

Pendekatan LSTM Berbasis Deep Learning dalam Memprediksi Fluktuasi Harga Cabai

Aryka Anisa Pertiwi*, Nisa Hanum Harani, Cahyo Prianto

Sekolah Vokasi, Teknik Informatika, Universitas Logistik dan Bisnis Internasional, Bandung, Indonesia

Email: ¹*aryaanisap22@gmail.com, ²nisa@ulbi.ac.id, ³cahyo@ulbi.ac.id

Email Penulis Korespondensi: aryaanisap22@gmail.com

Submitted: 25/07/2025; Accepted: 01/09/2025; Published: 04/09/2025

Abstrak—Fluktuasi harga cabai yang signifikan di Indonesia menimbulkan ketidakstabilan ekonomi, terutama bagi konsumen dan pelaku pasar. Penelitian ini bertujuan membangun model prediksi harga cabai harian menggunakan algoritma *Long Short-Term Memory* (LSTM) berbasis deep learning untuk menangkap pola musiman dan ketergantungan jangka panjang dalam data historis. Penelitian ini mengikuti pendekatan CRISP-DM, mencakup tahapan pemahaman bisnis, pemrosesan data, pelatihan model, hingga implementasi ke dalam *dashboard* berbasis web. *Dataset* yang digunakan berasal dari tahun 2022–2024, mencakup fitur seperti harga sebelumnya (Harga Sebelumnya), jenis cabai (Sub-Variant), transformasi waktu dalam bentuk sin-cos (Tanggal sin dan Tanggal cos), serta variabel kondisi pasar (Faktorisasi dan Kondisi) sebagai variabel bebas, dengan target prediksi Harga Hari Ini sebagai variabel terikat. Model regresi LSTM menunjukkan performa yang tinggi dengan nilai R^2 sebesar 0,9567, MAE sebesar 1.402,92, dan RMSE sebesar 2.595,98. Nilai R^2 yang mendekati 1 ini mengindikasikan bahwa terdapat hubungan yang sangat signifikan antara variabel bebas dan variabel terikat, serta model mampu menjelaskan hampir seluruh variansi harga aktual dalam data. Selain itu, model klasifikasi status harga (naik, turun, stabil) juga dikembangkan untuk memperkuat sistem pendukung keputusan. Implementasi sistem ke dalam *dashboard* interaktif memungkinkan pengguna untuk memprediksi harga secara *real-time*. Aplikasi ini dibangun menggunakan framework *Streamlit* dan bahasa pemrograman Python, dan dapat dijalankan secara lokal maupun daring, sehingga memberikan solusi praktis dalam pemantauan dan pengambilan keputusan harga komoditas pangan.

Kata Kunci: Prediksi Harga; Cabai; LSTM; Deep Learning; CRISP-DM

Abstract—The significant fluctuation in chili prices in Indonesia leads to economic instability, particularly for consumers and market stakeholders. This study aims to develop a daily chili price prediction model using the Long Short-Term Memory (LSTM) algorithm based on deep learning, designed to capture seasonal patterns and long-term dependencies in historical data. The research adopts the CRISP-DM approach, encompassing business understanding, data processing, model training, and implementation into a web-based dashboard. The dataset, collected from Pagar Alam City between 2022 and 2024, includes features such as previous prices, chili sub-variants, sinusoidal time transformations, and market conditions. The LSTM regression model demonstrated high performance, achieving an R^2 score of 0.9567, a MAE of 1,402.92, and an RMSE of 2,595.98. Additionally, a classification model was developed to predict price status (increase, decrease, stable) as a decision-support tool. The deployment of this system into an interactive dashboard enables real-time price predictions. These results indicate that the LSTM-based approach is not only technically accurate but also offers a practical solution for commodity price monitoring and decision-making in the food sector.

Keywords: Price Prediction; Chili; LSTM; Deep Learning; CRISP-DM

1. PENDAHULUAN

Harga cabai sebagai salah satu komoditas pangan strategis di Indonesia mengalami fluktuasi yang signifikan sepanjang tahun. Perubahan harga yang tajam, terutama pada momen tertentu seperti hari besar keagamaan, musim panen, atau gangguan pasokan, menimbulkan ketidakpastian bagi konsumen, distributor, dan pengambil kebijakan. Ketidakstabilan ini berdampak pada perencanaan logistik, ketahanan pangan, dan daya beli masyarakat, terutama kelompok menengah ke bawah [1][2]. Berbagai pendekatan telah dilakukan untuk memprediksi harga bahan pokok, termasuk penggunaan model statistik seperti regresi linier dan AutoRegressive Integrated Moving Average (ARIMA). Meskipun metode tersebut unggul dalam prediksi jangka pendek dan kemudahan interpretasi, beberapa penelitian menunjukkan bahwa model konvensional memiliki keterbatasan dalam mengenali pola fluktuasi yang bersifat musiman dan nonlinier. Model statistik konvensional seperti ARIMA memiliki keterbatasan dalam menangani kompleksitas harga komoditas pertanian yang dipengaruhi oleh pola musiman dan nonlinier. Penelitian tersebut membuktikan bahwa pendekatan berbasis deep learning, seperti LSTM dan GRU, secara konsisten menghasilkan akurasi prediksi yang lebih tinggi dibandingkan ARIMA, terutama dalam menghadapi volatilitas harga produk seperti bawang dan tomat yang sangat dipengaruhi oleh faktor eksternal dan temporal [3].

Sebagai solusi atas keterbatasan tersebut, algoritma Long Short-Term Memory (LSTM) yang termasuk dalam kategori Recurrent Neural Network (RNN) menjadi alternatif yang unggul. LSTM memiliki kemampuan untuk mempertahankan informasi jangka panjang melalui mekanisme tiga gerbang utama: input, forget, dan output. Hal ini menjadikannya efektif dalam menangkap pola harga yang kompleks dan bersifat sekuensial. Gunadi et al. (2025) dalam studinya terhadap prediksi harga cabai rawit di Kota Metro, menunjukkan bahwa model LSTM berhasil mempelajari pola musiman harga dengan error validasi yang sangat kecil, yaitu sebesar 0.00249 [4]. Penelitian serupa oleh Tami dan Owda (2023) juga mendukung efektivitas LSTM, di mana model yang diterapkan pada berbagai harga komoditas pangan berhasil mencapai nilai akurasi tinggi dengan R^2 sebesar 98,2% serta MAPE yang rendah, menunjukkan kinerja superior dibandingkan model regresi linier dan ARIMA [5].

Berbagai penelitian telah membuktikan efektivitas algoritma *Long Short-Term Memory* (LSTM) dalam konteks prediksi harga komoditas pertanian. Zhang et al. [6] menggunakan model LSTM untuk memprediksi harga

harian enam jenis sayuran di pasar grosir Beijing dan mencatat performa terbaik dengan nilai Mean Absolute Percentage Error (MAPE) sebesar 7,3%, lebih unggul dibandingkan model ARIMA dan XGBoost. Model Long Short-Term Memory (LSTM) telah terbukti efektif dalam menangani prediksi harga komoditas pertanian, terutama dalam menghadapi pola musiman dan ketergantungan jangka panjang. LSTM juga menunjukkan keunggulan dalam stabilitas hasil prediksi pada data yang bersifat nonlinier dan volatil, menjadikannya sangat cocok untuk domain harga pangan yang sering berubah. Efektivitas ini semakin diperkuat dalam studi-studi lanjutan yang mengintegrasikan LSTM ke dalam arsitektur hybrid, seperti CEEMDAN-LSTM dan TCN-LSTM, untuk meningkatkan akurasi prediksi melalui pemrosesan awal sinyal masukan[7]. Studi serupa juga menggabungkan CEEMDAN dengan Time Delay Neural Network (TDNN) dalam prediksi harga komoditas pertanian dan membuktikan bahwa pendekatan ini menghasilkan nilai RMSE dan MAPE yang lebih rendah dibandingkan ARIMA, SVR, dan GBM, memperkuat bukti bahwa kombinasi decomposisi sinyal dan deep learning sangat menjanjikan dalam domain ini [8]. Menyempurnakan temuan sebelumnya, Jayadianti et al. [9] memanfaatkan model LSTM untuk memprediksi harga sepuluh komoditas pangan strategis nasional, termasuk cabai, dan mencatat bahwa model tersebut mampu mencapai akurasi di atas 95% dengan error prediksi kurang dari 1%.

