

Implementasi Model Long Short Term Memory (LSTM) Pada Proyeksi Harga Saham (Studi Kasus: PT. Pertamina Geothermal Energy (Persero))

Dian Christien Arisona^{1*}, Agusrawati², Makkulau³, Irma Yahya⁴, Gusti Ngurah
Adhi Wibawa⁵, Baharuddin⁶, Putri Riski Fahyuni⁷

¹²³⁴⁵⁶⁷Program Studi Statistika, Fakultas Matematika dan Ilmu pengetahuan
Alam, Universitas Halu Oleo, Kendari, 93232, Indonesia

* Corresponding author, email: darisona@uho.ac.id

Abstract

This research presents a comprehensive analysis of the Long Short Term Memory (LSTM) method in projecting the stock price of PT. Pertamina Geothermal Energy (Persero). Utilizing daily stock price data, the LSTM model achieves a high level of accuracy with a Mean Absolute Percentage Error (MAPE) value of 0.84%. The LSTM's gate mechanism (input, forget, output) enables it to store long-term information, controlling the flow of information to update memory, delete irrelevant data, and generate predictions. Optimized with backpropagation through time (BPTT) and activation functions, the LSTM model proves effective in investment decision making, providing valuable insights for investors and market players to anticipate stock price fluctuations. This research demonstrates the great potential of machine learning in financial analysis, particularly in stock price projection and time series analysis. The results indicate that LSTM can be a valuable tool for investors and financial analysts, enhancing their ability to make informed decisions.

Keywords: Stock Price Forecasting, Time Series, LSTM, Financial Analysis, Machine Learning.

Abstrak

Penelitian ini menyajikan analisis komprehensif mengenai penerapan metode Long Short Term Memory (LSTM) dalam memproyeksikan harga saham PT. Pertamina Geothermal Energy (Persero). Dengan memanfaatkan data harga saham harian, penelitian ini menunjukkan bahwa model LSTM dapat memproyeksikan harga saham dengan tingkat akurasi yang tinggi, dengan nilai Mean Absolute Percentage Error (MAPE) sebesar 0,84%. LSTM memiliki mekanisme *gate* (*input*, *forget*, *output*) yang memungkinkannya menyimpan informasi jangka panjang. Setiap *gate* mengontrol aliran informasi: *input gate* menentukan *update* memori, *forget gate* menghapus data tidak relevan, dan *output gate* menghasilkan prediksi. Model ini dioptimalkan dengan *backpropagation through time* (BPTT) dan fungsi aktivasi. Hasil penelitian ini menunjukkan bahwa model LSTM dapat menjadi alat bantu yang efektif dalam pengambilan keputusan investasi dan memberikan wawasan berharga bagi investor dan pelaku pasar dalam mengantisipasi fluktuasi harga saham. Penelitian ini menunjukkan potensi besar dari penerapan teknologi machine learning dalam analisis keuangan, khususnya dalam proyeksi harga saham dan analisis time series.

Kata Kunci: Proyeksi harga saham, Time Series, LSTM, Analisis Keuangan, Machine Learning.

1. Pendahuluan

Peramalan (*forecasting*) merupakan proses memperkirakan nilai masa depan dari suatu variabel atau deret waktu berdasarkan data historis dan informasi yang tersedia. Data deret waktu (*time series*) adalah salah satu metode yang paling umum digunakan dalam peramalan, terutama untuk memprediksi nilai di masa depan berdasarkan pola historis. Pada pasar, peramalan menjadi alat yang penting yang membantu investor dan analis membuat keputusan yang lebih baik, mengidentifikasi peluang, dan mengelola risiko [1].

Metode *time series* seperti ARIMA (*AutoRegressive Integrated Moving Average*), *exponential smoothing*, *Moving Average* (MA), *decomposition* dan sebagainya telah banyak digunakan untuk peramalan *time series*. Namun, metode ini memiliki keterbatasan dalam menangani data yang sangat kompleks, tidak linear, dan data harus memenuhi asumsi seperti kestasioneran dan *white noise*. Oleh karena itu, metode tersebut tidak selalu cocok pada data asli yang seringkali tidak memenuhi asumsi. Dalam beberapa dekade terakhir, metode berbasis *machine learning* dan *deep learning* telah mendapatkan perhatian yang lebih besar karena kemampuannya dalam menangani kompleksitas, non-linearitas dalam data serta tidak memerlukan asumsi yang harus dipenuhi [2].

