

# PENERAPAN ALGORITMA LSTM UNTUK PREDIKSI HARGA SAHAM BYD

**APPLICATION OF LSTM ALGORITHM FOR BYD STOCK PRICE PREDICTION**

**Devinka Marta Legawa<sup>1\*</sup>, Muhammad Rafif Rizqullah<sup>2</sup>, Anggraini Puspita Sari<sup>3</sup>**

<sup>1,2,3</sup>Informatika, Fakultas Ilmu Komputer, UPN Veteran Jawa Timur, Surabaya

**Dikirim:**

**Direvisi:**

**Diterima:**

\* Email Korespondensi:  
[anggraini.puspita.if@upnijatim.ac.id](mailto:anggraini.puspita.if@upnijatim.ac.id)



**Abstrak:** Prediksi harga saham merupakan tantangan kompleks karena volatilitas pasar yang tinggi, terutama pada sektor teknologi. Penelitian ini ingin menunjukkan efektivitas LSTM pada saham dengan volatilitas tinggi, karena belum banyak eksplorasi saham teknologi volatilitas tinggi seperti saham BYD. Pengembangkan model prediksi harga saham BYD menggunakan algoritma LSTM dengan optimasi parameter untuk mencapai akurasi tinggi dan menganalisis pengaruh variasi epoch terhadap performa prediksi. Evaluasi dilakukan dengan metrik MAE, RMSE, dan R-squared. Model optimal dicapai pada epoch 50 dengan MAE 0.000340, RMSE 0.000464, dan R<sup>2</sup> 0.947522. Berdasarkan nilai R<sup>2</sup> yang diperoleh, model ini mampu menjelaskan 94.75% variasi pergerakan harga saham BYD, menunjukkan kemampuan prediksi yang sangat baik pada data dengan volatilitas tinggi. Kontribusi penelitian ini meliputi prediksi yang efektif pada saham teknologi bervolatilitas tinggi, hasil optimasi parameter epoch pada arsitektur LSTM untuk data finansial, dan insight penerapan deep learning dalam financial forecasting sektor otomotif teknologi.

**Kata kunci:** LSTM, Stock Price Prediction, BYD Stock, Deep Learning, Time Series

**Abstract:** Stock price prediction presents a complex challenge due to high market volatility, particularly in the technology sector. This research aims to demonstrate the effectiveness of LSTM on high-volatility stocks, as there has been limited exploration of high-volatility technology stocks such as BYD stock. This study develops a BYD stock price prediction model using LSTM algorithm with parameter optimization to achieve high accuracy and analyzes the effect of epoch variations on prediction performance. Evaluation was conducted using MAE, RMSE, and R-squared metrics. The optimal model was achieved at epoch 50 with MAE of 0.000340, RMSE of 0.000464, and R<sup>2</sup> of 0.947522. Based on the obtained R<sup>2</sup> value, this model can explain 94.75% of the variation in BYD stock price movements, demonstrating excellent predictive capability on high-volatility data. The contributions of this research include effective prediction on high-volatility technology stocks, epoch parameter optimization results on LSTM architecture for financial data, and insights into deep learning applications in financial forecasting for the automotive technology sector.

**Keywords:** LSTM, Stock Price Prediction, BYD Stock, Deep Learning, Time Series

## 1. PENDAHULUAN

Prediksi harga saham telah menjadi salah satu tantangan paling kompleks dalam dunia keuangan karena sifat pasar modal yang volatil dan dipengaruhi oleh berbagai faktor eksternal. Harga saham sangat sulit diprediksi pergerakannya karena dipengaruhi oleh perubahan finansial, peraturan industri, kondisi internal perusahaan, dan dampak ekonomi negara [1]. Volatilitas yang tinggi ini dapat mengakibatkan kerugian besar dalam waktu singkat bagi investor yang tidak memiliki strategi prediksi yang akurat [2]. Penelitian ini memilih saham BYD sebagai bahan penelitian didasarkan oleh beberapa pertimbangan. BYD merupakan perusahaan teknologi otomotif

terkemuka yang bergerak pada sektor kendaraan listrik, meningkatnya kompetisi pasar dan sentimen investor terhadap isu energi terbarukan global. Perusahaan ini memiliki kapitalisasi pasar besar dengan volume perdagangan tinggi, menjadikannya representatif saham teknologi yang baik, namun memiliki karakteristik fluktuasi yang menantang dalam pembuatan model prediksi akurat.