Meskipun LSTM telah terbukti menjadi metode yang menjanjikan dalam prediksi deret waktu, sebagian besar penelitian awal masih terbatas pada pendekatan univariat yang hanya memanfaatkan data historis seperti harga dan waktu. Galphade et al. (2024) menunjukkan bahwa penggunaan data multivariat yang mencakup berbagai faktor eksternal dapat secara signifikan meningkatkan akurasi model LSTM dalam peramalan energi terbarukan, membuktikan bahwa integrasi fitur tambahan dapat menangkap pola yang lebih kompleks secara efektif[10].

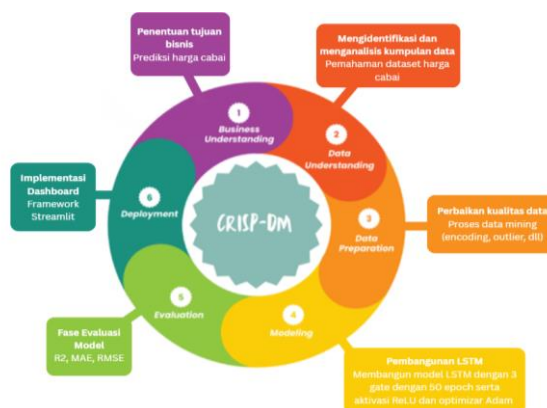
Pendekatan ini belum sepenuhnya mempertimbangkan kompleksitas pasar nyata yang melibatkan interaksi antar faktor seperti kondisi distribusi, stok pasokan, kebijakan harga pemerintah, maupun perilaku konsumen. Oleh karena itu, penelitian ini mengusulkan pengayaan fitur input dan pendekatan end-to-end berbasis LSTM yang lebih adaptif terhadap variasi pasar, serta mampu diimplementasikan langsung dalam sistem pendukung keputusan berbasis web. Pada penelitian ini, model dikembangkan menggunakan variabel yang lebih lengkap dan representatif, seperti Harga Sebelumnya, Sub-Variant, Tanggal sin, Tanggal cos, Faktorisasi, dan Kondisi. Seluruh variabel tersebut dirancang untuk merefleksikan kondisi aktual pasar secara lebih komprehensif. Tidak hanya berhenti pada tahap pemodelan, penelitian ini juga mengimplementasikan sistem prediksi ke dalam sebuah *dashboard* interaktif berbasis web menggunakan framework *Streamlit* dan bahasa pemrograman Python.

Aplikasi ini memungkinkan pengguna (baik dari kalangan pemerintah, pelaku usaha, maupun konsumen) untuk mengakses prediksi harga secara *real-time*, hanya dengan memasukkan parameter seperti jenis cabai, harga sebelumnya, dan kondisi pasar. Output yang diberikan mencakup harga perkiraan, status harga (naik, turun, atau stabil), serta grafik tren perubahan harga. Penelitian ini dirancang secara end-to-end menggunakan pendekatan CRISP-DM[11], dimulai dari pemahaman bisnis, eksplorasi data, preprocessing, pelatihan model, hingga deployment aplikasi. Tujuan utamanya adalah menghasilkan model prediktif harga cabai yang akurat, stabil, dan siap digunakan dalam pengambilan keputusan berbasis data. Dengan adanya sistem ini, diharapkan proses antisipasi terhadap fluktuasi harga cabai dapat dilakukan lebih cepat dan tepat, serta mendukung digitalisasi sistem informasi pangan nasional di era transformasi teknologi saat ini.

2. METODOLOGI PENELITIAN

2.1. Tahapan Penelitian

Penelitian ini menggunakan pendekatan Cross Industry Standard Process for Data Mining (CRISP-DM) sebagai kerangka utama dalam proses analisis data dan pembangunan model prediksi harga cabai. Pendekatan ini dipilih karena fleksibel dan mampu mengakomodasi tahapan analitik secara sistematis, mulai dari pemahaman bisnis hingga implementasi model ke dalam sistem yang aplikatif[11]. Tahapan penelitian ini terdiri atas enam langkah utama, seperti terlihat pada Gambar 1:



Gambar 1. Diagram Alur Metodologi Penelitian CRISP-DM

- a. *Business Understanding* – Harga cabai sebagai salah satu bahan pokok menunjukkan fluktuasi tinggi akibat faktor musiman, distribusi, dan kondisi pasar. Ketidakstabilan ini mempersulit pengambilan keputusan oleh pemerintah, pelaku distribusi, maupun konsumen. Tujuan penelitian ini adalah:
 1. Memprediksi nilai numerik Harga Hari Ini cabai.
 2. Mengklasifikasikan Status Harga (naik, turun, stabil).
- b. *Data Understanding* – Penjelajahan awal dilakukan terhadap dataset harga cabai harian dari tahun 2022–2024. Analisis deskriptif digunakan untuk memahami distribusi harga, keberadaan nilai hilang, dan pola musiman. Variabel seperti *Harga Sebelumnya*, *Sub-Variant*, *Tanggal*, dan *Status Harga* dianalisis secara visual dan statistik untuk mengidentifikasi pola awal dan keterkaitan antar fitur.
- c. *Data Preparation* – Proses praproses mencakup encoding fitur kategorikal (seperti kondisi dan status harga), pembersihan outlier dengan metode IQR, transformasi siklikal pada variabel tanggal, dan normalisasi menggunakan Min-Max Scaling. Dataset kemudian dibagi menjadi data latih (80%) dan uji (20%) tanpa pengacakan untuk menjaga kronologi time series. Fitur utama yang digunakan antara lain: Sub-Variant, Harga Sebelumnya, Tanggal_sin, Tanggal_cos, Faktorisasi, dan Kondisi.
- d. *Modeling* – Model LSTM dibangun dengan 50 unit memori dan satu lapisan output Dense. Pelatihan dilakukan selama 50 epoch dengan batch size 32, menggunakan fungsi aktivasi ReLU dan optimizer Adam. Fungsi loss yang digunakan adalah Mean Squared Error (MSE). Model diimplementasikan menggunakan TensorFlow dan Keras.
- e. *Evaluation* – Evaluasi dilakukan terhadap data uji menggunakan metrik MAE, RMSE, dan R^2 . Hasil evaluasi menunjukkan performa prediksi yang baik:
 1. MAE = 1.402,92
 2. RMSE = 2.595,98
 3. $R^2 = 0.9567$
 Selain itu, validasi silang 5-fold menghasilkan rata-rata R^2 sebesar 0.9496, menandakan stabilitas model dalam memprediksi harga cabai.
- f. *Deployment* – Model akhir diimplementasikan ke dalam dashboard berbasis *Streamlit*, yang memungkinkan pengguna memasukkan input (jenis cabai, harga sebelumnya, tanggal, dan kondisi pasar) untuk memperoleh prediksi harga harian serta status harga. Hasil prediksi divisualisasikan dalam bentuk grafik tren harga dan disajikan dalam antarmuka yang mudah digunakan.

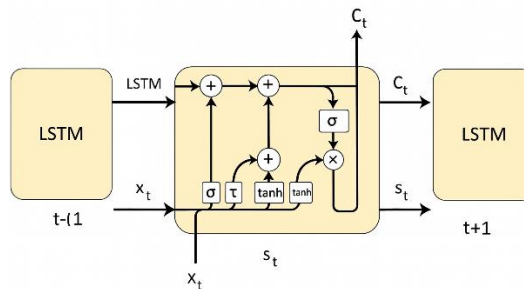
Setiap tahapan dilakukan secara iteratif [12], yaitu melalui proses pengulangan yang terencana dan bertahap untuk memperbaiki kualitas hasil pada setiap siklus. Misalnya, pada tahap data preparation, dilakukan penyesuaian encoding, deteksi dan pembersihan outlier, serta transformasi waktu secara berulang hingga distribusi data dinilai stabil dan representatif [13]. Demikian pula, pada tahap *modeling*, proses pelatihan model dilakukan dengan berbagai percobaan konfigurasi arsitektur, seperti jumlah neuron, fungsi aktivasi, dan batch size, serta tuning parameter lain seperti jumlah epoch. Penyesuaian ini penting karena masing-masing konfigurasi dapat memengaruhi performa prediksi secara signifikan, sebagaimana dijelaskan oleh Kim et al. dalam kajian komprehensif mengenai arsitektur LSTM untuk data deret waktu [14]. Evaluasi performa dilakukan setelah setiap percobaan dengan mengamati nilai MAE, RMSE, dan R^2 , kemudian hasil evaluasi tersebut menjadi dasar untuk iterasi selanjutnya. Proses iteratif ini memungkinkan peneliti untuk mengevaluasi dan memperbaiki kesalahan dari percobaan sebelumnya, sehingga model yang dihasilkan tidak hanya akurat, tetapi juga stabil dan andal saat digunakan pada data baru.

