

PERAMALAN HARGA SAHAM DENGAN METODE LONG SHORT TERM MEMORY (LSTM)

Mathew Judianto ¹⁾ Teny Handhayani ²⁾ Janson Hendryli ³⁾

^{1,2,3)} Teknik Informatika, FTI, Universitas Tarumanagara

Jl. Letjen S Parman No. 1, Jakarta 11440 Indonesia

email : mathew.535200047@stu.untar.ac.id¹⁾, tenyh@fti.untar.ac.id²⁾, jansonh@fti.untar.ac.id³⁾

ABSTRACT

Sistem prediksi harga saham di Indonesia menggunakan algoritma machine learning, yaitu Long Short Term Memory (LSTM). Data yang dipilih sebagai faktor pengaruh dalam sistem ini adalah harga Brent Crude Oil. Tujuan utama dari sistem ini adalah memberikan nilai prediksi saham untuk 7 hari ke depan, yang dapat membantu masyarakat umum dan investor untuk lebih memitigasi risiko yang timbul dari pembelian saham. Data yang digunakan adalah harga penutupan Brent Crude Oil dan harga penutupan PT. Astra International Tbk yang diperoleh dari situs Yahoo Finance yang terdiri dari data dari Januari 2018 hingga Januari 2023. Data yang digunakan telah dinormalisasi dan diproses sebelum dimasukkan ke dalam model LSTM. Hasil terbaik dapat dicapai dengan menggunakan 80% data sebagai dataset pelatihan, timestep 20, 120 epochs, dan batch size 64. Dengan pengembangan sistem ini, diharapkan masyarakat umum dan investor di Indonesia dapat lebih memitigasi risiko dari pembelian saham yang dianggap sebagai investasi berisiko tinggi.

Key words

saham, Long Short Term Memory, Brent Crude Oil

1. Pendahuluan

Kondisi perekonomian global sangat dipengaruhi oleh interaksi kompleks antar banyak variabel ekonomi, seperti inflasi, peristiwa geopolitik, dan harga dari energi serta komoditas [1]. Diantara faktor tersebut, harga minyak mentah merupakan salah satu faktor energi yang memiliki dampak besar pada berbagai sektor perekonomian, seperti sektor transportasi, pertambangan, dan energi yang merupakan pilar penting bagi kegiatan perekonomian. Salah satu tantangan utama yang dihadapi oleh investor dalam menentukan sektor yang akan dijadikan instrumen investasi adalah kesulitan dalam memilih waktu yang baik untuk membeli saham di sektor tertentu. Ini disebabkan oleh sektor-sektor tersebut sangat dipengaruhi oleh faktor eksternal yang seringkali sulit diprediksi.

Sektor transportasi sangat terdampak oleh fluktuasi harga minyak mentah karena ketergantungan yang tinggi pada bahan bakar berbasis minyak. Perusahaan transportasi, termasuk penerbangan dan logistik, terkait

erat dengan ketersediaan dan harga bahan bakar fosil. Kenaikan harga minyak mentah menyebabkan meningkatnya biaya bahan bakar transportasi, berpotensi merugikan profitabilitas perusahaan dan mempengaruhi harga saham mereka [2].

Ketika melibatkan faktor-faktor tersebut, menjadi sangat penting bagi para investor untuk dapat menentukan waktu yang tepat dalam melakukan investasi pada sektor yang relevan. Fluktuasi harga minyak, bersama dengan efeknya yang merata ke berbagai sektor, seperti transportasi, pertambangan, dan energi, dapat dijadikan landasan untuk melakukan peramalan harga saham dalam sektor-sektor ini. Ini adalah langkah kunci dalam membantu para investor mengidentifikasi peluang investasi yang menguntungkan.