Long Short-Term Memory (LSTM) sebagai pengembangan dari RNN telah terbukti mampu mengatasi masalah vanishing gradient dengan mekanisme gate yang dapat mengekstraksi informasi dari data jangka panjang dan sequential. Model LSTM mempunyai tingkat konsisten yang

mengungguli model RNN pada tingkat performance dan stability, spesifik pada data ekonomi dan finansial Indonesia [3]. Penggunaanya telah terbukti mampu menangani berbagai jenis data time series termasuk komoditas pertanian [4]. Pengembangan arsitektur hybrid CNN-LSTM juga menunjukkan hasil yang menjanjikan dengan peningkatan akurasi hingga 42.16%. hasil performa perlu diuji pada saham dengan karakteristik volatilitas teknologi seperti BYD. Keberhasilan model CNN-LSTM dalam menangani data time series dengan model kompleks [5]. Multiple layer yang digunakan LSTM mempunyai akurasi yang lebih tinggi dibandingkan single layer LSTM [6].

Penelitian terdahulu telah menunjukkan efektivitas LSTM dalam prediksi saham di Indonesia. Multiple layer LSTM pada lima emiten syariah mencapai MAPE terbaik 1.51 pada Kalbe Farma hingga 2.66 [6]. Optimasi hyperparameter LSTM pada Jakarta Islamic Index menghasilkan RMSE 5.20877667554 dan MAPE 0.08658576985 [7]. LSTM menunjukkan kemampuan yang baik dalam menangani efek typhoon dibandingkan CNN. Delay dalam proses untuk menghasilkan prediksi pada kondisi perubahan cepat kecepatan dan arah angin dapat dengan model LSTM [8]. Ada beberapa hubungan negatif antara epoch dan RMSE, dari beberapa saham pertambangan dengan emiten paling tinggi -0,873683 dengan 200 epochs [1]. Penerapan pada Bank Central Asia mencapai RMSE 40.85 dan MAPE 0.71% [2].

Long Short-Term Memory (LSTM) pengembangan dari Recurrent Neural Network (RNN). Pemilihan LSTM didasarkan pada kemampuannya mengatasi masalah *vanishing gradient* dan mengekstraksi informasi dari data deret waktu yang panjang dan berurutan[7]. LSTM telah terbukti unggul dalam kinerja dan stabilitas dibandingkan RNN, khususnya untuk data ekonomi dan keuangan, termasuk data deret waktu komoditas pertanian. Penggunaan *multiple layer* pada LSTM juga menunjukkan akurasi yang lebih tinggi dibandingkan *single layer* LSTM. BYD saham dengan karakteristik volatilitas tinggi, Volatilitas pasar sangat mempengaruhi performa prediksi [9].

Prediksi harga saham BYD menggunakan algoritma LSTM mengekstraksi informasi dari data deret waktu yang panjang dan berurutan dari saham BYD, optimasi parameter untuk mencapai akurasi prediksi yang tinggi. Tujuan spesifik meliputi: implementasi arsitektur LSTM dua layer dengan optimasi parameter LSTM untuk data finansial dengan karakteristik fluktuasi ekstrim, dan memberikan insight tentang penerapan deep learning dalam financial forecasting untuk sektor teknologi otomotif; evaluasi performa model menggunakan metrik MAE, RMSE, dan R-squared; analisis pengaruh variasi epoch terhadap akurasi prediksi; dan pengembangan model prediksi yang diharapkan mempunyai variasi pergerakan harga

saham. Penelitian ini diharapkan dapat memberikan manfaat praktis bagi investor dalam pengambilan keputusan investasi pada saham teknologi dan memberikan kontribusi teoritis dalam pengembangan metode prediksi finansial menggunakan machine learning.

## 2. BAHAN DAN METODE PENELITIAN

Perubahan nilai saham BYD Company Limited menunjukkan variasi yang signifikan di pasar saham internasional, mencerminkan perannya sebagai salah satu penghasil kendaraan listrik terkemuka dari China. Analisis perbandingan mengenai kinerja saham perusahaan otomotif listrik menunjukkan bahwa meskipun BYD menunjukkan pola pertumbuhan yang lebih konsisten dibandingkan Tesla, perusahaan ini masih mengalami fluktuasi harga yang signifikan [10]. Faktor-faktor fundamental perusahaan seperti efisiensi pengelolaan biaya, stabilitas aliran kas, dan struktur kewajiban keuangan terbukti menjadi determinan utama yang mempengaruhi tingkat volatilitas saham BYD [11]. Metodologi analisis hierarki dalam penilaian risiko keuangan mengidentifikasi bahwa komponen arus kas dari aktivitas operasional memberikan kontribusi dominan terhadap fluktuasi keuangan perusahaan, sementara lamanya siklus penagihan piutang menciptakan tekanan likuiditas yang berdampak pada ketidakstabilan harga saham [12]. Implementasi model prediksi berbasis time series analisis membuktikan adanya pola clustering dalam volatilitas return harian BYD, menunjukkan kecenderungan bahwa episode fluktuasi tinggi diikuti oleh periode ketidakstabilan serupa [13]. Sektor teknologi otomotif, khususnya kendaraan listrik seperti BYD, penggunaan data historis harga saham saja tanpa mempertimbangkan faktor eksternal seperti sentimen pasar [14], fokus pada satu jenis saham yaitu BYD, dan periode data yang terbatas pada tahun 2002-2025, dengan fokus pada pure LSTM architecture untuk memberikan baseline yang solid sebelum eksplorasi arsitektur yang lebih kompleks di penelitian mendatang.