Recurrent Neural Networks (RNN) adalah salah satu metode *deep learning* yang dirancang khusus untuk data berurutan seperti *time series*. Namun, dalam aplikasi praktis RNN tidak dapat menyimpan informasi sebelumnya dengan baik ketika interval waktu yang panjang karena masalah *vanishing gradient*. Untuk mengatasi kelemahan ini dan meningkatkan kinerja RNN, terdapat dua varian khusus arsitektur RNN yaitu *Long Short Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) [3]. LSTM dapat mengatasi masalah utama yang sering ditemui pada RNN melalui arsitektur uniknya yang mempunyai 3 gate yaitu *forget gate*, *input gate*, dan *output gate* yang memungkinkan jaringan untuk mempertahankan informasi yang relevan dari waktu sebelumnya dan melupakan informasi yang tidak relevan [4].

Berdasarkan penelitian yang dilakukan oleh Karyadi dan Santoso (2022) tentang prediksi kualitas udara di Kota Bandung dalam kurun waktu tahun 2019 menggunakan metode LSTM, bidirectional LSTM dan GRU menjelaskan bahwa penerapan model LSTM dan LSTM *Bidirectional* menunjukkan hasil yang lebih bagus dibandingkan model GRU. Hal ini mengacu pada hasil RMSE model LSTM dan LSTM *Bidirectional* (LSTM-Bi) yang lebih kecil dari RMSE model GRU. Selanjutnya penelitian yang dilakukan oleh Yulisa, dkk (2023) tentang peramalan nilai ekspor migas di Indonesia menggunakan model LSTM dan GRU. Hasil penelitian tersebut menjelaskan bahwa penggunaan arsitektur model LSTM lebih baik karena menghasilkan nilai MAPE sebesar 12.8% sedangkan arsitektur model GRU menghasilkan nilai MAPE sebesar 13.3%.

Berdasarkan sifat, karakteristik dan kelebihan metode LSTM pada penelitian sebelumnya, penelitian ini akan menggunakan metode LSTM untuk memproyeksikan harga saham PT. Pertamina Geothermal Energy. Penelitian yang dilakukan oleh Prasetya, dkk (2020) tentang pemodelan dan peramalan data saham dengan analisis *time series*

menggunakan *python* menyatakan bahwa pergerakan harga saham di Indonesia bersifat fluktuatif, oleh karena itu perlu dilakukan pemodelan dan prediksi untuk mengetahui kondisi dan mempersiapkan strategi untuk menghadapi penurunan atau pelonjakan harga saham. Fluktuasi harga saham ini memberi peluang kepada investor untuk mendapatkan keuntungan maupun kerugian [5]. Metode tradisional memiliki keterbatasan dalam menangani data yang berfluktuasi dan tidak linear, maka untuk menganalisis data ini dapat digunakan metode LSTM. Dengan ketersediaan informasi harga-harga saham masa lalu memberikan peluang secara teoritis untuk dapat dianalisis, diprediksi, diramalkan dan digunakan sebagai bahan pertimbangan informasi harga saham di waktu mendatang [5].

Berdasarkan pembahasan diatas, penulis mengajukan penelitian dengan judul “Implementasi Model *Long Short Term Memory* (LSTM) pada Proyeksi Harga Saham (Studi Kasus: PT. Pertamina Geothermal Energy (Persero))”.

2. Material dan Metode

2.1 Dataset

Data penelitian ini terdiri dari 412 data harga saham PT. Pertamina Geothermal Energy (Persero) yang diperoleh dari laman <https://www.investing.com/equities/pt-pertamina-geothermal-energy-historical-data> dengan rentang waktu dari Februari 2023 hingga Oktober 2024.

