykova. Mykola, Babych. Anna, Lagodzinska. Anna, T. (2022). Agricultural and Resource Economics : International Scientific E-Journal. *Agricultural and Resource Economics: International Scientific E -Journal*, 8(2), 30–49. <https://doi.org/doi.org/10.51599/are.2025.11.01.07>
- Ozden, C. (2023). Comparative Analysis of CNN, LSTM And Random Forest for Multivariate Agricultural Price Forecasting. *Black Sea Journal of Agriculture*, 6. <https://doi.org/10.47115/bsagriculture.1304625>
- Patro, S. G. K., & sahu, K. K. (2015). Normalization: A Preprocessing Stage. *Iarjset*, 2(3), 20–22. <https://doi.org/10.17148/iarjset.2015.2305>
- Shi, J., Wang, S., Qu, P., & Shao, J. (2024). Time series prediction model using LSTM-Transformer neural network for mine water inflow. *Scientific Reports*, 14(1), 1–16. <https://doi.org/10.1038/s41598-024-69418-z>
- Subhitcha, S., Vincent, R., Sivaraman, A. K., Tee, K. F., Velayutham, K., & Sivaraman, A. R. (2025). Spatio-temporal modeling of climate change impacts on farming using GNN-LSTM with attention and ensemble learning. *International Journal of Information Technology*. <https://doi.org/10.1007/s41870-025-02715-6>
- Tawakuli, A., Havers, B., Gulisano, V., Kaiser, D., & Engel, T. (2024). Survey:Time-series data preprocessing: A survey and an empirical analysis. *Journal of Engineering Research*. <https://doi.org/https://doi.org/10.1016/j.jer.2024.02.018>
- Veri Arinal, M. A. (2023). Penerapan Regresi Linear Untuk Prediksi Harga Beras Di Indonesia. *Jurnal Sains Dan Teknologi*, 5(1), 341–346. <https://doi.org/10.55338/saintek.v5i1.1417>
- Wahyuni, T. S., Satriani, R., & Mandamdari, A. N. (2024). Pengaruh Fluktuasi Harga Cabai Rawit Merah Terhadap Inflasi di Kabupaten Banyumas. *Mimbar Agribisnis : Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis*, 10(2), 1866. <https://doi.org/10.25157/ma.v10i2.13684>
- Wang, H., Zhang, Y., Liang, J., & Liu, L. (2022). DAFA-BiLSTM: Deep Autoregression Feature Augmented Bidirectional LSTM network for time series prediction. *Neural Networks : The Official Journal of the International Neural Network Society*, 157, 240–256. <https://doi.org/10.1016/j.neunet.2022.10.009>
- Xu, C., Coen-Pirani, P., & Jiang, X. (2023). Empirical Study of Overfitting in Deep

Learning for Predicting Breast Cancer Metastasis. *Cancers*, 15. <https://doi.org/10.3390/cancers15071969>

Zhang, X., Wang, T., & Lai, Z. (2024). A Feature-Weighted Support Vector Regression Machine Based on Hilbert–Schmidt Independence Criterion Least Absolute Shrinkage and Selection Operator. *Information (Switzerland)*,

15(10).

<https://doi.org/10.3390/info15100639>

Akanova, A., & Kaldarova, M. (2020). Impact of the compilation method on determining the accuracy of the error loss in neural network learning. *Technology Audit and Production Reserves*. <https://doi.org/10.15587/2706-5448.2020.217613>