### Long Short Term-Memory (LSTM)

Long Short-Term Memory (LSTM) adalah Recurrent Neural Network (RNN) yang ditingkatkan dan dirancang khusus untuk mengatasi vanishing gradient. Arsitektur LSTM yang disusun dari tiga gate utama: input gate (i), forget gate (f), dan output gate (o) [15]. Secara matematis, LSTM dapat diformulasikan sebagai:

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

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$h_t = o_t * \tanh(C_t) \quad (5)$$

Fungsi sigmoid, adalah weight matrices, adalah bias vectors, dan t adalah cell state. Kemampuan LSTM untuk mempertahankan informasi jangka panjang melalui cell state memungkinkannya menangkap pola temporal yang kompleks dalam data finansial [16]. Penerapan LSTM pada sektor perbankan juga menunjukkan hasil yang menjanjikan, di mana prediksi saham Bank Central Asia mencapai akurasi tinggi dengan nilai optimal pada epoch 5 dan batch size 1, menghasilkan Mean Absolute Error (MAE) sebesar 96.92 [17]. Berbagai penelitian yang telah dilakukan menunjukkan bahwa jumlah epoch sangat berpengaruh bagaimana performa LSTM dalam prediksi saham. Penelitian pada saham syariah menunjukkan bahwa optimasi epoch dapat mencapai performa yang sangat baik, dengan model menggunakan arsitektur multiple layer LSTM menghasilkan MAPE bervariasi dari 1.51 hingga 2.66 bergantung pada jumlah iterasi pelatihan [16]. Penerapan LSTM pada sektor perbankan menunjukkan bahwa epoch optimal bergantung pada kompleksitas data, dengan variasi epoch 5 menghasilkan konfigurasi terbaik untuk saham Bank Central Asia, menghasilkan Mean Absolute Error (MAE) sebesar 96.92 [17]. Dalam memprediksi harga daging ayam ras segar dengan hasil evaluasi RMSE sebesar 0.0937 dan R2 score sebesar 0.5949 [18]. Optimasi nilai parameter, khususnya bagaimana menentukan berapa jumlah epoch, nilai parameter terbaik untuk model. Penelitian menunjukkan bahwa pengembangan sistem prediksi saham menggunakan LSTM telah mencapai tingkat akurasi yang sangat tinggi, dengan beberapa penelitian melaporkan akurasi prediksi hingga 99% melalui pemilihan epoch yang tepat [19]. Penerapan pada Jakarta Islamic Index menunjukkan bahwa LSTM mampu menangkap karakteristik khusus saham syariah dengan optimasi epoch yang sesuai, memberikan hasil prediksi yang akurat untuk indeks yang mengelompokkan saham-saham syariah [20]. Penelitian lain menunjukkan bahwa epoch yang tidak optimal dapat menyebabkan underfitting atau overfitting, dimana epoch terlalu sedikit menghasilkan model yang tidak belajar dengan cukup, sementara epoch berlebihan dapat menyebabkan model menghafal noise dalam data training. Berbagai aspek ini menjadi pertimbangan penting untuk meningkatkan performa prediksi. Evaluasi performa model menggunakan beberapa metrik standar.

#### *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) adalah sebuah cara untuk mengukur keakuratan sebuah model peramalan[21]. Nilai Mean Absolute Error yang lebih rendah menunjukkan bahwa prediksi lebih mendekati nilai aktual [22]. MAE dinyatakan dalam persamaan sebagai berikut:

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

Di mana  $n$  mewakili jumlah data,  $f_i$  adalah nilai hasil prediksi, dan  $y_i$  merupakan nilai aktual.