Dalam mengatasi permasalahan diatas, maka dirancangnya sistem prediksi ini yang akan dikembangkan sebagai aplikasi berbasis web yang akan menggunakan bahasa pemrograman Python dengan framework Django. Sistem ini akan menerapkan model Deep Learning Recurrent Neural Network yaitu Long Short Term Memory (LSTM). Sistem ini akan memberikan informasi mengenai hasil prediksi serta memberikan saran atau rekomendasi mengenai waktu yang baik untuk membeli atau menjual serta akan menjelaskan hasil prediksi berdasarkan visualisasi grafik yang ditampilkan.

Dalam konteks ini, LSTM akan menjadi solusi yang efektif karena tantangan utama adalah menangkap hubungan yang kompleks antara faktor-faktor penentu dan pergerakan harga saham. Hubungan ini perlu dianalisis secara mendalam dan dalam jangka waktu yang panjang menggunakan data yang tersedia. Kemampuan LSTM untuk memahami dan memodelkan keterkaitan data jangka panjang menjadikannya alat yang sangat berguna dalam mengatasi tantangan tersebut.

Penelitian sejenis dengan menggunakan metode LSTM telah banyak dilakukan, diantaranya Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) menggunakan Long Short Term Memory (LSTM) oleh Roby Julian dan Muhammad Rizky [3]. Selanjutnya, Prediksi Data Time Series Saham Bank BRI dengan Mesin Belajar LSTM (Long Short Term Memory) oleh Adhitio Satyo Bayangkari Karmo [4]. Kemudian,

Forecasting Saham Syariah dengan Menggunakan LSTM oleh Ahmad Fauzi [5]. Selanjutnya, Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma *Long-Short Terms Memory* LSTM oleh Prismahardi Aji Riyantoko, Tresna Maulana Fahrudin, Kartika Maulida Hindrayani, dan Eristya Maya Safitri [6].

2. Sistem Persamaan Linier

2.1 Metode Long Short Term Memory (LSTM)

Long Short-Term Memory (LSTM) merupakan jenis *Recurrent Neural Network* dengan 3 gate yaitu Forget Gate, Input Gate, dan Output Gate yang telah dimodifikasi yang diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997 [10]. Perubahan yang dilakukan pada metode LSTM ini adalah dengan menambahkan sel memori yang dapat menyimpan informasi dalam waktu yang lama. Berikut merupakan tahap metode *Long Short Term Memory* (LSTM) [9]:

Pertama, *forget gate* berfungsi untuk melakukan penghapusan informasi yang tidak lagi berguna dalam cell state. Berikut persamaan untuk *forget gate*[9]:

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

Keterangan :
 f_t = forget gate
 σ = fungsi sigmoid
 W_f = bobot pada forget gate
 h_{t-1} = hasil output timestep t-1
 x_t = input pada timestep x
 b_f = bias pada forget gate

Kedua, *input gate* berfungsi untuk mengontrol berapa banyak informasi yang harus disimpan dalam *cell state*, berikut persamaan untuk *input gate* [9]:

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

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \dots\dots\dots(3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot C'_t \dots\dots\dots(4)$$

Keterangan :
 i_t = input gate
 σ = fungsi sigmoid
 W_i = bobot pada input gate
 h_{t-1} = hasil output timestep t-1
 x_t = input pada timestep x
 b_i = bias pada input gate
 C'_t = kandidat cell state baru
 \tanh = fungsi tanh
 W_c = bobot operasi update cell state
 b_c = bias pada operasi update cell state

Output Gate berfungsi untuk mengontrol berapa banyak informasi yang harus disimpan dalam cell state, berikut persamaan untuk input gate [9]:

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

$$h_t = O_t \cdot \tanh(C_t) \dots\dots\dots(6)$$

Keterangan :
 O_t = output gate
 σ = fungsi sigmoid
 W_o = bobot untuk output gate
 h_{t-1} = hasil output timestep t-1
 x_t = input pada timestep x
 b_o = bias pada output gate
 h_t = hasil output pada time step t
 \tanh = fungsi tanh