Dalam penelitian ini, digunakan model *Long Short-Term Memory* (LSTM) untuk memproses data deret waktu harga cabai harian. Metode deep learning seperti Long Short-Term Memory (LSTM) dirancang untuk secara otomatis mempelajari dan memodelkan ketergantungan temporal jangka panjang dalam deret waktu, terutama pada data yang kompleks dan tidak linier. Menurut Mojtahedi et al. (2025), LSTM memiliki keunggulan dalam menangkap dinamika temporal yang sulit ditangani oleh metode konvensional, dan telah diadopsi secara luas dalam berbagai bidang teknik dan prediksi berbasis waktu karena fleksibilitasnya dalam mengelola urutan data yang panjang [15]. Pemodelan LSTM dilakukan setelah data disesuaikan ke dalam format input tiga dimensi [samples, timesteps, features], melalui proses *reshape* data yang telah dinormalisasi [16]. Penyesuaian ini penting agar LSTM dapat belajar pola musiman, hubungan temporal antar waktu, serta ketergantungan jangka panjang dari data historis harga cabai yang kompleks dan fluktuatif.

2.2. Penerapan dan Pengujian Model

Pada Gambar 2 ditampilkan ilustrasi arsitektur internal dari satu unit Long Short-Term Memory (LSTM) yang diimplementasikan dalam model prediksi harga cabai. Arsitektur ini dirancang secara khusus untuk mengatasi keterbatasan model jaringan saraf konvensional dalam menangani dependensi jangka panjang pada data deret waktu. Struktur LSTM tersebut terdiri dari tiga komponen utama yang disebut sebagai gerbang: *forget gate*, *input gate*, dan *output gate*. Ketiga gerbang ini berperan penting dalam mengatur dan mengendalikan aliran informasi yang masuk dan keluar dari sel memori. *Forget gate* berfungsi untuk menentukan informasi mana dari status sebelumnya yang harus dilupakan, sedangkan *input gate* mengatur informasi baru yang akan disimpan dalam sel memori. Sementara itu, *output gate* bertugas mengontrol informasi yang akan diteruskan ke langkah waktu berikutnya. Interaksi dinamis antar gerbang ini berlangsung secara simultan dalam setiap langkah waktu, memungkinkan LSTM untuk mempertahankan informasi historis yang relevan dan secara selektif menyaring data yang tidak signifikan.

Kemampuan ini menjadikan LSTM sangat efektif dalam mengidentifikasi dan mempelajari pola musiman atau fluktuatif yang kompleks, seperti yang terdapat pada pergerakan harga komoditas cabai dari waktu ke waktu.



Gambar 2. Arsitektur LSTM

Arsitektur Long Short-Term Memory (LSTM) terdiri dari serangkaian gerbang yang mengatur aliran informasi dalam unit memori. Setiap unit memproses data berdasarkan urutan waktu dan mempertahankan informasi penting sambil melupakan yang tidak relevan. Proses ini diawali dari gerbang pelupa (forget gate) yang menentukan bagian dari memori sebelumnya yang perlu disingkirkan. Rumus forget gate dituliskan sebagai berikut:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Dalam persamaan tersebut, f_t merupakan vektor keluaran sigmoid yang nilainya berkisar antara 0 dan 1, berfungsi mengatur seberapa banyak informasi dari cell state sebelumnya C_{t-1} yang akan dipertahankan. Informasi ini ditentukan berdasarkan kombinasi dari input saat ini x_t dan hidden state sebelumnya h_{t-1} , yang dikalikan dengan bobot W_f dan ditambahkan bias b_f . Setelah menentukan informasi yang akan dilupakan, langkah berikutnya adalah memutuskan informasi baru apa yang akan disimpan dalam cell state melalui gerbang input (input gate), yang dirumuskan sebagai:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Input gate menghasilkan vektor i_t yang akan mengatur kontribusi dari informasi baru ke dalam memori jangka panjang. Namun, informasi baru ini terlebih dahulu diolah dalam bentuk kandidat memori menggunakan fungsi aktivasi tanh pada gabungan input dan hidden state sebelumnya, sebagaimana dirumuskan berikut:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

Kandidat memori \tilde{C}_t inilah yang kemudian dikombinasikan dengan hasil dari gerbang input untuk memperbarui cell state yang menjadi inti penyimpanan informasi dalam LSTM. Proses pembaruan cell state dijelaskan dalam rumus:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

Persamaan ini menunjukkan bahwa cell state saat ini C_t merupakan hasil penjumlahan antara bagian memori lama yang dipertahankan (melalui $f_t * C_{t-1}$) dan informasi baru yang ditambahkan (melalui $i_t * \tilde{C}_t$). Selanjutnya, untuk menghasilkan keluaran pada timestep saat ini, digunakan gerbang output (output gate) yang menghitung vektor kontrol output melalui fungsi sigmoid sebagai berikut:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

Nilai o_t digunakan untuk menentukan bagian dari cell state yang akan ditampilkan sebagai output. Akhirnya, hidden state h_t , yang juga menjadi output LSTM pada waktu t , dihitung dengan mengalikan o_t dengan fungsi aktivasi tanh dari cell state terkini, sebagaimana ditunjukkan oleh persamaan:

$$h_t = o_t * \tanh(C_t) \quad (6)$$

Dengan demikian, melalui keenam langkah ini, unit LSTM mampu menjaga konteks informasi dalam urutan data yang panjang, sekaligus mengatasi permasalahan vanishing gradient yang sering terjadi pada arsitektur Recurrent Neural Network (RNN) standar. Dalam penelitian ini, LSTM digunakan untuk memodelkan data harga komoditas cabai yang bersifat fluktuatif dan bergantung pada pola musiman. Setiap unit memori memproses input harga hari sebelumnya, kategori cabai, dan fitur waktu, lalu memutuskan informasi mana yang harus disimpan atau dilupakan. Hasil dari proses internal ini menghasilkan representasi tersembunyi h_t dan cell state C_t , yang digunakan untuk prediksi harga pada hari berikutnya. Pendekatan ini lebih adaptif dibandingkan model statistik konvensional karena mampu menangkap ketergantungan jangka panjang serta dinamika pasar yang berubah-ubah. Implementasi LSTM juga didukung oleh proses pra-proses data seperti normalisasi, encoding, dan transformasi tanggal, sehingga model dapat belajar dengan lebih efisien. Setelah seluruh data melalui tahap pra-pemrosesan dan disiapkan dengan format yang sesuai untuk model deret waktu, pembangunan model LSTM dilakukan dengan menggunakan kerangka



Sequential dari Keras. Model ini terdiri dari satu lapisan utama LSTM yang bertugas untuk menangkap pola-pola urutan dalam data historis, kemudian dilanjutkan oleh satu lapisan Dense yang berfungsi sebagai lapisan output untuk menghasilkan prediksi harga pada waktu tertentu. Struktur dan konfigurasi lengkap dari arsitektur model yang digunakan dapat dilihat secara lebih rinci pada Tabel 1 berikut.

Tabel 1. Arsitektur Model LSTM.