#### *Root Mean Squared Error (RMSE)*

Root Mean Squared Error (RMSE) adalah ukuran perbedaan antara nilai prediksi dan observasi model [23]. RMSE berguna untuk mendeteksi pencilan. Nilai RMSE dapat berkisar dari 0 hingga  $\infty$  [24]. Nilai RMSE yang lebih rendah mengindikasikan akurasi model yang lebih tinggi dalam peramalan. Persamaan 2 mengilustrasikan rumus RMSE:  $n$  adalah jumlah data,  $Y_t$  adalah nilai aktual, dan  $\hat{Y}_t$  adalah nilai perkiraan [14].

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (7)$$

#### *R-Square*

R-Squared mengacu pada jumlah varians dalam variabel dependen yang dapat dijelaskan oleh variabel independen [25]. R-Squared mengevaluasi kemampuan model untuk memprediksi hasil. Koefisien determinasi berkisar dari 0 hingga 1, dengan 0 mewakili tingkat terendah dan 1 menunjukkan tingkat tertinggi.

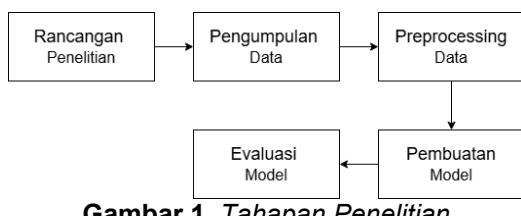
$$R^2 = 1 - \frac{\sum_{i=1}^m (x_i - y_i)^2}{\sum_{i=1}^m (\bar{y} - y_i)^2} \quad (8)$$

#### *MinMaxScaler*

Normalisasi data menggunakan fungsi MinMax adalah salah satu proses dalam preprocessing memiliki fungsi untuk meningkatkan performa hasil, terutama pada data yang terdapat variabel numerik yang memiliki perbedaan rentang yang cukup jauh dari atribut lainnya [25]. Teknik ini memastikan bahwa semua fitur memiliki kontribusi yang seimbang dalam proses training model LSTM, menghindari dominasi fitur dengan skala nilai yang lebih besar terhadap fitur dengan skala yang lebih kecil.

$$X_{scaled} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (9)$$

Penelitian ini menggunakan metode eksperimen yang memanfaatkan algoritma Long Short-Term Memory (LSTM) untuk meramalkan arah pergerakan harga saham BYD. Rencana penelitian sudah dirancang agar dapat memperoleh hasil melalui serangkaian langkah yang terorganisir dari pengumpulan data hingga pelatihan model. Dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

## 2.1 Rancangan Penelitian

Penelitian yang melakukan eksperimen untuk prediksi data time series. Variabel dependen adalah harga penutupan (close price) saham BYD, sedangkan variabel independen adalah sekuens data historis harga penutupan selama 60 hari sebelumnya. Eksperimen yang dilakukan dengan membandingkan kinerja model LSTM dengan beberapa algoritma lain.

## 2.2 Teknik Pengumpulan Data

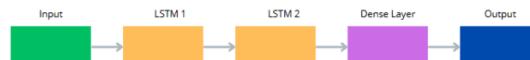
Data sekunder berupa data harga saham BYD diperoleh dari platform Kaggle. Dataset mencakup periode 31 Juli 2002 hingga 16 Maret 2025 dengan jumlah data 4047 mempunyai frekuensi harian, memberikan rentang waktu lebih dari 22 tahun untuk analisis pola time series. Dataset memiliki struktur kolom: Date, Open, High, Low, Close, Volume. Penelitian ini menggunakan kolom 'Close' sebagai target prediksi karena merepresentasikan nilai akhir perdagangan harian yang paling relevan untuk analisis investasi.

## 2.3 Preprocessing Data

Pada proses *preprocessing* data dalam kode ini, langkah pertama yang dilakukan adalah pemilihan fitur multivariat dari data historis saham, yaitu kolom *Open*, *High*, *Low*, dan *Close*. Data tersebut disalin dari *DataFrame* ke dalam variabel baru untuk menjaga integritas data asli. Langkah selanjutnya adalah normalisasi data menggunakan metode *MinMaxScaler* yang akan mengubah setiap nilai dalam fitur-fitur tersebut ke dalam rentang [0, 1]. Proses normalisasi ini sangat penting untuk menyeimbangkan skala antar fitur, mempercepat konvergensi selama pelatihan model, serta mencegah dominasi salah satu fitur terhadap hasil pembelajaran. Setelah proses normalisasi, dilakukan pemisahan data menjadi dua bagian utama, yaitu data latih dan data uji, dengan perbandingan 80% untuk pelatihan dan 20% untuk pengujian. Pemisahan ini bertujuan untuk mengevaluasi performa model secara objektif terhadap data yang tidak digunakan selama proses pelatihan. Keseluruhan langkah preprocessing ini bertujuan untuk mempersiapkan data historis saham secara sistematis dan konsisten, sehingga siap digunakan dalam pelatihan model prediksi berbasis *time series*.

