

Implementasi Metode LSTM untuk Prediksi Harga Saham PT Indofood CBP Sukses Makmur TBK

Endritha Pramudya^{a1}, Dwi Retnoningsih^{a2}, Diyah Ruswanti^{b3}

^aInformatika, Universitas Sahid Surakarta
Jl. Adi Sucipto No. 54, Jajar, Laweyan, Surakarta, Indonesia
¹endrithapramudya834@gmail.com
²dw1retno.usahid@gmail.com
³dyahruswanti@usahidsolo.ac.id

Abstract

High stock price fluctuations make stock prices difficult to predict accurately. Therefore, a predictive analysis approach that utilizes historical data in addition to machine learning methods is needed to help estimate price movements more effectively. This study aims to determine the performance of the Long Short-Term Memory (LSTM) method in predicting the stock prices of PT Indofood CBP Sukses Makmur Tbk based on historical data. LSTM is a type of artificial neural network that is effective in processing time series data due to its ability to capture long term relationships between data. Historical data is used to train the LSTM model. The results show that the LSTM model is effective in predicting stock prices, with an average accuracy of 80.5%. Sukses Makmur Tbk based on historical data. LSTM is one type of artificial neural network that is effective in processing time series data due to its ability to capture long-term relationships between data points. The data used consists of ICBP stock closing prices from January 2019 to May 2025. The methods used include data cleaning, data normalization, data splitting, model design, prediction, denormalization, and evaluation using the Mean Absolute Percentage Error (MAPE) metric. The research results demonstrate that the LSTM model performs well in recognizing time series data patterns, as indicated by the lowest MAPE value of 1.43, at the combination of 100 epochs and a batch size of 32.

Keywords: time series, long short-term memory, machine learning, stock price prediction, PT Indofood shares.

1. Pendahuluan

Era digital yang semakin berkembang membuat akses terhadap informasi dan layanan investasi menjadi lebih mudah dan cepat. Kemudahan akses melalui platform digital serta meningkatnya literasi keuangan turut berkontribusi pada pertumbuhan jumlah investor di Indonesia. Investasi saham menjadi alternatif menarik karena potensi keuntungan yang signifikan dibandingkan opsi investasi lain [1]. Salah satu saham yang menarik perhatian banyak investor adalah PT Indofood CBP Sukses Makmur Tbk (ICBP), yang beroperasi di sektor barang konsumsi primer. Saat ini ICBP menjadi salah satu komponen indeks LQ45, sebuah indeks yang terdiri dari 45 perusahaan dengan likuiditas dan kapitalitas pasar terbaik. Saham ICBP menempati peringkat ke-14 dalam daftar perusahaan dengan kapitalis pasar terbesar di Bursa Efek Indonesia (BEI) pada Januari 2025. Meskipun memiliki potensi yang menjanjikan harga saham ICBP tetap mengalami fluktuasi yang disebabkan oleh beberapa faktor, seperti kenaikan bahan baku, kebijakan pemerintah, dan kondisi makro ekonomi. Analisis prediktif berbasis data historis diperlukan sebagai upaya menghadapi tantangan tersebut, sekaligus memberikan wawasan yang lebih mendalam terkait pergerakan harga saham di masa mendatang. Pemilihan metode LSTM dalam penelitian ini didasarkan pada kemampuannya dalam mengelola data deret waktu dengan mempertahankan informasi jangka panjang melalui mekanisme gerbang, seperti *input gate*, *forget gate*, dan *output gate*. Mekanisme ini memungkinkan LSTM mengatasi permasalahan *vanishing gradient* yang sering dialami oleh RNN konvensional, sehingga model dapat memahami pola-pola data historis yang kompleks dengan lebih konsisten dan efektif dalam melakukan prediksi. Sejumlah penelitian