Layer Type	Units	Activation	Output Shape
LSTM	50	ReLU	(None, 50)
Dense	1	Linear	(None, 1)

Model dilatih selama 50 epoch dengan batch size sebesar 32. Optimasi dilakukan dengan Adam optimizer, sedangkan fungsi loss yang digunakan adalah Mean Squared Error. Untuk mengevaluasi performa model, digunakan empat metrik evaluasi regresi, yakni:

1. MAE (Mean Absolute Error)
2. MSE (Mean Squared Error)
3. RMSE (Root Mean Squared Error)
4. R² (Coefficient of Determination)

Root Mean Squared Error (RMSE) adalah salah satu metrik evaluasi yang digunakan untuk mengukur seberapa besar rata-rata kesalahan antara nilai aktual dan nilai prediksi dari suatu model. RMSE dihitung dengan mengambil akar kuadrat dari rata-rata kuadrat selisih antara nilai observasi sebenarnya (y_i) dan nilai yang diprediksi oleh model (\hat{y}_i). Secara matematis, RMSE dirumuskan sebagai[17]:

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

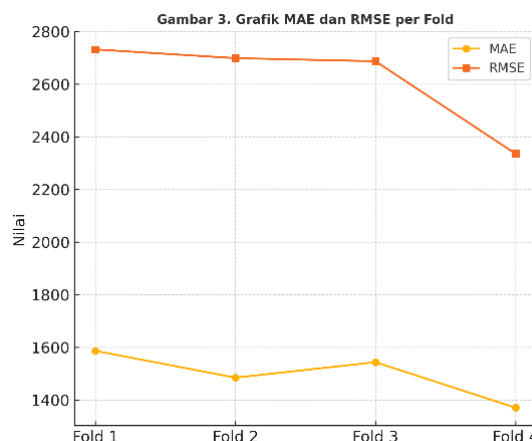
Di mana n merupakan jumlah total data. Nilai RMSE yang lebih kecil menunjukkan bahwa hasil prediksi model lebih mendekati nilai aktual, sehingga menandakan tingkat kesalahan prediksi yang rendah. Karena menggunakan kuadrat dari selisih nilai, RMSE akan memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar (outlier), sehingga metrik ini sangat sensitif terhadap deviasi ekstrem. Oleh karena itu, RMSE sering digunakan dalam evaluasi model prediksi yang memerlukan presisi tinggi, seperti dalam pemodelan harga, suhu, atau permintaan.

Selanjutnya, model divalidasi menggunakan 4-Fold Cross Validation untuk memastikan kestabilan dan generalisasi. Hasil evaluasi disajikan pada Tabel 2 sebagai berikut:

Tabel 2. Hasil Evaluasi Model LSTM per Fold.

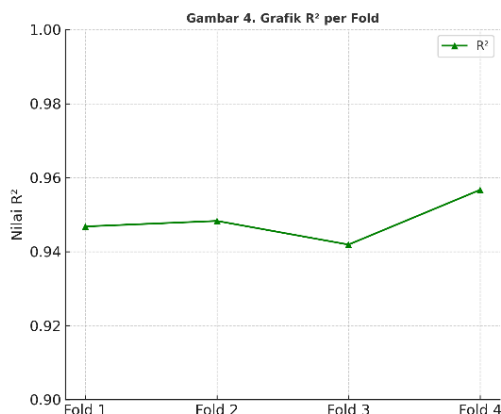
Fold	MAE	RMSE	R ²
Fold 1	1587.01	2732.17	0.9468
Fold 2	1485.08	2699.00	0.9483
Fold 3	1543.96	2687.28	0.9419
Fold 4	1371.01	2336.04	0.9567
Fold 5	1366.81	2525.70	0.9546

Hasil validasi silang menunjukkan bahwa model LSTM memiliki performa yang konsisten dan akurat, dengan rata-rata R² mencapai 0.9496. Visualisasi performa model pada tiap fold ditampilkan pada grafik pada gambar 3 dan gambar 4 berikut:



Gambar 3. Grafik MAE dan RMSE per Fold

Pada Gambar 3 menampilkan nilai MAE dan RMSE dari model LSTM pada masing-masing fold. Terlihat bahwa Fold 4 memiliki nilai MAE dan RMSE paling rendah, menunjukkan performa prediksi terbaik dibanding fold lainnya. Secara umum, grafik memperlihatkan bahwa model memiliki performa yang cukup stabil antar fold, dengan penurunan signifikan pada Fold 4.

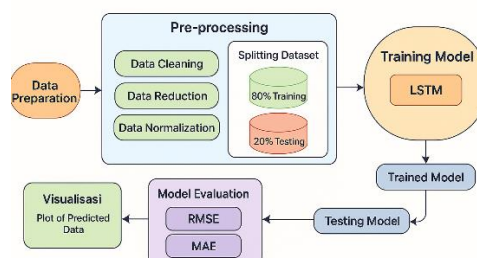


Gambar 4. Grafik R2 per Fold

Gambar 4 menampilkan nilai koefisien determinasi (R^2) pada setiap fold. Nilai R^2 berada di atas 0.94 untuk seluruh fold, dengan nilai tertinggi pada Fold 4 yaitu sekitar 0.9567. Hal ini menunjukkan bahwa model LSTM mampu menjelaskan sebagian besar variasi data aktual, dan performanya cukup stabil antar fold.

2.3. Implementasi Model Prediksi LSTM

Proses implementasi model prediksi harga cabai dalam penelitian ini dilakukan dengan menggunakan algoritma Long Short-Term Memory (LSTM) sebagai pendekatan utama. Gambar 5 menunjukkan alur kerja model prediktif yang dikembangkan berdasarkan tahapan aktual pada sistem.



Gambar 5. Alur Penelitian

Berdasarkan alur tersebut, proses implementasi dilakukan sebagai berikut:

- Data historis yang telah dibersihkan dari missing value, outlier, serta dilakukan encoding dan transformasi sin-cos untuk fitur tanggal, kemudian dinormalisasi menggunakan Min-Max Scaling.
- Data yang telah dibersihkan dibagi menjadi dua bagian: 80% untuk pelatihan (training) dan 20% untuk pengujian (testing). Pembagian dilakukan secara kronologis, tanpa randomisasi, agar urutan waktu tetap terjaga dan representatif terhadap kondisi nyata prediksi deret waktu.
- Pembuatan Model LSTM, dimana model dibangun menggunakan pustaka Keras dengan struktur:
 - LSTM Layer berisi 50 unit memori dengan fungsi aktivasi ReLU
 - Dense Layer sebagai output regresi dengan satu neuron
 - Optimizer: Adam
 - Loss function: Mean Squared Error (MSE)
 - Epoch: 50
 - Batch size: 32

Input model diformat ke dalam struktur tiga dimensi [samples, timesteps, features] sebagaimana disyaratkan oleh arsitektur LSTM.

- Pada setiap langkah waktu (time step), data masukan diproses melalui struktur internal sel Long Short-Term Memory (LSTM) yang terdiri atas tiga komponen utama, yaitu forget gate, input gate, dan output gate. Forget gate berfungsi untuk menentukan informasi mana dari memori sebelumnya yang harus dilupakan agar tidak membebani proses prediksi berikutnya. Selanjutnya, input gate bertugas mengatur dan memilih informasi baru yang relevan

untuk disimpan ke dalam memori jangka panjang. Setelah memori diperbarui, output gate akan menghasilkan keluaran berdasarkan informasi yang telah diproses secara selektif. Melalui mekanisme tiga gerbang ini, LSTM mampu mempertahankan, memperbarui, serta membuang informasi secara adaptif, sehingga sangat efektif dalam mengenali dan menangkap pola fluktuatif harga cabai harian, baik dalam jangka pendek maupun jangka panjang.

e. Setelah pelatihan selesai, model diuji terhadap data uji menggunakan metrik evaluasi regresi:

1. MAE (Mean Absolute Error): 1.402,92
2. RMSE (Root Mean Squared Error): 2.595,98
3. R^2 (Coefficient of Determination): 0,9567

Selain itu, 5-Fold Cross Validation juga diterapkan untuk memastikan kestabilan model. Rata-rata R^2 yang diperoleh dari validasi silang adalah 0,9496, menunjukkan bahwa model tidak hanya akurat, tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru.