## 2.4 Model LSTM

Model LSTM yang ditunjukkan pada Gambar 2, yang terdiri dari dua LSTM (masing-masing satu lapisan) yang terdiri dari lima komponen berurutan.



Gambar 2. Model LSTM

Model ini dirancang untuk tugas prediksi deret waktu dengan pendekatan *sequence-to-one*, di mana model mempelajari pola temporal dari data historis untuk memprediksi satu nilai di masa depan. Arsitektur dimulai dari Lapisan Input yang menerima data dalam bentuk tiga dimensi: (batch size, 60, 4), di mana 60 adalah *sequence length* (jumlah langkah waktu atau hari historis yang diamati), dan 4 merepresentasikan fitur multivariat (*Open*, *High*, *Low*, *Close*). Data kemudian diproses oleh LSTM Layer pertama dengan 64 unit dan return\_sequences=True, menghasilkan output berdimensi (batch size, 60, 64). Output ini melewati Lapisan Dropout dengan tingkat regularisasi 20% untuk mencegah overfitting, lalu diteruskan ke LSTM Layer kedua dengan 32 unit. Setelah melalui Dropout kedua dengan tingkat yang sama, data dialirkkan ke Dense Layer dan akhirnya ke Output Layer untuk menghasilkan prediksi harga penutupan. Arsitektur ini efektif dalam menangkap ketergantungan temporal dan dinamika kompleks dalam data saham multivariat.

## 2.5 Hyperparameter Tuning

Optimasi hyperparameter dilakukan dengan grid search pada kombinasi parameter: learning rate 0.001, batch size 32, epochs 25, 50, 100, 150, dan dropout rate 0.2. Proses tuning dibungkus dalam fungsi hyperparameter tuning yang menggunakan Early Stopping untuk menghentikan pelatihan saat tidak ada peningkatan. Evaluasi didasarkan pada validation loss, dan model terbaik dipilih dari kombinasi dengan loss terendah. Hasil menunjukkan bahwa jumlah epoch sangat berpengaruh, di mana epoch lebih besar memberi hasil lebih baik selama tidak terjadi overfitting.

## 2.6 Evaluasi Model

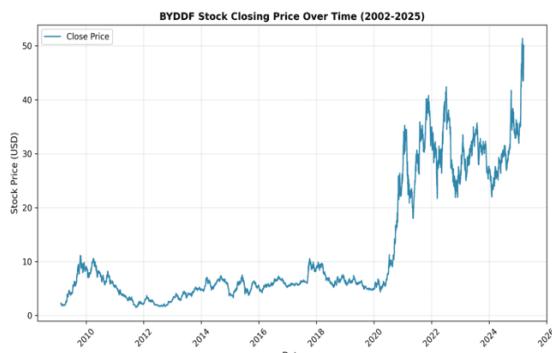
Model dievaluasi menggunakan tiga metrik: MAE, RMSE, dan R-Squared ( $R^2$ ). MAE menunjukkan rata-rata kesalahan absolut, RMSE menekankan kesalahan besar, dan  $R^2$  mengukur seberapa baik model menjelaskan variansi data. Ketiganya memberikan gambaran akurasi dan kinerja model secara menyeluruh.

## 3. HASIL DAN PEMBAHASAN

Pada bagian ini akan membahas implementasi Long Short-Term Memory dalam memprediksi pergerakan harga saham BYD.

### 3.1 Visualisasi Dataset Saham BYD kolom Close

Gambar 3 menampilkan visualisasi data harga penutupan (Close) saham BYD dari periode 31 July 2002 - 16 Maret 2025. Plot ini menunjukkan tentang pergerakan harga saham BYD dalam rentang waktu yang dipilih. Dapat dilihat pada pergerakan harga saham BYD terdapat beberapa harga puncak dan harga terendah, dapat disimpulkan saham BYD memiliki volatilitas yang tinggi. Pola data ini menunjukkan tantangan yang dihadapi dalam memprediksi pergerakan saham di masa depan dengan akurasi yang tinggi.



**Gambar 3.** Harga Saham BYD

### 3.2 Visualisasi Data Prediksi

Rentang data yang dipakai pada visualisasi data dimulai dari tanggal 13 Oktober 2021 - 14 Maret 2025. Berikut hasil pengujian dan visualisasi data berdasarkan beberapa epoch.