sebelumnya telah membuktikan keefektifan LSTM dalam memprediksi harga saham. Hanafiah et al [2], memprediksi harga saham BBNi dengan dua skenario *epoch* (10 dan 20) dan *batch size* 32, di mana *epoch* 20 menghasilkan nilai MAE 0.0150 dan MAPE 0.0257 yang lebih baik dibandingkan *epoch* 10. Riyantoko et al [3], menggunakan LSTM pada saham sektor perbankan, dengan hasil terbaik RMSE 57,31 pada saham BRI menggunakan *optimizer* Adam dan 100 *epoch*. Julian & Pribadi [4], meneliti sektor pertambangan, menunjukkan peningkatan akurasi seiring penambahan *epoch*, dengan hasil optimal RMSE 31,71 pada *epoch* 200 untuk emiten TINS. Kartika & Karmilasari [5], memprediksi harga saham PT Aneka Tambang Tbk menggunakan satu layer LSTM dengan 64 neuron, memperoleh hasil terbaik pada *epoch* 100 dengan RMSE 0.3304 dan R^2 sebesar 0.9902. Penelitian ini akan menggunakan model LSTM untuk memprediksi harga saham ICBP. Penelitian ini dilakukan dengan melakukan uji coba berbagai konfigurasi model LSTM melalui pengaturan parameter seperti jumlah *epoch*, dan *batch size* untuk memperoleh akurasi prediksi yang optimal. Dengan prediksi yang akurat, investor diharapkan memperoleh dasar pertimbangan yang kuat dalam pengambilan keputusan investasi. Tujuan dari penelitian ini adalah mengetahui performa metode *Long Short-Term Memory* (LSTM) dalam memprediksi harga saham PT Indofood CBP Sukses Makmur Tbk berdasarkan data historis.

2. Metode Penelitian

2.1. Pengumpulan Data

Pengumpulan data dilakukan dengan memanfaatkan data historis harga saham yang tersedia secara publik pada Yahoo Finance. Data diunduh menggunakan library *yfinance* di Python, kemudian disimpan dalam format CSV agar dapat dibaca dan diolah pada tahap *pre-processing*. Adapun data yang dikumpulkan mencakup periode Januari 2019 hingga Mei 2025 yang terdiri dari 1559 baris. Dataset ini terdiri dari kolom *Date*, *Open*, *High*, *Low*, *Close*, *Adj Close*, dan *Volume*. Dalam penelitian ini, kolom *Close* digunakan sebagai variabel utama untuk prediksi karena mencerminkan harga penutupan saham yang umum dijadikan acuan dalam analisis.

2.2. Normalisasi

Normalisasi merupakan proses penskalaan ulang data dari nilai aslinya ke dalam rentang tertentu. Normalisasi dilakukan pada *dataset* untuk meminimalkan *error*, mengubah data aktual menjadi nilai dengan rentang 0-1 dengan teknik *MinMaxScaler* [4]. Pengubahan nilai ke dalam rentang 0-1 pada normalisasi *min-max* memiliki tujuan agar nilai yang terdapat pada setiap data memiliki proporsi yang sama dalam pemrosesannya [6]. Rumus dari normalisasi *min-max* ditunjukkan pada persamaan 1.

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

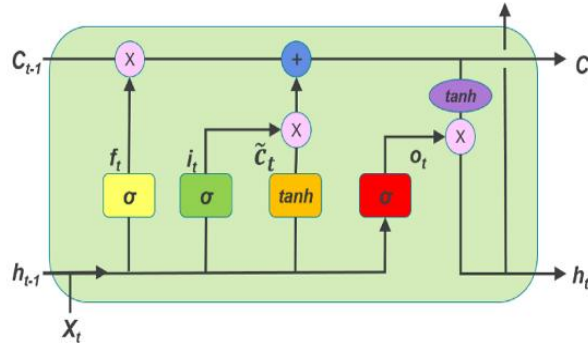
Dimana X' adalah data hasil normalisasi, X adalah nilai data yang akan dinormalisasi, X_{max} adalah nilai maksimal, dan X_{min} adalah nilai minimal.

