

Stock Market Index Prediction using Bi-directional Long Short-Term Memory

Muhammad Althaf Majid ^{1*}, Prilyandari Dina Saputri ^{2*}, Soehardjoepri ^{3*}

* Institut Teknologi Sepuluh Nopember
prilyandaridina@its.ac.id ²

Article Info

Article history:

Received 2024-02-03
Revised 2024-05-21
Accepted 2024-06-06

Keyword:

bi-LSTM,
Global Index,
IHSG,
LSTM.

ABSTRACT

The IHSG (Indonesia Stock Exchange Composite Index) is a stock price index in the Indonesia Stock Exchange (BEI) that serves as an indicator reflecting the performance of company stocks through stock price movements. Therefore, IHSG becomes a reference for investors in making investment decisions. Advanced stock exchanges generally have a strong influence on other stock exchanges. Several studies have proven the influence of one global index on another. Global index is a term that refers to each country's index to represent the movement of its country's stock performance. Forecasting IHSG can be one of the analyses that help investors make wise decisions when investing. To obtain an IHSG forecasting model, an appropriate and suitable method is needed, especially for data that has a large amount. LSTM is a development of Recurrent Neural Network (RNN) which has the ability to remember information in a longer period of time, while Bi-LSTM is a development of LSTM which has the ability to remember information longer and can understand more complex patterns than LSTM. This research provide the IHSG forecasting based on global index factors. The results showed that the best Bi-LSTM model (6-9-1) had a better performance in predicting and forecasting JCI movements with a MAPE value of 0.572314% better than the best LSTM model (4-10-1) which had a MAPE of 0.74326%. With forecasting based on the Bi-LSTM model, it is expected to help investors in making decisions on the Indonesia Stock Exchange (IDX).



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Pasar modal Indonesia terus berkembang dengan pertumbuhan jumlah investor mencapai 11.581.153 hingga Agustus 2023. Investasi di pasar modal dianggap populer dan meyakinkan karena dapat membantu individu dan perusahaan memperoleh penghasilan pasif di masa depan. Namun, investasi tidak terlepas dari risiko, dan setiap instrumen investasi memiliki tingkat return dan risiko yang berbeda. Saham, sebagai instrumen dengan tingkat risiko tinggi, memerlukan pemilihan yang bijak dan analisis yang kuat. Indeks Harga Saham Gabungan (IHSG) menjadi acuan utama bagi investor di Indonesia, mencerminkan performa saham-saham perusahaan di Bursa Efek Indonesia (BEI).

Beberapa indeks harga saham global, seperti NASDAQ Composite Index (IXIC), New York Stock Exchange

Composite Index (NYSE), Shanghai SE Composite Index (SSEC), Tokyo Stock Price Index (TOPIX), Financial Times Stock Exchange All Shares Index (FTSE All Shares) dapat memiliki dampak yang positif dan signifikan terhadap IHSG [1], [2]. Menurut penelitian yang pernah dilakukan oleh Jonathan [3], SSE Composite Index (SSEC) memberikan dampak yang signifikan dan positif terhadap IHSG. Indeks-indeks harga saham global, terutama di kawasan asia dapat digunakan sebagai faktor peramalan terhadap Indeks Harga Saham Gabungan (IHSG) karena terbukti mampu menggambarkan dampaknya terhadap IHSG.

Pada umumnya, peramalan IHSG dilakukan menggunakan metode *Auto-Regressive Integrated Moving Average* (ARIMA) [4], *Long Short-Term Memory* (LSTM) [5], *Gated Recurrent Unit* (GRU) [6]. Salah satu metode yang baik untuk peramalan IHSG adalah *Recurrent Neural Network* (RNN).