**Gambar 4.** Epoch 25

MAE: 0.000370  
RMSE: 0.000490  
 $R^2$ : 0.938766

**Gambar 5.** Evaluasi Matriks Epoch 25

Model LSTM dengan epoch 25 menunjukkan kemampuan prediksi yang cukup baik, tetapi belum optimal, seperti yang ditunjukkan oleh visualisasi data pada Gambar 4 dan evaluasi metrik pada Gambar 5. Dari grafik prediksi, terlihat bahwa garis prediksi (warna ungu) mampu mengikuti pola umum pergerakan harga aktual (warna merah). Namun, ada beberapa waktu di mana ada jarak yang cukup besar, terutama di daerah yang sangat volatil. Hasil evaluasi metrik menunjukkan bahwa model dapat menjelaskan variasi 93.88% dari data, dengan MAE sebesar 0.000370, RMSE sebesar 0.000490, dan

$R^2$  sebesar 0.938766. Meskipun nilai ini cukup baik, masih ada ruang untuk peningkatan akurasi. Ini karena periode yang agak pendek ini menyebabkan model belum sepenuhnya mempelajari pola kompleks dalam data time series saham BYD yang memiliki tingkat volatilitas tinggi.



**Gambar 6.** Epoch 50

MAE: 0.000282  
RMSE: 0.000388  
 $R^2$ : 0.961599

**Gambar 7.** Evaluasi Matriks Epoch 50

Epoch 50 menunjukkan kinerja terbaik dari semua variasi yang diuji, seperti yang ditunjukkan pada Gambar 6 dan Gambar 7, yang menunjukkan evaluasi metrik. Grafik prediksi menunjukkan kesesuaian yang sangat baik antara nilai prediksi dan yang sebenarnya, dengan garis prediksi yang hampir sempurna mengikuti pola pergerakan harga saham BYD. Model bekerja dengan baik dengan MAE terendah sebesar 0.000282, RMSE terendah sebesar 0.000388, dan  $R^2$  tertinggi sebesar 0.961599. Nilai  $R^2$  sebesar 96.16% menunjukkan bahwa model mampu menjelaskan hampir seluruh variasi dalam data, yang menunjukkan tingkat akurasi prediksi yang sangat tinggi. Epoch 50 tampaknya memiliki keseimbangan pembelajaran yang cukup untuk menangkap pola kompleks dalam data tanpa overfitting. Ini menjadikannya konfigurasi yang ideal untuk prediksi harga saham BYD yang sangat volatil.



**Gambar 8.** Epoch 100

MAE: 0.000308  
RMSE: 0.000419  
 $R^2$ : 0.955114

**Gambar 9.** Evaluasi Matriks Epoch 100

Menurut visualisasi yang ditunjukkan pada Gambar 8 dan evaluasi metrik yang ditunjukkan pada Gambar 9, model dengan epoch 100

menunjukkan performa yang lebih rendah daripada epoch 50. Terlepas dari fakta bahwa grafik prediksi terus menunjukkan kemampuan untuk mengikuti pola data aktual, ada tanda-tanda awal overfitting, di mana model mulai terlalu menyesuaikan diri dengan suara dalam data pelatihan. Hasil evaluasi menunjukkan bahwa MAE turun 0.000308, RMSE turun 0.000419, dan R2 turun 0.955114, masing-masing. Nilai R2 turun menjadi 95.51% menunjukkan bahwa kemampuan model untuk menggeneralisasi data baru mulai menurun. Dengan peningkatan epoch yang berlebihan dari titik optimal, model tidak hanya belajar pola yang relevan tetapi juga mulai menghafal karakteristik khusus dari pelatihan data yang tidak dapat digeneralisasi dengan baik pada penilaian data.



**Gambar 10.** Epoch 150

MAE: 0.000302

RMSE: 0.000419

R<sup>2</sup>: 0.955231

**Gambar 11.** Epoch 150

Epoch 150 tidak menghasilkan peningkatan performa yang signifikan dan bahkan cenderung mempertahankan level overfitting yang telah terjadi pada epoch 100, seperti yang ditunjukkan oleh visual pada Gambar 10 dan evaluasi metrik pada Gambar 11. Pola yang mirip dengan Epoch 100 ditemukan pada grafik prediksi. Ini berarti bahwa model masih dapat mengikuti tren umum dalam data aktual, tetapi tidak seakurat Epoch 100. Hasil evaluasi menunjukkan MAE sebesar 0.000302, RMSE sebesar 0.000419, dan R2 sebesar 0.955231, yang hampir sama dengan hasil epoch 100. Melihat metrik ini stabil di bawah epoch 50 menunjukkan bahwa model telah mencapai titik konvergensi yang suboptimal karena overfitting. Fenomena ini menunjukkan bahwa penambahan epoch yang berlebihan tidak selalu menghasilkan peningkatan kinerja sebaliknya, penambahan epoch yang berlebihan dapat menyebabkan model menjadi kurang generalisatif dalam memprediksi data baru. Berikut merupakan tabel perbandingan hasil evaluasi setiap epoch.