2.4. Implementasi Dashboard Prediktif

Dalam sistem prediksi harga berbasis machine learning, penyampaian hasil kepada pengguna akhir merupakan komponen penting yang memungkinkan model digunakan secara nyata dalam proses pengambilan keputusan. Salah satu pendekatan yang efektif adalah penggunaan *dashboard* interaktif, yaitu antarmuka visual berbasis web yang menyajikan hasil prediksi secara *real-time* dan memungkinkan interaksi langsung dari pengguna. *Dashboard* berperan sebagai penghubung antara model prediktif yang kompleks dengan kebutuhan pengguna non-teknis seperti pelaku pasar, pembuat kebijakan, maupun masyarakat umum. Penelitian terbaru menegaskan bahwa visualisasi prediktif berbasis web dapat meningkatkan pemahaman pengguna terhadap fluktuasi harga dan mempercepat proses pengambilan keputusan berbasis data. Sebuah studi melalui pengembangan aplikasi TradeForecast memperlihatkan bahwa integrasi prediksi AI dengan visualisasi interaktif dalam antarmuka web dapat meningkatkan keterlibatan pengguna non-teknis dalam memahami pergerakan pasar secara *real-time*[18]. Iskandar dan Arifin juga menunjukkan pentingnya antarmuka yang adaptif untuk menyajikan prediksi harga cabai kepada pengguna lokal melalui sistem digital [19].

Dalam penelitian ini, *dashboard* dikembangkan menggunakan framework *Streamlit*, yang dipilih karena sifatnya ringan, fleksibel, dan mudah diintegrasikan dengan model prediktif berbasis Python. Framework ini memungkinkan penggabungan antara form input pengguna dan visualisasi interaktif hasil prediksi. Form input dalam *dashboard* mencakup variabel seperti Harga Sebelumnya, Sub-Variant, Tanggal (sin/cos), Faktorisasi, dan Kondisi Pasar. Output yang diberikan meliputi prediksi harga selama tujuh hari ke depan, status harga (naik/turun/stabil), serta grafik tren harga harian. Penerapan *dashboard* ini merupakan bagian dari tahap deployment dalam pendekatan CRISP-DM. Menurut Shimaoka et al., deployment adalah fase penting dalam pengembangan sistem analitik karena menjembatani hasil teknis dengan penggunaan nyata di lapangan [11]. Hal ini diperkuat oleh Elkabalawy et al., yang menunjukkan keberhasilan penerapan model prediktif berbasis CRISP-DM dalam skenario *real-world* dengan antarmuka yang intuitif [20]. Pendekatan serupa juga digunakan oleh Agustin et al. dalam sistem prediksi harga cabai di Garut, yang menyajikan hasil regresi dalam format digital untuk memudahkan pemanfaatan hasil prediksi oleh pengguna non-ahli[21].

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi dan evaluasi dari model prediksi harga cabai harian menggunakan algoritma *Long Short-Term Memory* (LSTM). Proses dimulai dari pengumpulan dan pembersihan data, diikuti dengan tahap preprocessing yang mencakup encoding data kategorikal, transformasi waktu secara siklikal, normalisasi, dan pembentukan format input tiga dimensi. Seluruh tahapan mengikuti pendekatan CRISP-DM, yang memungkinkan proses analitik dilakukan secara terstruktur dari pemahaman bisnis hingga implementasi akhir ke dalam sistem aplikasi.

3.1 Hasil Eksperimen Model Regresi

Dataset yang digunakan terdiri dari lebih dari seribu catatan harga harian komoditas cabai. Sebelum digunakan dalam pelatihan model, data melalui tahapan pembersihan dan transformasi untuk memastikan kualitas dan kesesuaiannya dengan kebutuhan model deep learning. Proses pembersihan data mencakup penghapusan nilai duplikat, penanganan nilai hilang (missing values) pada kolom-kolom penting seperti Harga Hari Ini, Sub-Variant, dan Status Harga, serta deteksi dan penghapusan outlier pada data numerik menggunakan metode Interquartile Range (IQR). Tujuan pembersihan ini adalah untuk menghilangkan anomali yang dapat mengganggu proses pelatihan model.

Sementara itu, tahapan transformasi data meliputi beberapa langkah penting, yaitu:

- a. Encoding variabel kategorikal seperti Sub-Variant, Faktorisasi, dan Kondisi ke dalam bentuk numerik menggunakan metode *Label encoding*, agar bisa dibaca oleh model.
- b. Transformasi waktu secara siklikal, di mana fitur tanggal diubah ke bentuk sinusoidal (menghasilkan dua fitur baru: Tanggal_sin dan Tanggal_cos) untuk merepresentasikan siklus bulanan yang berulang.
- c. Normalisasi fitur numerik ke dalam rentang 0–1 menggunakan metode *Min-Max Scaling* agar bobot fitur seimbang dan model lebih cepat konvergen.

d. Reshape data ke format tiga dimensi [samples, timesteps, features] agar sesuai dengan struktur input yang dibutuhkan oleh arsitektur LSTM.

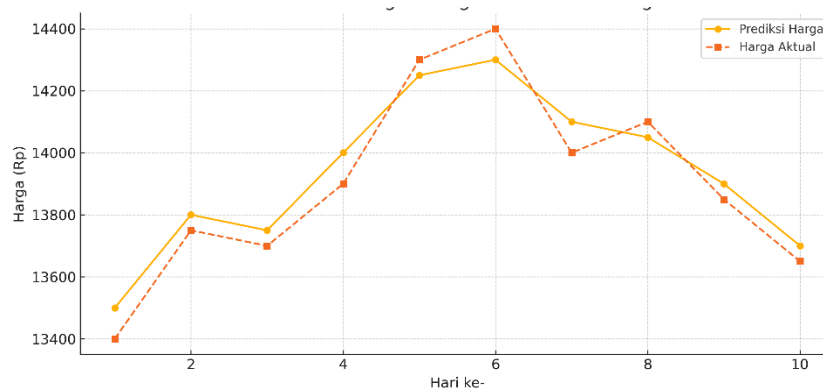
Setelah seluruh proses tersebut, data dibagi menjadi 80% data latih dan 20% data uji, dengan pembagian dilakukan secara kronologis tanpa pengacakan, sesuai dengan karakteristik data deret waktu. Model LSTM yang dibangun menunjukkan performa sangat baik dalam memprediksi harga harian cabai. Evaluasi terhadap data uji menghasilkan nilai Mean Absolute Error (MAE) sebesar 1.402,92, Root Mean Squared Error (RMSE) sebesar 2.595,98, dan koefisien determinasi (R-squared) sebesar 0,9567. Nilai R^2 yang mendekati 1 menunjukkan bahwa model mampu menjelaskan lebih dari 95% variasi harga cabai berdasarkan data historis yang diberikan.

Tabel 3 menyajikan perbandingan antara data harga aktual dengan hasil prediksi model LSTM pada lima data uji. Selisih antara nilai aktual dan prediksi cukup kecil, menunjukkan bahwa model mampu memetakan pola data dengan baik. Secara umum, prediksi yang dihasilkan model LSTM menunjukkan kecenderungan mengikuti arah perubahan data aktual. Ketika harga mengalami kenaikan atau penurunan, model mampu menyesuaikan proyeksi dengan tren tersebut. Hal ini mengindikasikan bahwa model tidak hanya menghafal pola, tetapi juga berhasil memahami dinamika pergerakan harga berdasarkan urutan waktu. Keakuratan yang ditampilkan dalam tabel merupakan hasil dari proses pelatihan model dengan fitur-fitur yang telah diolah, seperti harga sebelumnya, representasi musiman dari tanggal, serta beberapa fitur tambahan yang relevan. Model LSTM terbukti efektif dalam mengenali pola-pola jangka pendek maupun kecenderungan musiman yang memengaruhi fluktuasi harga cabai. Dengan hasil prediksi yang relatif stabil dan selisih kesalahan yang kecil, model ini berpotensi diimplementasikan untuk sistem prediksi harga secara berkala, guna mendukung pengambilan keputusan dalam pengelolaan distribusi dan stabilisasi harga pangan.