Berbeda dengan *Artificial Neural Network* (ANN), *Recurrent Neural Network* (RNN) merupakan *deep learning* model dengan struktur yang sederhana dan pengembangan dari ANN. RNN dapat mempertahankan dan mengingat informasi masa lalu dengan tujuan menggunakan informasi tersebut untuk memprediksi informasi di masa berikutnya sehingga sering digunakan untuk melakukan *forecasting* dan pengolahan pada data-data *time series*. Namun, RNN akan menjadi kurang efektif untuk digunakan apabila pada data dengan jumlah besar. Hal ini karena secara teori, banyaknya langkah yang harus ditempuh sebanding dengan banyaknya *layer*. Selain itu, metode RNN juga tidak dapat mengatasi permasalahan *vanishing gradient* dan *exploding gradient* untuk jaringan RNN yang sangat dalam [7]. Terdapat beberapa cara yang dapat digunakan untuk permasalahan tersebut. Cara yang paling populer adalah dengan menggunakan metode *Long Short-Term Memory* (LSTM) dan *Bidirectional Long Short-Term Memory* (*bi-LSTM*).

Metode LSTM merupakan perluasan dari metode RNN, sedangkan metode *bi-LSTM* sendiri merupakan perluasan dari metode LSTM layaknya menggabungkan dua LSTM dipersatukan, namun dengan arah yang berlawanan. Kedua metode ini mampu mengatasi *vanishing gradient* dengan sangat bersih. Hal ini karena memori yang disimpan jauh lebih luas sehingga dapat mengingat informasi untuk waktu yang lebih lama dan memiliki kemampuan untuk membaca, menulis, dan menghapus informasi dari memori tersebut [7]. Model LSTM sudah cukup efektif dalam meramal data-data *time series* dengan jangka waktu yang panjang dan dapat memberikan akurasi yang baik, namun terkadang kurang efektif jika dibandingkan dengan model *bi-LSTM*. Hal ini karena *bi-LSTM* dapat menggabungkan dan mengalirkan informasi dari dua arah, baik dari masa lalu ke masa depan maupun dari masa depan ke masa lalu, sedangkan LSTM hanya dapat melakukannya dari masa lalu ke masa depan sehingga pemahaman konteksnya, cara mengatasi masalah hilangnya informasi, dan pemodelan dari model *bi-LSTM* lebih unggul [4].

Penelitian yang pernah dilakukan oleh Siami-Namini [4] dilakukan perbandingan antara model ARIMA, LSTM, dan *bi-LSTM* dalam melakukan peramalan harga saham dan diperoleh kesimpulan bahwa model *bi-LSTM* memiliki akurasi terbaik dibandingkan model lainnya. Penelitian yang pernah dilakukan oleh Gangwar dkk. [8], dilakukan perbandingan prediksi model antara SVM dan LSTM dalam memprediksi kecepatan angin dan diperoleh kesimpulan bahwa model LSTM memiliki akurasi terbaik dibandingkan model SVM. Penelitian yang pernah dilakukan oleh Lammoreno [9], dilakukan perbandingan prediksi model antara LSTM dan GRU dalam memprediksi IHSG berdasarkan faktor indeks global dan diperoleh kesimpulan bahwa model GRU memiliki akurasi lebih baik dibandingkan model LSTM. Penelitian yang pernah dilakukan oleh [10], dilakukan perbandingan prediksi model menggunakan CNN (*Multilayer*), CNN-RNN, CNN-LSTM, RNN (*Single layer*), LSTM (*Multilayer*) peramalan harga saham dan diperoleh

kesimpulan bahwa model RNN (*Singlelayer*) memiliki akurasi model paling baik dibandingkan dengan model lainnya. Penelitian dari Zaheer dkk. [10] menyatakan bahwa LSTM yang merupakan pengembangan dari RNN tidak selalu menghasilkan model prediksi yang lebih baik. Oleh karena itu, adanya model modifikasi dari LSTM, yaitu *bi-LSTM* diharapkan mampu untuk membuktikan bahwa *Bi-LSTM* memberikan performa jauh lebih baik dibandingkan LSTM. Selain itu, perbandingan model juga dilakukan untuk membuktikan bahwa model *Bi-LSTM* tidak selalu menjadi model yang lebih baik dibandingkan model LSTM. Hal ini karena adanya kemungkinan *overfitting* dan bentuk model yang lebih kompleks dibandingkan model LSTM.