**Tabel 1.** Matriks Evaluasi

Epoch	MAE	RMSE	R <sup>2</sup>
25	0.000370	0.000490	0.938766
<b>50</b>	<b>0.000282</b>	<b>0.000388</b>	<b>0.961599</b>
100	0.000308	0.000419	0.955114
150	0.000302	0.000419	0.955231

25	0.000370	0.000490	0.938766
<b>50</b>	<b>0.000282</b>	<b>0.000388</b>	<b>0.961599</b>
100	0.000308	0.000419	0.955114
150	0.000302	0.000419	0.955231

Menurut hasil evaluasi model pada tabel 1 menunjukkan bahwa epoch 50 adalah yang terbaik karena memiliki MAE terendah 0.000282, RMSE terendah 0.000388, dan R2 tertinggi 0.961599. Karena nilai MAE dan RMSE yang rendah, model dapat menjelaskan variasi data dengan baik. Setelah epoch 50, kinerja model mulai menurun, mungkin karena overfitting. Jadi, untuk model ini, epoch 50 adalah waktu terbaik.

#### 4. KESIMPULAN

Studi ini menunjukkan bahwa penyesuaian parameter epoch sangat penting untuk membuat model LSTM untuk prediksi harga saham dengan volatilitas signifikan, seperti BYD. Pada epoch ke-50, model dengan arsitektur dua lapis berkinerja terbaik, dengan akurasi prediksi 96,16% ( $R^2 = 0,961599$ ), MAE terendah (0,000282), dan RMSE terendah (0,000388). Namun, seiring dengan bertambahnya epoch di atas titik ideal (100 dan 150), kinerja menjadi semakin buruk karena overfitting. Temuan ini menunjukkan bahwa LSTM dapat mengelola ekuitas teknologi yang kompleks dan volatil ketika dioptimalkan dengan baik. Ini juga merupakan kontribusi signifikan terhadap penggunaan pembelajaran mendalam untuk peramalan keuangan di sektor teknologi otomotif.

#### DAFTAR PUSTAKA

- [1] R. Julian dan M. R. Pribadi, "Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM)," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 3, hal. 1570–1580, 2021, doi: 10.35957/jatisi.v8i3.1159.
- [2] A. Rosyd, A. Irma Purnamasari, dan I. Ali, "Penerapan Metode Long Short Term Memory (Lstm) Dalam Memprediksi Harga Saham Pt Bank Central Asia," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 1, hal. 501–506, 2024, doi: 10.36040/jati.v8i1.8440.
- [3] C. Alkahfi, A. Kurnia, dan A. Saefuddin, "Perbandingan Kinerja Model Berbasis RNN pada Peramalan Data Ekonomi dan Keuangan Indonesia," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, hal. 1235–1243, Jul 2024, doi: 10.57152/malcom.v4i4.1415.
- [4] M. Lim dan T. Handhayani, "Penerapan LSTM Dan GRU untuk Prediksi Harga Cabai Merah di kota Jawa Timur," *J. Inform. dan Tek. Elektro*