Tabel 3. Perbandingan Data Aktual dan Hasil Prediksi Model LSTM

No.	Data Test (Aktual)	Prediction Result LSTM
1.	29,000	29,182
2.	44,000	43,319
3.	60,000	61,051
4.	48,500	48,676
5.	35,000	34,447

Visualisasi perbandingan antara harga prediksi dan harga aktual dari data uji ditampilkan pada Gambar 6. Dari grafik tersebut, terlihat bahwa model berhasil mengikuti pola fluktuasi harga dengan deviasi yang relatif kecil. Meskipun terdapat beberapa titik yang menyimpang, secara umum prediksi cukup akurat dalam mengikuti tren aktual.



Gambar 6. Perbandingan Harga Prediksi dan Harga Aktual

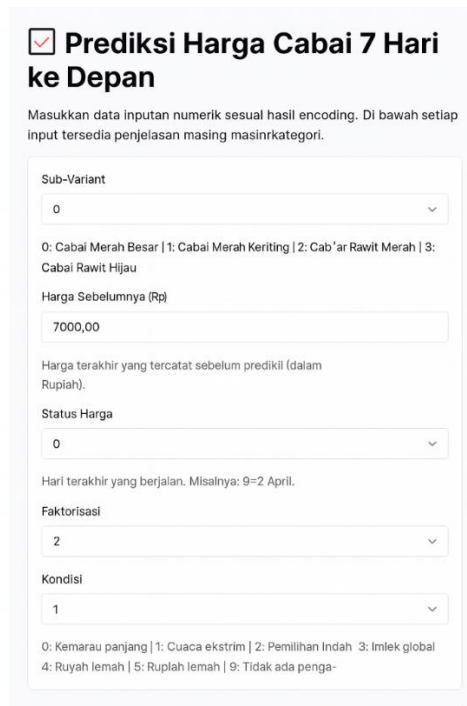
3.1.1 Validasi dan Evaluasi Stabilitas Model

Untuk memastikan model tidak hanya bekerja baik pada data uji tunggal, dilakukan pengujian validasi silang menggunakan skema 5-Fold Cross Validation. Proses ini membagi data menjadi lima bagian dan secara bergantian menggunakan satu bagian sebagai data validasi, sementara sisanya digunakan untuk pelatihan. Hasil evaluasi pada masing-masing fold ditampilkan dalam Tabel 2 di bagian 2.2, sementara grafik performa MAE, RMSE, dan R^2 disajikan pada Gambar 2 hingga Gambar 4 (di bagian metodologi sebelumnya). Rata-rata R^2 pada seluruh fold mencapai 0,9496, yang mengindikasikan kestabilan performa model dalam menjelaskan variansi data. Nilai MAE dan RMSE juga relatif konsisten, yang menunjukkan bahwa model tidak mengalami overfitting maupun underfitting yang signifikan pada subset data yang berbeda.

3.2 Implementasi Aplikasi Prediksi

Sebagai bentuk pemanfaatan hasil penelitian, model LSTM yang telah dilatih diintegrasikan ke dalam aplikasi *dashboard* berbasis *Streamlit*. Aplikasi ini memungkinkan pengguna untuk melakukan prediksi harga cabai dengan memasukkan beberapa parameter seperti jenis cabai (sub-varian), harga sebelumnya, tanggal (yang akan

ditransformasi secara internal), dan kondisi pasar. *Dashboard* ini dirancang untuk menampilkan hasil prediksi harga hari ini dalam bentuk angka, perubahan harga terhadap hari sebelumnya dalam persentase, serta klasifikasi status harga (naik, turun, atau stabil). Aplikasi dapat digunakan secara *real-time*, sehingga memberikan manfaat langsung bagi pemangku kebijakan, pelaku pasar, atau konsumen dalam memantau dan mengantisipasi dinamika harga pangan. Tampilan visual dari *dashboard* (Gambar 5) menampilkan antarmuka yang sederhana namun informatif. Penggunaan framework *Streamlit* memungkinkan integrasi yang efisien antara model machine learning dan tampilan antarmuka pengguna berbasis web. Model LSTM yang telah dilatih disimpan dalam format .pkl dan selanjutnya diintegrasikan ke dalam sebuah aplikasi *dashboard* berbasis web yang dikembangkan menggunakan framework *Streamlit*. Antarmuka *dashboard* ini ditampilkan pada Gambar 7.



Prediksi Harga Cabai 7 Hari ke Depan

Masukkan data inputan numerik sesuai hasil encoding. Di bawah setiap input tersedia penjelasan masing masing kategori.

Sub-Variant
0

0: Cabai Merah Besar | 1: Cabai Merah Keriting | 2: Cab'ar Rawit Merah | 3: Cabai Rawit Hijau

Harga Sebelumnya (Rp)
7000,00

Harga terakhir yang tercatat sebelum prediksi (dalam Rupiah).

Status Harga
0

Hari terakhir yang berjalan. Misalnya: 9=2 April.

Faktorisasi
2

Kondisi
1

0: Kemarau panjang | 1: Cuaca ekstrim | 2: Pemilihan Indah | 3: Imlek global
4: Ruyah lemah | 5: Ruyah lemah | 9: Tidak ada penga-

Gambar 7. *Dashboard*

Pada Gambar 7 terlihat bahwa model prediksi menerima enam variabel input dalam bentuk numerik, yang merupakan hasil dari proses encoding untuk memenuhi kebutuhan arsitektur LSTM yang hanya mampu mengolah data numerik. Keenam variabel tersebut terdiri atas: pertama, Sub-Variant, yaitu jenis cabai yang dikodekan menjadi nilai numerik 0 untuk Cabai Merah Besar, 1 untuk Cabai Merah Keriting, 2 untuk Cabai Rawit Merah, dan 3 untuk Cabai Rawit Hijau. Kedua, Harga Sebelumnya, yaitu nilai harga terakhir sebelum prediksi yang digunakan dalam satuan rupiah tanpa melalui proses encoding. Ketiga, Status Harga, yang mencerminkan tren harga sebelumnya dan dikodekan sebagai 0 untuk harga turun, 1 untuk harga stabil, dan 2 untuk harga naik. Keempat, Hari ke-, yang menunjukkan hari keberapa dalam setahun (1–365) dan digunakan untuk menangkap pola musiman melalui transformasi sinusoidal. Kelima, Faktorisasi, yaitu representasi faktor eksternal yang memengaruhi perubahan harga, seperti hari besar atau krisis global, yang dikodekan dalam rentang nilai 0 hingga 9. Terakhir, Kondisi, yang menggambarkan situasi pasar atau produksi, misalnya cuaca ekstrem atau lonjakan permintaan, dan juga dikodekan dalam rentang 0 hingga 9. Seluruh variabel ini dirancang agar model LSTM dapat secara efektif mengenali pola dan dinamika harga cabai berdasarkan konteks historis dan eksternal yang relevan.

Seluruh nilai ini diinput secara manual oleh pengguna melalui antarmuka prediksi, dan diolah oleh model untuk menghasilkan proyeksi harga cabai tujuh hari ke depan. *Dashboard* ini dirancang untuk dapat digunakan secara interaktif oleh berbagai pemangku kepentingan, seperti pemerintah, pelaku pasar, maupun konsumen akhir. Fitur utama yang tersedia dalam *dashboard* antara lain:

- Formulir input: Pengguna dapat memasukkan parameter prediksi seperti harga sebelumnya, sub-varian cabai, tanggal (dalam bentuk hari ke- dari awal tahun), status harga terakhir, serta kondisi dan faktor pasar yang relevan. Seluruh parameter disesuaikan dengan hasil proses encoding numerik dari tahap preprocessing.
- Output prediksi: Setelah input dikirimkan, sistem akan menghasilkan prediksi harga cabai untuk tujuh hari ke depan. Hasil ditampilkan dalam bentuk daftar harga harian lengkap dengan status harga (naik, turun, atau stabil), serta divisualisasikan dalam bentuk grafik tren (lihat Gambar 8 dan Gambar 9).
- Visualisasi grafik: Prediksi harga ditampilkan dalam grafik garis yang menunjukkan pola pergerakan harga harian, sehingga memudahkan pengguna dalam memahami tren yang diperkirakan oleh model.