2.2 Penanganan *Missing Value*

Missing value adalah informasi yang tidak tersedia untuk sebuah objek. *Missing value* terjadi karena informasi untuk sesuatu tentang objek tidak diberikan, sulit dicari, atau memang informasi tersebut tidak ada. Untuk menangani masalah *missing value* dapat menggunakan metode interpolasi. Interpolasi adalah metode matematika yang digunakan untuk memperkirakan nilai suatu titik di antara dua titik data yang sudah diketahui. Persamaan yang digunakan untuk interpolasi data sebagai berikut [6].

$$y_t = y_{t-1} + \frac{y_{t+1} - y_{t-1}}{x_{t+1} - x_{t-1}} \times (x_t - x_{t-1}) \quad (1)$$

dengan:

y_t : Nilai variabel *high stock* yang dicari

y_{t-1} : Nilai variabel *high stock* pada data $t - 1$

y_{t+2} : Nilai variabel *high stock* pada data $t + 1$

x_t : variabel *date* yang kosong

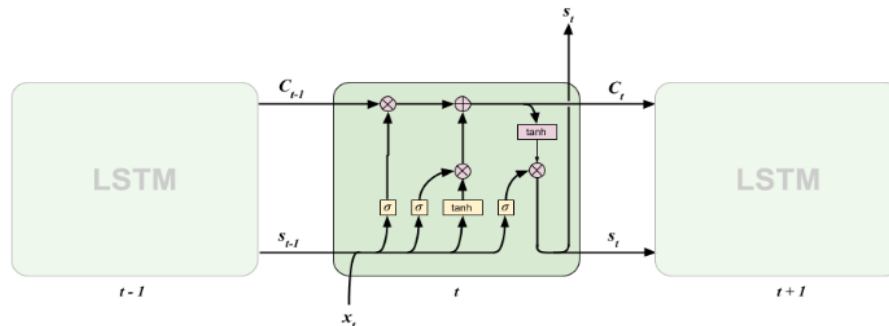
x_{t-1} : variabel *date* pada $t - 1$

x_{t+1} : variabel *date* $t + 1$

2.3 Long Short Term Memory (LSTM)

LSTM memiliki sifat mengembangkan mengikuti urutan waktu, dan banyak digunakan dalam *time series*. Model LSTM terdiri dari rangkaian sel memori unik yang menggantikan neuron lapisan tersembunyi dari RNN, dan kuncinya adalah keadaan sel-

sel memori [7]. Arsitektur LSTM terdiri dari beberapa komponen penting, yaitu *memory block*, *input gate*, *forget gate*, dan *output gate*. *Memory block* berfungsi sebagai penyimpan informasi, sedangkan *input gate*, *forget gate*, dan *output gate* berfungsi sebagai pengontrol aliran informasi masuk, keluar, dan dipertahankan dalam *memory block*. Arsitektur dari LSTM digambarkan dalam diagram pada Gambar 1.



Gambar 1. Arsitektur LSTM [8]

Long Short-Term Memory (LSTM) dapat secara selektif mengontrol aliran informasi dalam dan keluar dari memori jangka panjang dan memungkinkan model untuk mengatasi ketergantungan jangka panjang dan membuat prediksi yang akurat [8].

Hasil kalkulasi pada gerbang LSTM menghasilkan matriks dengan ordo $d \times 1$, sehingga untuk membuat prediksi memiliki satu nilai hasil peramalan perlu menyatukan semua nilai pada matriks hasil kalkulasi gerbang LSTM menggunakan *dense layer*. *Dense layer* merupakan perkalian matriks bobot dan *input* dengan tambahan vektor bias yang nantinya dapat dilatih ke *output*. Secara matematis *output* setiap lapisan *dense layer* didefinisikan dalam persamaan 2 [8].

$$y_t = \sigma([W_{dense} \times h_t] + b_{dense}) \quad (2)$$

dengan:

- y_t : Hasil Prediksi
- W_{dense} : Nilai Bobot dense
- h_t : Nilai Output
- b_{dense} : Bias dense

2.4 Kriteria Evaluasi Model

Untuk menilai kinerja model, beberapa metode evaluasi dapat digunakan, seperti *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE). Dalam penelitian ini, akan digunakan RMSE dan MAPE dalam evaluasi kinerja, Adapun RMSE dan MAPE dapat dilihat pada persamaan (3) dan persamaan (4) berikut:

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

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (4)$$

dengan:

y_i : Data aktual

$\hat{y}_i$: Nilai hasil ramalan

n : Banyak data

3. Hasil dan Diskusi

3.1 Deskripsi Data

Data harga saham PT. Pertamina Geothermal Energy (Persero) divisualisasikan dalam bentuk time series untuk mengidentifikasi pola yang terbentuk. Grafik yang dihasilkan berdasarkan data harian high stock dari tanggal 27 Februari 2023 hingga 7 Oktober 2024, dengan total 412 data.