2.2 Mean Squared Error (MSE)

MSE merupakan Jumlah selisih kuadrat antara data nilai yang dihasilkan oleh model dan data nilai aktual yang dibagi dengan jumlah titik historis, data tersebut lalu akan dikurangi dengan jumlah parameter dalam model. Jumlah parameter dalam model akan digunakan untuk mengurangi angka dari titik histori agar menghasilkan hasil yang konsisten [7]. Berikut ini persamaan *mean squared error* [8]:

$$MSE = \frac{\sum(Y_i - \hat{Y}_t)^2}{n} \dots\dots\dots(7)$$

Keterangan :
 Y_i = nilai aktual
 \hat{Y}_t = nilai prediksi dari variabel Y
 n = jumlah periode yang terlibat

2.3 Mean Absolute Percentage Error (MAPE)

MAPE merupakan rata rata dari perbedaan antara data nilai yang dihasilkan oleh model dengan data nilai aktual. Semakin kecil tingkat persentase tersebut, maka prediksi akan dinyatakan lebih akurat [7]. Berikut rumus dari persamaan *Mean Absolute Percentage Error*:

$$MAPE = \left(\frac{100}{n}\right) \sum \left| A_t - \frac{F_t}{A_t} \right| \dots\dots\dots(8)$$

Keterangan :
 A_t = nilai aktual pada periode -t
 F_t = nilai prediksi pada periode -t
 n = total periode prediksi

2.4 R-Squared (R²)

R-Squared (R²) merupakan ukuran yang memberikan informasi mengenai tingkat akurasi dari sebuah model. Dalam konteks regresi, ini adalah ukuran statistic dari seberapa baik garis regresi mendekati nilai actual, berikut rumus dari persamaan *R-Squared* (R²):

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \dots\dots\dots(9)$$

3. Hasil Percobaan

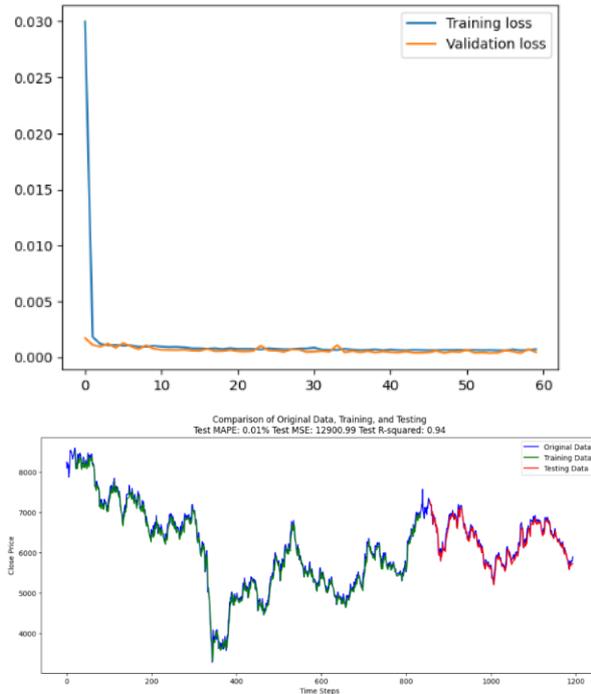
Hasil percobaan sistem peramalan harga saham sektor transportasi, energi, dan pertambangan dengan harga minyak sebagai faktor pengaruh menggunakan metode *Long Short Term Memory* (LSTM) akan dilakukan pengujian model terhadap sistem dengan dipertimbangkan berdasarkan hasil metrik evaluasi *Mean Squared Error* (MSE) dan *Mean Absolute Percentage Error* (MAPE) serta akan menggunakan hasil nilai *R-Squared* (R^2). Pengujian model akan dilakukan sebanyak 10 kali dengan nilai learning rate 0.01 dan layer 50 serta perubahan nilai *hyperparameter* lainnya seperti yang tertera pada tabel 1.