7 Prediksi Harga 7 Hari Berturut-turut

1 April - Rp 10,675 (Tak Stabil)

2 April - Rp 15,886 (Tak Stabil)

4 April - Rp 17,788 (Tak Stabil)

5 April - Rp 19,368 (Tak Stabil)

6 April - Rp 20,694 (Tak Stabil)

7 April - Rp 21,817 (Tak Stabil)

Gambar 8. Prediksi Harga Cabai 7 Hari Berturut-turut

Gambar 8 menunjukkan hasil prediksi harga komoditas selama tujuh hari berturut-turut (1–7 April) dengan tren kenaikan yang konsisten. Meskipun harga terus meningkat, seluruh periode diklasifikasikan sebagai tak stabil, menandakan adanya fluktuasi signifikan yang terdeteksi oleh model.

7 Grafik Prediksi



Gambar 9. Grafik Tren Harga Prediksi 7 Hari ke Depan

Grafik menunjukkan tren kenaikan harga komoditas selama 7 hari berturut-turut, dari Rp 10.675 menjadi Rp 21.817. Meskipun ada sedikit fluktuasi pada 3 April, pola keseluruhan mengindikasikan kondisi pasar yang tidak stabil. Implementasi sistem ini memungkinkan proses prediksi dilakukan secara *real-time*, berbasis data historis yang telah diproses melalui tahapan deep learning. Keunggulan dari pendekatan ini adalah kemampuannya untuk memberikan estimasi harga yang informatif dan responsif terhadap masukan pengguna. Dengan demikian, *dashboard* ini tidak hanya menjadi sarana visualisasi hasil model, tetapi juga berfungsi sebagai alat bantu pengambilan keputusan berbasis data dalam konteks pengendalian harga pangan strategis.

3.3 Pembahasan

Hasil penelitian menunjukkan bahwa pendekatan *Long Short-Term Memory* (LSTM) memiliki performa yang unggul dalam memodelkan harga cabai yang bersifat musiman dan sangat fluktuatif. Model LSTM yang dibangun dievaluasi menggunakan data uji dan menghasilkan nilai Mean Absolute Error (MAE) sebesar 1.402,92, Root Mean Squared Error (RMSE) sebesar 2.595,98, serta koefisien determinasi (R^2) sebesar 0,9567. Nilai MAE dan RMSE yang relatif kecil menunjukkan bahwa rata-rata kesalahan prediksi harga yang dihasilkan oleh model hanya berada pada kisaran dua hingga tiga ribu rupiah, yang tergolong rendah dan dapat diterima mengingat rentang harga cabai yang dalam praktiknya sering kali mengalami perubahan antara Rp10.000 hingga Rp100.000 per kilogram dalam waktu singkat. Sementara itu, nilai R^2 sebesar 0,9567 mengindikasikan bahwa lebih dari 95% variasi harga cabai harian dapat dijelaskan oleh variabel input dalam model, sehingga dapat disimpulkan bahwa model memiliki daya jelaskan (explained variance) yang sangat kuat. Dalam konteks regresi, nilai koefisien determinasi (R^2) di atas 0,90 secara umum dikategorikan sebagai sangat baik, karena menunjukkan bahwa model memiliki kemampuan yang tinggi dalam menjelaskan variabilitas data target. Model LSTM yang digunakan untuk memprediksi harga sayuran menghasilkan R^2 sebesar 0,958, yang mencerminkan kecocokan model terhadap data historis yang sangat tinggi dan performa prediktif yang andal [6]. Studi lain oleh Fikri dan Nurochman [22] yang fokus pada prediksi harga cabai merah dan rawit dengan arsitektur LSTM melaporkan nilai RMSE sekitar Rp 1.751–1.889 dan MAPE terbaik sekitar 3,4%, menunjukkan bahwa rata-rata kesalahan harga hanya berjarak beberapa ribu rupiah dalam konteks jenis cabai lokal Indonesia. Model penelitian ini dengan MAE sebesar Rp 1.402,92 dan RMSE sebesar Rp 2.595,98, serta $R^2 = 0,9567$, juga menunjukkan performa prediktif yang lebih unggul dalam kesalahan mutlak meskipun MAPE tidak disebutkan secara eksplisit.

Meskipun hasil tersebut tergolong baik, pendekatan mereka hanya memanfaatkan variabel seperti harga sebelumnya dan kategori produk sebagai input, tanpa menyertakan representasi waktu siklikal maupun evaluasi model secara menyeluruh menggunakan teknik validasi silang. Penelitian ini memberikan kontribusi signifikan dalam aspek integrasi preprocessing dan sistematisasi proses *modeling*. Salah satu keunggulan utama terletak pada penggunaan transformasi waktu berbasis sin-cos, yang membantu model mengenali pola musiman bulanan yang tidak linier. Hal ini penting mengingat harga cabai sering menunjukkan pola fluktuatif yang berulang setiap tahun, terutama saat hari besar keagamaan dan perubahan musim. Selain itu, proses 5-Fold Cross Validation yang diterapkan menunjukkan bahwa model memiliki performa yang stabil pada data yang berbeda-beda, dengan rata-rata R^2 antar-fold yang tetap berada di atas 0,94, mengindikasikan generalizability yang baik.

Yang tak kalah penting, penelitian ini tidak hanya menghasilkan model yang akurat dari sisi evaluasi metrik, tetapi juga merealisasikan model tersebut ke dalam sebuah aplikasi *dashboard* prediktif berbasis web menggunakan framework *Streamlit* dan bahasa pemrograman Python. *Dashboard* ini memungkinkan pengguna untuk memasukkan input seperti jenis cabai, harga sebelumnya, dan kondisi pasar, lalu menghasilkan prediksi harga secara *real-time* lengkap dengan klasifikasi status harga dan visualisasi tren. Dengan demikian, hasil penelitian ini tidak hanya menjadi produk akademik, tetapi juga memiliki manfaat praktis sebagai sistem pendukung keputusan yang dapat digunakan langsung oleh pemangku kepentingan di sektor pertanian dan pangan. Namun, terdapat beberapa keterbatasan yang perlu dicatat. Model ini belum mengintegrasikan data eksternal dinamis seperti kondisi cuaca, perubahan regulasi, atau informasi stok pasar. Faktor-faktor tersebut berpotensi menurunkan akurasi dan ketepatan prediksi pada situasi pasar yang lebih ekstrem. Selain itu, kendala teknis seperti keterbatasan data yang bersih dan terstruktur juga memengaruhi potensi model dalam skenario produksi berskala besar. Meskipun demikian, hasil penelitian ini telah menunjukkan bahwa pendekatan deep learning berbasis LSTM dapat diterapkan secara nyata untuk prediksi harga komoditas strategis. Pendekatan ini tidak hanya memberikan hasil akurat, tetapi juga berkontribusi dalam transformasi digital di sektor pangan melalui integrasi teknologi prediktif ke dalam sistem berbasis data.