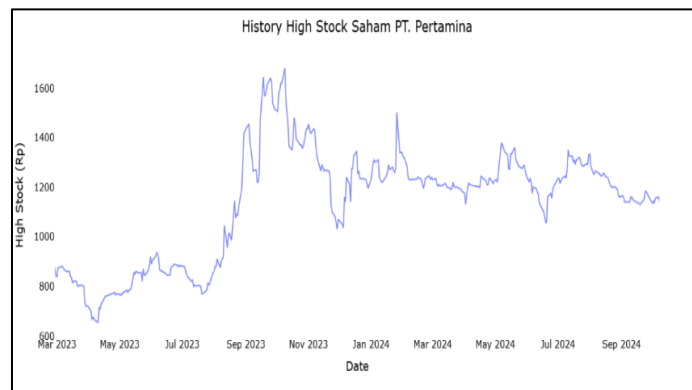


Gambar 2. Grafik Harga *High* Saham PT Pertamina dari 2023 –2024

Analisis data histori harga tertinggi (high stock) saham PT. Pertamina menunjukkan adanya fluktuasi harga dari waktu ke waktu. Namun, terdapat beberapa *missing value* pada data yang disebabkan oleh tidak tercatatnya harga saham pada hari libur nasional, yang ditandai dengan garis putus-putus pada grafik.

3.2 Penanganan Missing Value

Penanganan *missing value* dilakukan untuk menangani data yang hilang (*missing*) dalam dataset agar tidak mempengaruhi analisis statistik dan kinerja model. Setelah melakukan interpolasi, nilai yang kosong pada dataset akan terisi sehingga proses selanjutnya dapat dilanjutkan. Grafik data harga saham setelah dilakukan interpolasi data dapat dilihat pada Gambar 3.



Gambar 3. Grafik Harga *High* Saham PT Pertamina Setelah Interpolasi

Berdasarkan Gambar 3, grafik yang terbentuk tidak lagi menunjukkan garis putus-putus, yang menandakan bahwa tidak ada lagi data yang hilang. Oleh karena itu, data tersebut sudah lengkap dan dapat dianalisis lebih lanjut. Proses selanjutnya membagi data menjadi data latih dan data uji.

3.3 Metode Long Short Term Memory (LSTM)

3.3.1 Penentuan Hyperparameter

Tahap penentuan inisialisasi *hyperparameter* LSTM terdiri dari mendefinisikan *batch size*, *time step*, neuron dan jumlah *epoch* untuk iterasi dalam suatu jaringan LSTM. Untuk menentukan jumlah *neuron*, *epoch* dan algoritma optimasi yang tepat dapat dilihat dari nilai galat terkecil. Pada penelitian ini menggunakan persamaan RMSE sebagai ukuran galat dalam evaluasi model. Hasil pengujian dengan *hyperparameter* LSTM ditampilkan dalam Tabel 1.

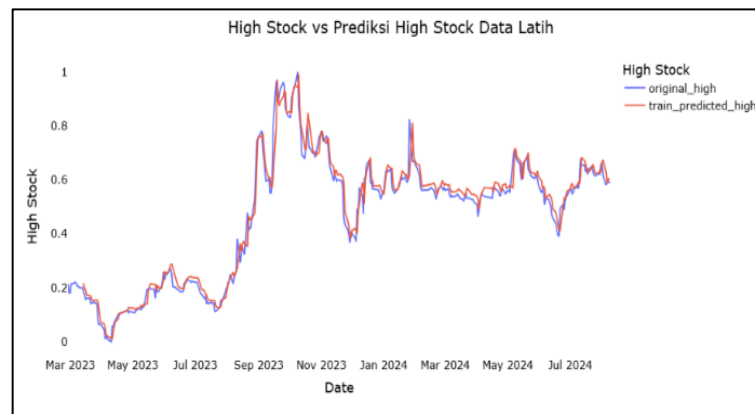
Tabel 1. Nilai RMSE dan MAE Berdasarkan Jumlah *Neuron* dan *Epoch*