Tabel 1 *Hyperparameter* Pengujian Model LSTM

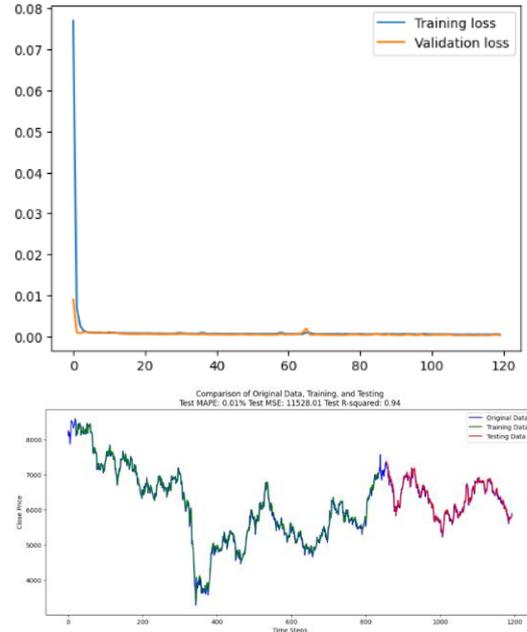
No.	Data Training (%)	Timestep	Epoch	Batch Size
1.	70	20	60	32
2.	70	20	120	64
3.	70	30	60	32
4.	70	30	120	64
5.	80	20	60	32
6.	80	20	120	64
7.	80	30	60	32
8.	80	30	120	64

Pada hasil pengujian, diperoleh grafik sebagai berikut:

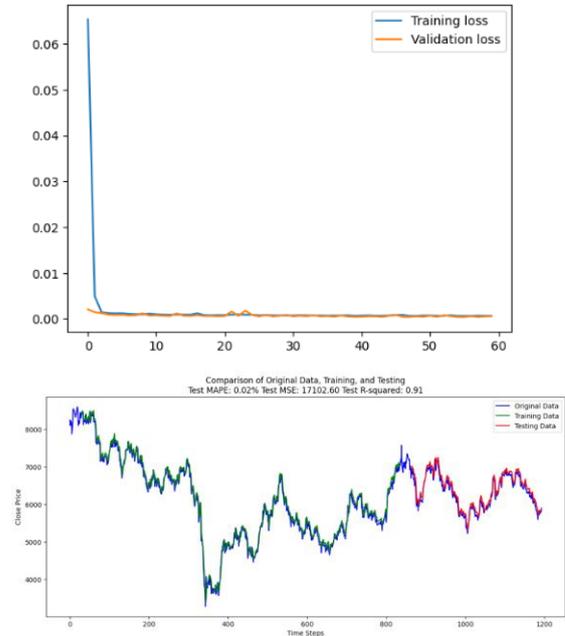
1. Hasil pengujian pertama



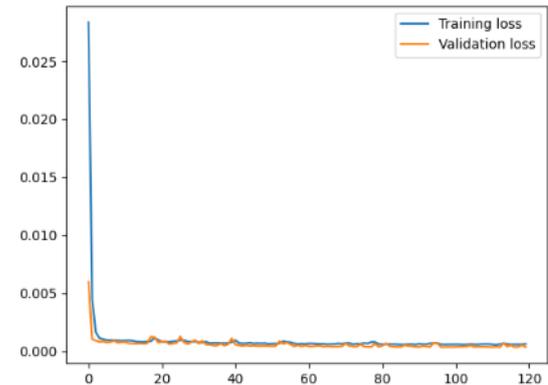
2. Hasil pengujian kedua

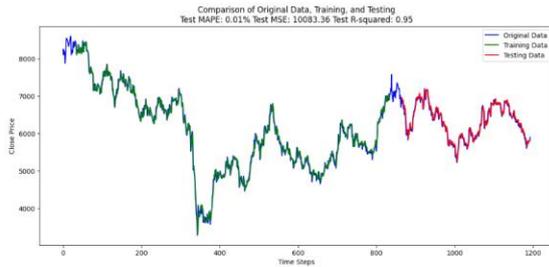


3. Hasil pengujian ketiga

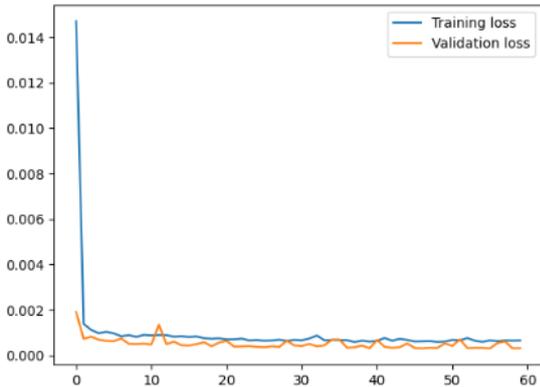


4. Hasil pengujian keempat

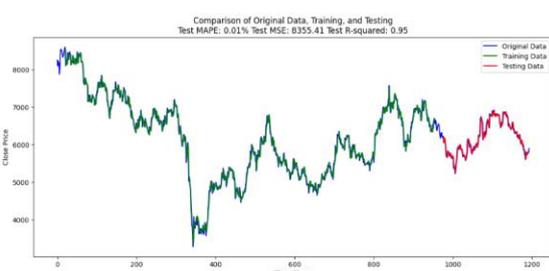
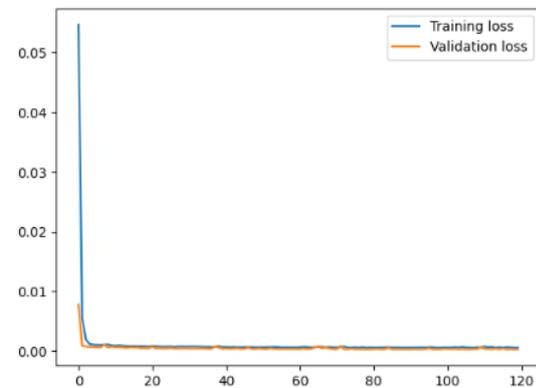




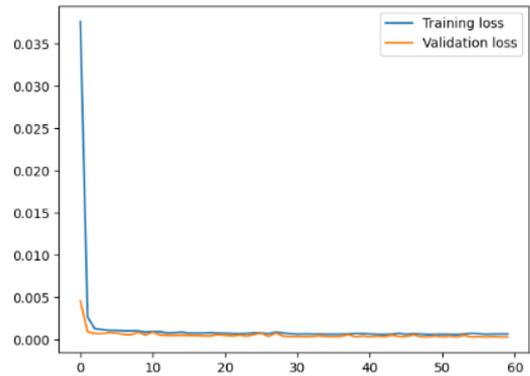
5. Hasil pengujian kelima



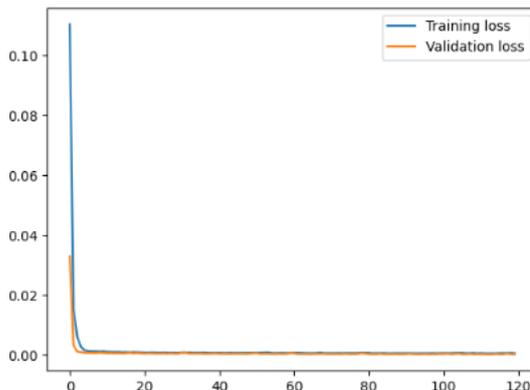
6. Hasil pengujian keenam



7. Hasil pengujian ketujuh



8. Hasil pengujian kedelapan



Nilai MSE, RMSE, dan R-Squared dari hasil 8x pengujian dapat dilihat pada tabel 2.

Tabel 2 Hasil Pengujian

Pengujian ke-	R-Squared	MSE	MAPE (%)
1.	0.94	12900.999	0.01
2.	0.94	11528.01	0.01
3.	0.91	17102.60	0.02
4.	0.95	10083.36	0.01
5.	0.95	8859.42	0.01
6.	0.95	8355.41	0.01
7.	0.95	8538.25	0.01