- Terap., vol. 13, no. 2, Apr 2025, doi: 10.23960/jitet.v13i2.6467.
- [5] G. Budiprasetyo, M. Hani'ah, dan D. Z. Aflah, "Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM)," *J. Nas. Teknol. dan Sist. Inf.*, vol. 8, no. 3, hal. 164–172, 2023, doi: 10.25077/teknosi.v8i3.2022.164–172.
- [6] D. R. Chandranegara, R. A. Afif, C. S. K. Aditya, W. Suharso, dan H. Wibowo, "Prediksi Harga Saham Jakarta Islamic Index Menggunakan Metode Long Short-Term Memory," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 1, hal. 129, 2023, doi: 10.26418/jp.v9i1.57561.
- [7] W. Feng, Y. Gao, dan J. Wu, "Yield and Volatility of BYD in Uncertain Market: Evidence from Fed's Monetary Policy," *BCP Bus. Manag.*, vol. 38, hal. 1000–1009, Mar 2023, doi: 10.54691/bcpbm.v38i.3818.
- [8] A. P. Sari, H. Suzuki, T. Kitajima, T. Yasuno, D. A. Prasetya, dan N. Nachrowie, "Prediction Model of Wind Speed and Direction using Convolutional Neural Network - Long Short Term Memory," in *2020 IEEE International Conference on Power and Energy (PECon)*, IEEE, Des 2020, hal. 356–361. doi: 10.1109/PECon48942.2020.9314474.
- [9] M. T. Islam, M. R. Islam, M. S. Faruque, S. M. D. U. Daiam, dan M. M. Islam, "Comparative Stock Performance Analysis of Leading Electric Vehicle Brands: Tesla, BYD, and NIO Using Python Programming Language," *Eur. J. Theor. Appl. Sci.*, vol. 2, no. 4, hal. 327–338, Jul 2024, doi: 10.59324/ejtas.2024.2(4).27.
- [10] "Research on Financial Development Capability of Listed Companies Based on Multiple Regression Model—Take BYD as an Example," *Acad. J. Bus. Manag.*, vol. 4, no. 11, 2022, doi: 10.25236/AJBM.2022.041108.
- [11] Y. Chu, A. Zhang, dan Q. Yu, "Financial Risk Evaluation and Control in New Energy Vehicle Enterprises: A Case Study of BYD," 2025, hal. 355–366. doi: 10.1007/978-3-031-94190-0\_30.
- [12] H. Zou, "Forecast Research of Stock Return Rate of the Three New Energy Auto Companies Based on Three Models," 2025, hal. 289–300. doi: 10.1007/978-96-3236-7\_27.
- [13] A. Mahani dan H. Margono, "Prediksi Sentimen Investor Pasar Modal Di Jejaring Sosial Menggunakan Text Mining," *Balanc. Econ. Business, Manag. Account. J.*, vol. 18, no. 2, hal. 32, 2021, doi: 10.30651/blc.v18i2.7226.
- [14] M. Khan dan Y. Hossni, "A comparative analysis of LSTM models aided with attention and squeeze and excitation blocks for activity recognition," *Sci. Rep.*, vol. 15, no. 1, hal. 3858, Jan 2025, doi: 10.1038/s41598-025-88378-6.
- [15] W. Firgiawan, D. Yustianisa, N. A. Nur, dan G. Gabrelia, "Hyperparameter Tuning for Optimizing Stunting Classification with KNN, SVM, and Naïve Bayes Algorithms," *J. Tekno Kompak*, vol. 19, no. 1, hal. 92, Okt 2024, doi: 10.33365/jtk.v19i1.4574.
- [16] E. S. Putri dan M. Sadikin, "Prediksi Penjualan Produk Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM dan ARIMA," *Format J. Ilm. Tek. Inform.*, vol. 10, no. 2, hal. 162, 2021, doi: 10.22441/format.2021.v10.i2.007.
- [17] J. Cahyani, S. Mujahidin, dan T. P. Fiqar, "Implementasi Metode Long Short Term Memory (LSTM) untuk Memprediksi Harga Bahan Pokok Nasional," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 2, hal. 346, 2023, doi: 10.26418/justin.v11i2.57395.
- [18] Y. Karyadi, "Prediksi Kualitas Udara Dengan Metoda LSTM, Bidirectional LSTM, dan GRU," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 1, hal. 671–684, 2022, doi: 10.35957/jatisi.v9i1.1588.
- [19] A. Hamied Nababan dan M. Y. Hutagalung, "Hyperparameter Tuning Pada Model Stance Detection Menggunakan GridSearchCV," *J. Sains dan Teknol.*, vol. 5, no. 1, hal. 205–209, 2023, doi.org:10.55338/saintek.v5i1.1505
- [20] 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, hal. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.
- [21] A. A. Suryanto, "Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi," *Saintekbu*, vol. 11, no. 1, hal. 78–83, 2019, doi: 10.32764/saintekbu.v11i1.298.
- [22] D. Chicco, M. J. Warrens, dan G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, hal. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.
- [23] A. Mastrucci, A. Marvuglia, E. Benetto, dan U. Leopold, "A spatio-temporal life cycle assessment framework for building renovation scenarios at the urban scale," *Renew. Sustain. Energy Rev.*, vol. 126, hal. 109834, Jul 2020, doi: 10.1016/j.rser.2020.109834.
- [24] F. Rustam *et al.*, "COVID-19 Future Forecasting Using Supervised Machine Learning Models," *IEEE Access*, vol. 8, hal. 101489–101499, 2020, doi: 10.1109/ACCESS.2020.2997311.
- [25] I. Permana dan F. N. S. Salisah, "Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation," *Indones. J. Inform. Res. Softw. Eng.*, vol. 2, no. 1, hal. 67–72, 2022, doi: 10.57152/ijirse.v2i1.311.