4. KESIMPULAN

Penelitian ini bertujuan untuk membangun model prediksi harga cabai harian berbasis algoritma Long Short-Term Memory (LSTM) dengan pendekatan CRISP-DM sebagai kerangka metodologis. Model dikembangkan untuk menjawab permasalahan fluktuasi harga cabai yang tinggi dan tidak menentu, dengan memanfaatkan keunggulan LSTM dalam menangkap pola musiman dan ketergantungan jangka panjang pada data deret waktu. Hasil pengujian menunjukkan bahwa model memiliki performa yang sangat baik, ditandai oleh nilai koefisien determinasi (R^2) sebesar 0,9567, serta nilai MAE dan RMSE yang rendah pada data uji. Hal ini menunjukkan bahwa model mampu memproyeksikan harga cabai dengan tingkat akurasi yang tinggi dan kesalahan yang relatif kecil. Evaluasi menggunakan skema 5-Fold Cross Validation menunjukkan bahwa model bersifat stabil dan dapat digeneralisasi terhadap data baru, memperkuat keandalannya untuk digunakan dalam konteks prediksi jangka pendek. Salah satu aspek penting dalam keberhasilan model ini adalah penerapan transformasi waktu secara siklikal menggunakan fungsi sinus dan cosinus. Transformasi ini memungkinkan model mengenali pola musiman tahunan secara lebih efektif dibandingkan representasi waktu linier, karena mampu menangkap siklus harga seperti kenaikan menjelang hari besar atau penurunan saat musim panen. Penelitian ini juga mengintegrasikan proses preprocessing yang komprehensif, termasuk pembersihan data, normalisasi fitur numerik, encoding variabel kategorikal, dan pembentukan struktur data tiga dimensi yang sesuai dengan arsitektur LSTM. Seluruh proses ini disusun secara sistematis dan mendukung pengembangan model secara end-to-end, dari data mentah hingga implementasi. Sebagai bentuk pemanfaatan praktis, model yang telah dilatih kemudian diintegrasikan ke dalam dashboard interaktif berbasis web menggunakan *Streamlit*, yang memungkinkan hasil prediksi diakses secara *real-time* oleh berbagai pemangku kepentingan, seperti pedagang, konsumen, maupun pengambil kebijakan. Meskipun model menunjukkan performa teknis yang sangat baik, penelitian ini masih memiliki keterbatasan dalam hal cakupan geografis, karena hanya menggunakan data dari satu wilayah. Untuk meningkatkan kemampuan generalisasi model terhadap kondisi pasar yang beragam, pengembangan lebih lanjut disarankan mencakup integrasi data multiregional dari berbagai daerah. Selain itu, penggunaan variabel input yang lebih kompleks dan kontekstual, seperti faktor cuaca, curah hujan, musim tanam, volume pasokan, dan pengaruh kebijakan pemerintah, juga dapat dipertimbangkan untuk meningkatkan akurasi dan ketahanan model dalam menghadapi dinamika pasar. Penggunaan pendekatan ensemble atau hybrid model juga menjadi strategi potensial untuk memperkuat performa prediksi secara keseluruhan.

REFERENCES

- [1] C. A. Rizqy, N. R. Ali, and K. R. Hayati, "Analisis Pengaruh Kenaikan Harga Bahan Pokok Terhadap Kebutuhan Rumah Tangga Dan Sebagai Tantangan Kegiatan Pkk Di Daerah Ketegan, Taman, Sidoarjo," *Causa J. Huk. Dan Kewarganegaraan*, vol. 3, no. 8, pp. 55–65, 2024, doi: <https://doi.org/10.3783/causa.v3i8.3284>.
- [2] M. R. Widarso and S. Djameluddin, "Analisis Harga Pangan Pokok Terhadap Ketahanan Pangan Di Indonesia," *J. Sos. Ekon. Pertan. dan Agribisnis*, vol. 21, no. 2, pp. 256–272, 2024, doi: <http://dx.doi.org/10.20961/sepa.v21i2.83821>.
- [3] R. L. Manogna, V. Dharmaji, and S. Sarang, "Enhancing agricultural commodity price forecasting with deep learning," *Sci. Rep.*, vol. 15, no. 1, pp. 1–24, 2025, doi: 10.1038/s41598-025-05103-z.
- [4] G. M. Gunadi and A. Perdana, "Forecasting Chili Prices in Metro City Using Long Short-Term Memory (LSTM)," *Int. J.*



- Softw. Eng. Comput. Sci.*, vol. 5, no. 1, pp. 215–225, 2025, doi: 10.35870/ijsecs.v5i1.3526.
- [5] M. Tami and A. Y. Owda, “Efficient commodity price forecasting using long short-term memory model,” *IAES Int. J. Artif. Intell.*, vol. 13, no. 1, pp. 994–1004, 2024, doi: 10.11591/ijai.v13.i1.pp994-1004.
- [6] Q. Zhang, W. Yang, A. Zhao, X. Wang, Z. Wang, and L. Zhang, *Short-term forecasting of vegetable prices based on LSTM model—Evidence from Beijing’s vegetable data*, vol. 19, no. 7 July. 2024. doi: 10.1371/journal.pone.0304881.
- [7] L. Heng, C. Hao, and L. C. Nan, “Load forecasting method based on CEEMDAN and TCN-LSTM,” *PLoS One*, vol. 19, no. 7 July, pp. 1–18, 2024, doi: 10.1371/journal.pone.0300496.
- [8] P. Pandit *et al.*, “Hybrid modeling approaches for agricultural commodity prices using CEEMDAN and time delay neural networks,” *Sci. Rep.*, vol. 14, no. 1, p. 26639, 2024, doi: 10.1038/s41598-024-74503-4.
- [9] H. Jayadianti, V. A. Permadi, and P. Partoyo, “LSTM forecast of volatile national strategic food commodities,” *J. Infotel*, vol. 15, no. 4, pp. 345–351, 2023, doi: 10.20895/infotel.v15i4.1037.
- [10] M. Galphade, V. B. Nikam, B. Banerjee, A. W. Kiwelekar, and P. Sharma, “Stacked LSTM for Short-Term Wind Power Forecasting Using Multivariate Time Series Data,” *Int. J. Interact. Multimed. Artif. Intell.*, vol. 9, no. 3, pp. 71–81, 2025, doi: 10.9781/ijimai.2024.07.002.
- [11] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, “The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks,” *SBC Rev. Comput. Sci.*, vol. 4, no. 1, pp. 28–43, 2024, doi: 10.5753/reviews.2024.3757.
- [12] N. Lundén, E. T. Bekar, A. Skoogh, and J. Bokrantz, “Domain Knowledge in CRISP-DM: An Application Case in Manufacturing,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 7603–7608, 2023, doi: 10.1016/j.ifacol.2023.10.1156.
- [13] B. L. Ortiz *et al.*, “Data Preprocessing Techniques for AI and Machine Learning Readiness: Scoping Review of Wearable Sensor Data in Cancer Care,” *JMIR mHealth uHealth*, vol. 12, p. e59587, 2024, doi: 10.2196/59587.
- [14] J. Kim, H. Kim, H. G. Kim, D. Lee, and S. Yoon, “A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges,” *Artif. Intell. Rev.*, vol. 58, no. 7, 2025, doi: 10.1007/s10462-025-11223-9.
- [15] F. F. Mojtahedi, N. Yousefpour, S. H. Chow, and M. Cassidy, “Deep Learning for Time Series Forecasting: Review and Applications in Geotechnics and Geosciences,” *Arch. Comput. Methods Eng.*, vol. 32, no. 6, pp. 3415–3445, 2025, doi: 10.1007/s11831-025-10244-5.
- [16] W. Witanti, S. Arie Anggara, and M. Melina, “Peramalan Harga Cabai Rawit Merah Menggunakan Attention Mechanism Berbasis Long Short-Term Memory,” *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 2, pp. 128–135, 2024, doi: 10.52158/jacost.v5i2.875.
- [17] T. O. Hodson, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” *Geosci. Model Dev.*, vol. 15, no. 14, pp. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.
- [18] H. Sonavane, G. By, and P. Pankaj, “TradeForecast : Stock Performance and Prediction Web Application,” 2009.
- [19] M. S. Iskandar and Rita Wahyuni Arifin, “Prediction of Red Chili Prices at Musi Market Using Android-Based Linear Regression Algorithm,” *J. Artif. Intell. Eng. Appl.*, vol. 4, no. 3, pp. 2215–2221, 2025, doi: 10.59934/jaiea.v4i3.1138.
- [20] M. Elkabalawy, A. Al-Sakkaf, E. Mohammed Abdelkader, and G. Alfalah, “CRISP-DM-Based Data-Driven Approach for Building Energy Prediction Utilizing Indoor and Environmental Factors,” *Sustainability*, vol. 16, no. 17, p. 7249, 2024, doi: 10.3390/su16177249.
- [21] Y. H. Agustin, F. Nuraeni, and R. Lestari, “Implementation of Multiple Linear Regression Algorithm in Predicting Red Chili Prices in Garut Regency,” *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 2, pp. 304–314, 2024, doi: 10.33480/jitk.v10i2.5882.
- [22] F. N. Fikri and N. Nurochman, “Performance Evaluation of Long Short-Term Memory for Chili Price Prediction,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 10, no. 1, pp. 33–47, 2025, doi: 10.14421/jiska.2025.10.1.33-47.