8.	0.95	9518.32	0.01
----	------	---------	------

Dapat dilihat pada tabel 2., bahwa pada pengujian pertama dan kedua memiliki hasil yang konsisten dengan *R-Squared* 0.94, MSE berkisar antara 11528.01 hingga 12900.99 dan MAPE sekitar 0.01%, meskipun parameter *timestep*, *epoch*, dan *batch size* nilainya bervariasi. Pada pengujian ketiga, terdapat seikit penurunan performa dengan nilai *R-Squared* sebesar 0.91, MSE sebesar 17102.60, dan MAPE sebesar 0.02%. Hal ini menunjukkan bahwa kombinasi parameter kurang optimal atau tidak sesuai dengan pola data. Pengujian keempat hingga kedelapan, terlihat peningkatan yang konsisten dalam performa model. *R-Squared* mencapai 0.95, nilai MSE yang menurun menjadi berkisaran antara 8355.41 hingga 10083.36, dan nilai MAPE yang stabil pada 0.01. Hal ini menandakan bahwa kombinasi parameter pada pengujian keempat hingga kedelapan memiliki hasil yang konsisten dan hasil yang baik.

4. Kesimpulan

Berdasarkan hasil percobaan dapat disimpulkan bahwa :

1. Model dengan parameter data training 70% dan 80% memiliki hasil yang konsisten baik dan tidak terdapat perbedaan signifikan dalam performa model.
2. Variasi *timestep* 20 menghasilkan performa yang lebih baik dengan nilai *R-Squared* yang tinggi dan nilai MSE yang rendah.
3. Model cenderung memperoleh pola yang lebih baik dengan iterasi yang lebih banyak.
4. Hasil pengujian terbaik terdapat pada pengujian keempat, dimana

REFERENSI

[1] P. Forgarty, "5 Key Factors Influencing the Global Economy According to Munich Re's 2023 Economic Outlook," Risk & Insurance, 21 February 2023. [Online]. Available: <https://riskandinsurance.com/5-key-factors-influencing-the-global-economy-according-to-munich-res-2023-economic-outlook/>. [Diakses 21 September 2023].

[2] M. R. Ramadan, "Pengaruh Inflasi, BI Rate dan Harga Minyak Dunia Terhadap Harga Saham Sektor Transportasi Periode Tahun 2013-2017," Universitas Widyatama, 2019. [Online]. Available: <https://repository.widyatama.ac.id/items/33497e7b-32ab-4116-9e7c-abef680df441>. [Diakses 21 September 2023].

[3] R. Julian dan 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. 8, no. 3, pp. 1570-1572, 2021.

[4] A. Fauzi, "Forecasting Saham Syariah Dengan Menggunakan LSTM," *Jurnal Lembaga Keuangan dan Perbankan*, vol. 4, no. 1, pp. 1, 66-68, 2019.

[5] A. S. B. Karno, "Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long ShortTerm Memory)," *JIFORTY*, vol. 1, no. 1, pp. 1-4, 2020.

[6] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani dan E. M. Safitri, "Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma Long-Short Term Memory (LSTM)," *Seminar Nasional Informatika*, vol. 1, no. 1, pp. 427-428, 2020.

[7] IBM, "Forecasting statistical details," 3 January 2023. [Online]. Available: <https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=forecasting-statistical-details>. [Diakses 22 September 2023].

[8] D. Kurniasih, "Efisiensi Relatif Estimator Fungsi Kernel Gaussian Terhadap Estimator Polinomial Dalam Peramalan USD Terhadap JPY," Universitas Negeri Semarang, November 2013. [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/ujm/article/view/3245/2985>. [Accessed 25 Agustus 2023].

[9] J. Schmidt, "What Are Stocks? How Do They Work?," *forbes*, 6 July 2023. [Online]. Available: <https://www.forbes.com/advisor/investing/what-are-stocks/>. [Diakses 22 September 2023].

Mathew Judianto, mahasiswa S1, program studi Teknik Informatika, Fakultas Teknologi Informasi Universitas tarumanagara.

Teny Handhayani, S.Kom., M.Kom., Ph.D., Dosen program studi Teknik Informatika, Universitas Tarumanagara.

Janson Hendryli, S.Kom., Dosen program studi Teknik Informatika, Universitas Tarumanagara.