2.3. Pemisahan Data

Pemisahan data adalah proses membagi dataset menjadi dua atau lebih subset, umumnya terdiri dari data *training* dan data *testing*. Data *training* adalah bagian yang digunakan untuk melatih model, sedangkan data *testing* merupakan data yang diuji dalam model akhir dan dibandingkan dengan kumpulan data sebelumnya [7]. Dalam penelitian ini pembagian data dilakukan menggunakan metode *non-random split* atau *time series split* karena data harga saham bersifat *time series*. Data dibagi menjadi dua bagian yaitu 80% untuk data latih dan 20% untuk data uji.

2.4. Long Short-Term Memory (LSTM)

LSTM merupakan pengembangan dari RNN yang diperkenalkan oleh *Sepp Hochreiter* dan *Jürgen Schmidhuber* pada tahun 1997. LSTM mampu mengatasi *vanishing gradient* atau keadaan nilai gradien adalah 0 atau dekat dengan 0 dengan mekanisme *gate* [4]. *Vanishing gradient* merupakan keadaan yang disebabkan oleh nilai gradien yang selalu mengecil hingga *layer* terakhir sehingga nilai bobot tidak berubah dan menyebabkan model tidak pernah memperoleh hasil yang lebih baik [8].



Gambar 1. Arsitektur Jaringan LSTM [9]

Gambar 1 menunjukkan arsitektur gerbang pada Long Short-Term Memory (LSTM) yang digunakan untuk mengatur aliran informasi di dalam jaringan. LSTM memiliki tiga jenis *gate*, yaitu *input gate*, *forget gate*, dan *output gate*. Ketiga gerbang ini adalah gerbang analog yang menggunakan fungsi sigmoid dengan rentang nilai antara 0 hingga 1 [10].

Forget gate akan menentukan informasi mana yang akan dibuang dari sel memori (*cell state*). *Forget gate* dihitung menggunakan fungsi sigmoid yang menerima h_{t-1} dan x_t yang menghasilkan nilai antara 0 hingga 1. Jika nilai mendekati 0, maka informasi dari *cell state* akan dihapus, dan jika nilai mendekati 1, maka informasi akan disimpan. Persamaan yang digunakan untuk menghitung *forget gate* pada persamaan 2:

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

Input gate menggunakan fungsi aktivasi sigmoid untuk menentukan bagian mana yang akan diperbaharui. Fungsi aktivasi tanh digunakan untuk membuat vektor dari nilai kandidat baru. *Input gate* dihitung menggunakan Persamaan 3 dan 4

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (4)$$

Kemudian *cell state* diperbaharui melalui Persamaan 5.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

Gerbang terakhir, yaitu *output gate* yang berperan untuk menentukan nilai pada bagian memory cell mana yang akan dikeluarkan dengan menggunakan fungsi aktivasi sigmoid [11]. Kemudian *cell state* diperbaharui menggunakan fungsi tanh kemudian dikalikan dengan *output* dari *output gate* untuk menghasilkan *hidden state* yang baru. Persamaan *output gate* dan *hidden state* diuraikan pada Persamaan 6 dan 7.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * \tanh(C_t) \quad (7)$$

Tabel 1. Arsitektur Model

Model: "sequential"		
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 64)	16,896
dropout (Dropout)	(None, 64)	0
dense (Dense)	(None, 1)	65

Model terdiri dari tiga layer, seperti yang ditunjukkan pada Tabel 1. Pertama LSTM *layer* dengan 64 *neuron* yang mempelajari pola jangka panjang pada data *time series* sekaligus menjadi *input layer* yang menerima data dalam bentuk jumlah sampel, panjang urutan waktu (*timestep*), dan fitur. Penelitian ini menggunakan *timestep* 60, sehingga model mempelajari 60 hari data historis untuk memprediksi harga pada hari ke-6. Kedua, yaitu *dropout layer* untuk mengurangi risiko *overfitting* dengan mengabaikan sejumlah unit secara acak selama pelatihan. Terakhir, yaitu *Dense layer (fully connected)* dengan satu unit *output* untuk menghasilkan prediksi harga saham berikutnya.