Berdasarkan uraian yang telah disebutkan di atas, pada penelitian ini dilakukan peramalan data harga penutupan (*close price*) harian IHSG berdasarkan faktor indeks global periode 1 Januari 2018 – 30 Juni 2023 menggunakan metode *Bi-LSTM*. Peramalan dengan menggunakan metode LSTM juga dilakukan untuk mengetahui perbandingan hasil peramalan antara model LSTM dan *bi-LSTM*. Dengan adanya penelitian ini, diharapkan dapat membantu investor dalam pengambilan keputusan di Bursa Efek Indonesia (BEI).

II. METODE PENELITIAN

A. Data dan Variabel Penelitian

Data yang akan digunakan dalam penelitian ini merupakan data sekunder berbentuk deret waktu (*time series*), yaitu data harga penutupan (*close price*) harian indeks saham global selama 4 tahun 6 bulan (1 Januari 2018 – 30 Juni 2023) yang dikumpulkan melalui situs *finance.yahoo.com*. Variabel penelitian yang akan digunakan dalam penelitian ini disajikan pada Tabel I.

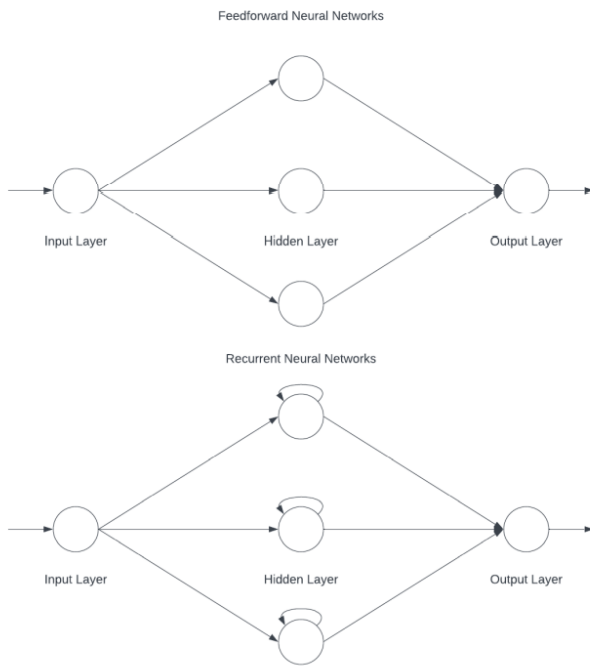
TABEL I
VARIABEL PENELITIAN

Variabel	Definisi
IHSG (Y)	Harga Penutupan Harian Indeks Harga Saham Gabungan
IXIC (X_1)	Harga Penutupan Harian NASDAQ Composite
FTSE (X_2)	Harga Penutupan Harian FTSE All Share Index
SSEC (X_3)	Harga Penutupan Harian Shanghai SE Composite Index

B. Recurrent Neural Network

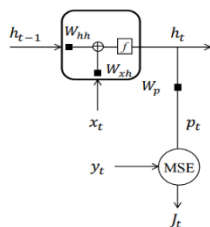
Recurrent Neural Network atau RNN merupakan salah satu jenis arsitektur *Neural Network* dengan struktur sederhana. RNN adalah model pengembangan dari FFNN atau *Feed Forward Neural Network* yang mana dirancang khusus untuk menganalisis data-data yang mengandung deret waktu. Namun, perbedaannya adalah *Recurrent Neural Network* dapat mempertahankan dan mengingat informasi masa lalu dengan tujuan menggunakan informasi tersebut untuk memprediksi informasi di masa berikutnya. Hal ini menyebabkan RNN sering digunakan para peneliti untuk melakukan *forecasting* dan pengolahan pada data-data *time series* [7]. Secara umum, walaupun RNN mampu untuk memahami urutan data, RNN masih memiliki kelemahan. Kelemahan dari RNN adalah kesulitan untuk melakukan