Neuron	Epoch	RMSE
20	200	30.84
20	300	34.96
20	400	28.28

Berdasarkan Tabel 1, *hyperparameter* yang optimal yang menghasilkan *Mean Absolute Error* (MAE) terkecil berada pada *epoch* sebesar 400. Sehingga, dapat disimpulkan bahwa model dengan pengaturan *hyperparameter* ini mampu memberikan hasil peramalan yang paling akurat pada data Saham PT Pertamina.

3.3.2 Pemodelan

Setelah mendapatkan nilai *hyperparameter* yang tepat, selanjutnya melakukan pemodelan menggunakan nilai tersebut. Berikut grafik perbandingan data aktual dan data hasil peramalan pada data latih setelah dilakukan *tunning hyperparameter*.

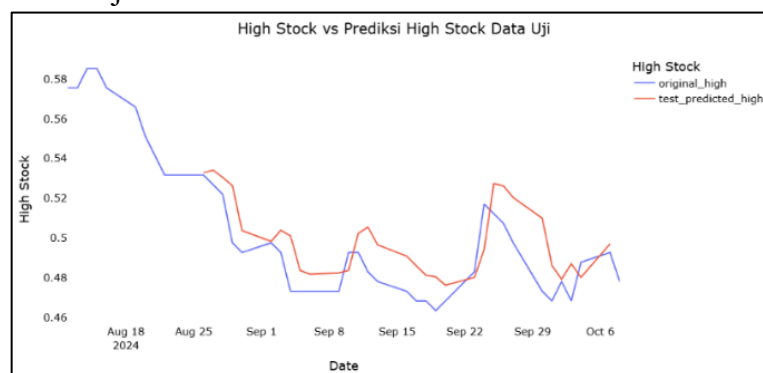


Gambar 4. Grafik *High Stock* vs Prediksi *High Stock* Data Latih

Gambar 4 menunjukkan perbandingan antara data aktual dan hasil peramalan high stock saham PT. Pertamina untuk data latih. Garis merah mewakili data prediksi, sedangkan garis biru mewakili data aktual. Dari grafik tersebut, terlihat bahwa hasil peramalan mengikuti pola data aktual dengan cukup baik, meskipun terdapat beberapa data di mana peramalan sedikit berbeda dari data aktual. Hal ini menunjukkan bahwa model peramalan yang digunakan mampu menangkap pola data harga saham dengan akurasi yang cukup tinggi selama periode pelatihan.

3.3.3 Evaluasi Model

Setelah tahap pemodelan, tahap selanjutnya yaitu evaluasi model dengan cara melakukan peramalan menggunakan data uji untuk melihat seberapa baik model bekerja pada data yang belum pernah dilihat sebelumnya. Dalam beberapa periode, hasil peramalan cukup mendekati data aktual, sementara pada periode lainnya terdapat perbedaan yang lebih signifikan. Berikut grafik perbandingan data aktual dan data hasil peramalan pada data uji.



Gambar 5. Grafik *High Stock* vs Prediksi *High Stock* Data Uji

Gambar 5 menunjukkan perbandingan antara data aktual dan hasil peramalan harga tertinggi saham PT. Pertamina untuk data uji. Dari Gambar 5 terlihat bahwa data hasil prediksi mengikuti pola data aktual, meskipun terdapat beberapa perbedaan di mana

peramalan sedikit lebih tinggi atau lebih rendah daripada data aktual. Hal ini menunjukkan bahwa model peramalan memiliki kemampuan yang cukup baik dalam memperamalan data uji, meskipun masih ada ruang untuk perbaikan dalam beberapa periode tertentu.