2.5. Evaluasi

Evaluasi model dilakukan menggunakan metrik *Mean Absolute Percentage Error* (MAPE) dan Koefisien Determinasi (R^2). MAPE adalah suatu metrik evaluasi yang dihitung dengan cara mencari rata-rata dari selisih mutlak antara nilai aktual dan nilai prediksi, lalu dibagi dengan nilai aktual [12]. MAPE digunakan untuk mengukur tingkat kesalahan antara nilai aktual dan nilai prediksi dalam bentuk persentase. Persamaan MAPE dapat dilihat pada Persamaan 8.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (8)$$

Koefisien determinasi (R^2) adalah metrik yang dihitung berdasarkan jumlah kesalahan kuadrat, dan digunakan untuk menunjukkan berapa persen variasi nilai target yang dapat dijelaskan oleh model prediksi [7]. R^2 juga berfungsi untuk menilai tingkat akurasi prediksi. Nilai R^2 berada dalam rentang 0 sampai 1, di mana semakin mendekati 1, maka model dianggap memiliki performa yang semakin baik [5]. R^2 dapat dihitung menggunakan Persamaan 2.10. R^2 digunakan untuk mengukur tingkat akurasi model dalam menjelaskan variasi data.

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

3. Hasil dan Pembahasan

3.1. Dataset

Data yang dikumpulkan mencakup periode Januari 2019 hingga Mei 2025 dengan total data berjumlah 1559. Kolom *Date* berfungsi sebagai penanda waktu yang merepresentasikan urutan data dalam bentuk deret waktu (*time series*). Fokus analisis adalah pada harga penutupan (*Close*) yang digunakan sebagai variabel utama dalam pelatihan dan pengujian model prediksi berbasis *time series*. Sebagian *dataset* harga saham ICBP yang diperoleh dari Yahoo Finance ditunjukkan pada **Gambar 2**.

	Date	Close	High	Low	Open	Volume
0	2019-01-01	9240.448242	9240.448242	9240.448242	9240.448242	0
1	2019-01-02	9196.236328	9328.874352	9019.385630	9284.661677	1844700
2	2019-01-03	9350.981445	9439.406802	9240.449750	9306.768767	3283000
3	2019-01-04	9373.086914	9373.086914	9284.661566	9350.980577	3964600
4	2019-01-07	9152.023438	9527.831163	9041.491753	9461.512153	5581500
5	2019-01-08	8997.279297	9152.023658	8886.747610	9152.023658	10049900
6	2019-01-09	9063.597656	9107.810328	9019.384985	9107.810328	2295000

Gambar 2. Sampel Data Harga Saham ICBP

3.2. Hasil Evaluasi

Penelitian ini dilakukan beberapa percobaan dengan menggunakan nilai *epoch* dan *batch size* yang berbeda untuk memprediksi harga saham ICBP. Percobaan ini dilakukan untuk menentukan nilai *epoch* dan *batch size* efektif untuk melatih model agar menghasilkan tingkat kesalahan prediksi yang rendah yang diukur menggunakan metrik MAPE. Berikut adalah hasil dari percobaan tersebut:

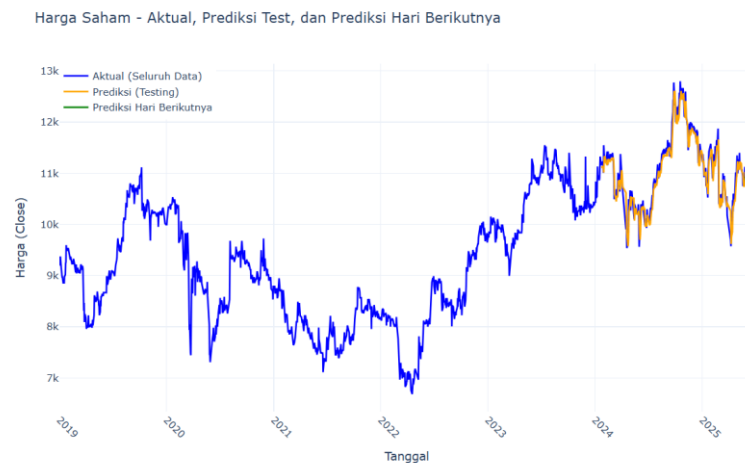
Tabel 2. Perbandingan Jumlah *Epoch* dan *Batch Size*

<i>Epoch</i>	<i>Batch Size</i>	MAPE	R ²	Akurasi
10	32	2.09	0.8265	82.6%
20	32	2.11	0.8290	82.9%
50	32	1.74	0.8834	88.3%
100	32	1.43	0.9203	92.03%
10	64	2.18	0.8197	81.9%
20	64	2.12	0.8180	81,8%
50	64	1.78	0.8708	87.08%
100	64	1.57	0.8988	89.8%

Tabel 2 menunjukkan bahwa model yang paling optimal menggunakan parameter *epoch* sebesar 100 dan *batch size* sebesar 32 memberikan performa terbaik dengan nilai MAPE terkecil sebesar 1.43, koefisien determinasi (R^2) sebesar 0.9203, dan akurasi prediksi sebesar 92.03%. Nilai MAPE menunjukkan bahwa rata-rata kesalahan prediksi terhadap nilai aktual hanya sebesar 1.43 yang mengindikasikan bahwa model memiliki performa yang cukup baik dalam melakukan prediksi harga saham.

3.3. Hasil Prediksi

Setelah dilakukan pelatihan dan evaluasi model terhadap data historis, dilakukan prediksi harga saham untuk 30 hari ke depan terhitung sejak hari terakhir pada dataset yang digunakan. Prediksi ini menggunakan konfigurasi terbaik yang diperoleh dari percobaan, yaitu *epoch* 100 dan *batch size* 32, sehingga hasil yang dihasilkan mencerminkan potensi arah pergerakan harga saham dalam jangka waktu tersebut.



Gambar 3. Grafik Hasil Prediksi

Gambar 3 menampilkan grafik yang membandingkan tiga garis utama, yaitu garis biru sebagai representasi data aktual harga saham, garis kuning sebagai hasil prediksi model terhadap data uji (*testing*), serta garis hijau yang menggambarkan prediksi model terhadap data yang berada di luar rentang pelatihan dan pengujian (prediksi nyata selama 30 hari ke depan). Berdasarkan grafik tersebut, dapat diamati bahwa garis kuning mengikuti arah pergerakan garis biru dengan cukup baik. Hal ini mengindikasikan bahwa model memiliki kemampuan yang baik dalam mengenali pola pergerakan harga saham yang tercermin pada data aktual selama periode pengujian. Sementara itu, garis hijau menunjukkan kelanjutan dari pola tren yang telah terbentuk sebelumnya, yaitu kecenderungan penurunan harga saham secara bertahap setelah akhir periode pengujian.

4. Kesimpulan

Penelitian ini dilakukan untuk mengetahui performa model *Long Short-Term Memory* (LSTM) dalam melakukan prediksi harga saham PT Indofood CBP Sukses Makmur Tbk (ICBP) berdasarkan data historis periode Januari 2019 hingga Mei 2025. Model LSTM dirancang dengan konfigurasi satu *layer* LSTM berisi 64-unit *neuron*. Eksperimen dilakukan dengan berbagai kombinasi parameter *epoch* (10, 20, 50, dan 100) serta *batch size* (32 dan 64). Hasil evaluasi menunjukkan bahwa model mampu mengenali pola data *time series* dengan baik, ditandai dengan nilai *Mean Absolute Percentage Error* (MAPE) terendah sebesar 1.43 pada kombinasi *epoch* 100 dan *batch size* 32. Nilai tersebut mengindikasikan bahwa model memiliki performa prediksi yang cukup akurat terhadap data harga saham yang digunakan.