pemodelan terutama untuk pemodelan dengan jumlah data yang besar dan waktu yang jangka panjang sehingga sering timbul permasalahan berupa lenyapnya informasi (*vanishing gradient*) atau meledaknya informasi (*exploding gradient*) pada proses pelatihan data [4]. *Vanishing gradient* adalah keadaan ketika gradien menjadi sangat kecil selama *backpropagation* sehingga lapisan atau *layer* dalam jaringan tidak berubah secara signifikan karena gradien terlalu kecil untuk membuat perubahan pada parameter, sedangkan *exploding gradient* adalah keadaan ketika gradien menjadi sangat besar selama *backpropagation* sehingga pembaruan parameter menjadi terlalu besar dan melewati titik konvergensi yang optimal. Arsitektur dari FFNN dan RNN disajikan pada Gambar 1 [11].



Gambar 1. Contoh Arsitektur FFNN dan RNN

RNN memiliki tiga layer, yaitu *input layer*, *hidden layer*, dan *output layer*. Gambaran struktur sel di dalam setiap *hidden neuron* pada *hidden layer* RNN disajikan pada Gambar 2.



Gambar 2. Struktur Sel *Hidden Layer Recurrent Neural Network*

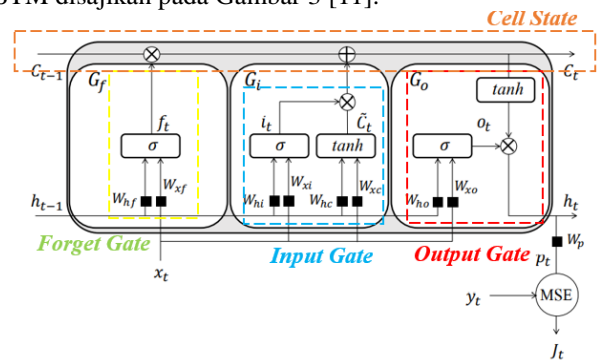
Setiap *hidden layer* memiliki fungsi aktivasi masing-masing yang berfungsi untuk menentukan keberlanjutan perjalanan *neuron* atau sel pada *hidden layer*. Selain itu juga terdapat bobot dan bias pada *input* saat ini dan *state*

sebelumnya. Persamaan yang digunakan pada RNN dapat dituliskan sebagai [11].

$$h_t = \tanh(W_{hh} \cdot h_{t-1} + W_{xh} \cdot x_t + b_x + b_h). \quad (1)$$

C. Long Short Term Memory

Long Short-Term Memory atau LSTM merupakan sebuah model yang dikembangkan untuk mengatasi permasalahan yang terdapat pada RNN. LSTM memiliki komponen sel memori yang berfungsi mengingat dan menggunakan informasi yang ada dalam urutan data dengan jarak waktu yang lebih jauh. Gambaran struktur sel dan arsitektur pada LSTM disajikan pada Gambar 3 [11].



Gambar 3. Arsitektur dan Struktur Sel *Hidden Neuron LSTM*

Long Short-Term Memory terdapat lima komponen utama. *Forget gate* yang berfungsi untuk menyeleksi informasi mana yang harus dihapus. Informasi yang bertahan dari *cell state* sebelumnya akan diteruskan ke *cell state* pada waktu ke- t [12]. Persamaan dari *forget gate* dapat dituliskan sebagai

$$f_t = \sigma(W_{hf} \cdot h_{t-1} + W_{xf} \cdot x_t + b_f). \quad (2)$$

Komponen *input gate* dan *candidate state*. *Input gate* berfungsi untuk menentukan informasi yang harus ditambahkan ke dalam *cell state*, informasi ini akan dikumpulkan dalam *candidate state* berfungsi untuk dilakukan pembaruan informasi [12]. Persamaan dari *input gate* dan *candidate state* dapat dituliskan sebagai