3.3.4 Akurasi Peramalan

Ketepatan hasil peramalan pada data latih dan data uji dapat dievaluasi dengan menggunakan *Mean Absolute Percentage Error* (MAPE). Dengan menggunakan persamaan 4, diperoleh nilai MAPE untuk data latih sebesar 1.94% sedangkan untuk data uji diperoleh nilai MAPE sebesar 0.83%. Nilai MAPE dibawah 10% dianggap sebagai kriteria yang sangat baik dalam keakuratan peramalan.

3.3.5 Proyeksi harga Saham Harian

Setelah mendapatkan nilai *hyperparameter* yang optimum, selanjutnya melakukan peramalan dengan model tersebut untuk sepuluh hari kedepan. Hasil peramalan menggunakan metodel LSTM ditampilkan pada Tabel 2.

Tabel 2. Proyeksi High Stock Saham

Tanggal	Prediksi <i>High Stock</i>
09-10-24	1151.15
10-10-24	1150.98
11-10-24	1148.67
14-10-24	1145.72
15-10-24	1142.60

Prediksi ini mencerminkan tren penurunan harga, yang menunjukkan potensi pertumbuhan saham dalam periode tersebut. Pergerakan harga yang diprediksi ini dapat mencerminkan dinamika pasar yang stabil atau adanya faktor yang mendukung penurunan harga saham pada jangka waktu tersebut. Pemodelan menggunakan data uji ini memberikan gambaran yang dapat digunakan untuk evaluasi tren harga saham di masa mendatang.

4. Kesimpulan

Model yang telah dioptimalkan dengan *hyperparameter* terbaik dan optimasi Adam dapat digunakan untuk memproyeksikan *high stock* saham PT. Pertamina Geothermal Energy pada periode-periode mendatang, yang dibuktikan dengan akurasi peramalan yang tinggi pada data uji. Model LSTM yang digunakan untuk meramalkan harga saham PT. Pertamina menunjukkan performa yang sangat baik dengan nilai MAPE sebesar 1.94% untuk data latih dan 0.83% untuk data uji. Nilai tersebut menunjukkan akurasi yang sangat baik karena berada di bawah 10%.

Daftar Pustaka

- [1] Widiyanto, M. H., Mayasari, R., & Garno. Implementasi Time Series Pada Data Penjualan Di Gaikindo Menggunakan Algoritma Seasonal ARIMA. *Jurnal Mahasiswa Teknik Informatika*, 7(3), 1501-1506, 2023.
- [2] Arwansyah, A., Suryani, S., Hasrif, S. Y., Usman, U., Ahyuna, A., & Samsu Alam. Time Series Forecasting Menggunakan Deep Gated Recurrent Units. *Digital Transformation Technology*, 4(1), 410–416, 2022.
- [3] Yulisa, P. N., Haris, M., & Arum, P. R. Peramalan Nilai Ekspor Migas di Indonesia dengan Model Long Short Term Memory(LSTM) dan Gated Recurrent Unit(GRU). *Jurnal Statistika*. 16(1), 328 -341, 2023.
- [4] Karyadi, Y. Prediksi Kualitas Udara Dengan Metoda LSTM, Bidirectional LSTM, dan GRU. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(1), 671-684, 2022.
- [5] Anggraeni, D. T. Forecasting Harga Saham Menggunakan Metode Simple Moving Average Dan Web Scrapping. *Jurnal Ilmiah Matrik*, 21(3), 234-241, 2019.
- [6] Lamabelawa, M. I. J. Analisis Perhitungan Metode Interpolasi Pada Data Time Series Kemiskinan Di NTT. *Jurnal (HOAQ) – Teknologi Informasi*, 8(1), 640-646, 2018.
- [7] Rahmah, A. L. *Analisis Model Multivariate Long Short-Term Memory Untuk Prakiraan Kualitas Udara Dki Jakarta Berdasarkan Data Tahun 2010-2022*. Jakarta: Program Studi Sistem Informasi Fakultas Sains Dan Teknologi Universitas Islam Negeri Syarif Hidayatullah, 2024.
- [8] Ramadhan, M. F. *Prediksi Harga Bitcoin Menggunakan Metode Long Short Term Memory*. Jambi: Program Studi Sistem Informasi Fakultas Sains Dan Teknologi Universitas Jambi, 2024.