Daftar Pustaka

- [1] A. Rosyd, A. Irma Purnamasari, and I. Ali, "Penerapan Metode Long Short Term Memory (LSTM) dalam Memprediksi Harga Saham PT Bank Central Asia," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, 2024, doi: 10.36040/jati.v8i1.8440.
- [2] A. Hanafiah, Y. Arta, H. O. Nasution, and Y. D. Lestari, "Penerapan Metode Recurrent Neural Network dengan Pendekatan Long Short-Term Memory (LSTM) Untuk Prediksi Harga Saham," *Bulletin of Computer Science Research*, vol. 4, no. 1, pp. 27–33, Dec. 2023, doi: 10.47065/bulletincsr.v4i1.321.
- [3] P. Aji Riyantoko, T. Maulana Fahrudin, K. Maulida Hindrayani, and E. Maya Safitri, "Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma Long-Short Terms Memory (LSTM)," *Seminar Nasional Informatika (SEMNASIF)*, vol. 1, no. 1, pp. 427–435, 2020, Accessed: Jan. 15, 2025. [Online]. Available: <http://www.jurnal.upnyk.ac.id/index.php/semnasif/article/view/4135>
- [4] R. Julian and M. R. Pribadi, "Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM)," *Jurnal Teknik*

- Informatika dan Sistem Informasi*, vol. Volume 8, no. No 3, pp. 1596–1606, Aug. 2021, [Online]. Available: <http://jurnal.mdp.ac.id>
- [5] S. D. Kartika and Karmilasari, "Implementasi Long Short-Term Memory pada Prediksi Harga Saham PT Aneka Tambang Tbk," *Jurnal Ilmiah Komputasi*, vol. Volume 21, no. No 1, pp. 33–42, 2022, doi: 10.32409/jikstik.21.1.2815.
- [6] G. I. Marthasari, S. A. Astiti, and Y. Azhar, "Prediksi Data Time-series menggunakan Jaringan Syaraf Tiruan Algoritma Backpropagation Pada Kasus Prediksi Permintaan Beras," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. No 3, pp. 187–193, 2021, doi: 10.30591/jpit.v6i3.2627.
- [7] R. Sarno, S. Sabila, M. Malikhah, D. Purbawa, and M. S. Ardani, *Machine Learning dan Deep Learning-Konsep dan Pemrograman Python*. Penerbit Andi, 2023.
- [8] Rowan, L. Muflikhah, and I. Cholissodin, "Peramalan Kasus Positif COVID-19 di Jawa Timur menggunakan Metode Hybrid ARIMA-LSTM," vol. 6, no. 9, pp. 4146–4153, 2022, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [9] K. W. Sari, D. P. Rini, R. F. Malik, and I. S. B. Azhar, "Klasifikasi Teks Multilabel pada Artikel Berita Menggunakan Long Short-Term Memory dengan Word2Vec," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 1, no. No 2, pp. 276–285, 2020, doi: 10.29207/RESTI.V4I2.1655.
- [10] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 2091–2100. doi: 10.1016/j.procs.2020.03.257.
- [11] N. Selle, N. Yudistira, and C. Dewi, "Perbandingan Prediksi Penggunaan Listrik dengan Menggunakan Metode Long Short Term Memory (LSTM) dan Recurrent Neural Network (RNN)," vol. 9, no. 1, pp. 155–162, 2022, doi: 10.25126/jtiik.202295585.
- [12] G. Budiprasetyo, M. Hani'ah, and D. Z. Aflah, "Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM)," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. Volume 8, no. No 3, pp. 164–172, Jan. 2023, doi: 10.25077/teknosi.v8i3.2022.164-172.

Halaman ini sengaja dibiarkan kosong