$$\tilde{c}_t = \tanh(W_{hc} \cdot h_{t-1} + W_{xc} \cdot x_t + b_c), \quad (3)$$

$$i_t = \sigma(W_{hi} \cdot h_{t-1} + W_{xi} \cdot x_t + b_i). \quad (4)$$

Komponen *cell state*, yaitu komponen sel memori yang berfungsi untuk menggabungkan perhitungan *input gate*, *candidate state*, dan *forget gate* [12]. *Cell state* nantinya akan menyimpan dan mempertahankan informasi tersebut untuk pelatihan model, dengan persamaan [12]

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

Komponen *Output gate* berfungsi untuk mengendalikan banyaknya informasi dari *cell state* yang perlu diteruskan ke *hidden state* melalui fungsi aktivasi *sigmoid*, sedangkan *hidden state* mengandung informasi yang telah diolah dan akan diteruskan ke *neuron* berikutnya melalui fungsi aktivasi *tanh* untuk mengubah rentang nilainya menjadi antara -1 dan

1 [12]. Persamaan dari *output gate* dan *hidden state* dapat dituliskan sebagai [7].

$$O_t = \sigma(W_{oh} \cdot h_{t-1} + W_{ox} \cdot x_t + b_o) \quad (6)$$

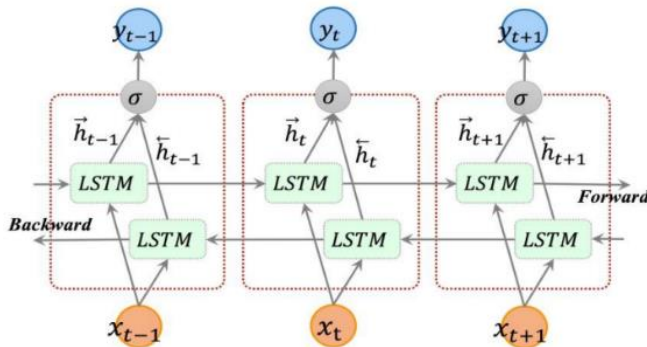
$$h_t = O_t \cdot \tanh(C_t) \quad (7)$$

Keterangan :

- O_t = output waktu t
- h_t = Hidden State waktu t
- C_t = Cell State waktu t
- f_t = Forget state waktu t
- \tilde{C}_t = Candidate state waktu t
- W_{oh} = Bobot output pada hidden state pada output gate
- x_t = Input waktu t
- W_{ox} = Bobot input pada output gate
- b_o = Bias pada output gate

D. Bidirectional Long Short-Term Memory

Bidirectional Long Short-Term Memory atau *Bi-LSTM* merupakan variasi pengembangan dari *Long Short-Term Memory*. Perbedaannya adalah model ini menggabungkan dua model LSTM, yaitu *forward LSTM* dan *backward LSTM* yang bergerak berlawanan arah waktu bersamaan [4]. Arsitektur dari model *Bi-LSTM* disajikan pada Gambar 4 [12].



Gambar 4. Arsitektur Model Bi-LSTM

Output akhir pada *Bi-LSTM* ditulis sebagai h_t dimana h_t dirumuskan sebagai.

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (8)$$

Keterangan :

- h_t = Hidden state waktu t
- \vec{h}_t = Hidden state waktu t *Forward LSTM*
- \overleftarrow{h}_t = Hidden state waktu t *Backward LSTM*

III. HASIL DAN PEMBAHASAN

A. Pre-processing Data

Tahapan awal dalam proses analisis adalah proses penggabungan data, yang bertujuan untuk melakukan sinkronisasi antara data-data saham dengan waktu yang sesuai sehingga memiliki rentang waktu dan jumlah yang sama. Hasil penggabungan data diilustrasikan pada Tabel II.

TABEL II
OBSERVASI DATA PENELITIAN

Periode	Observasi	IHSG	IXIC	FTSE	SSEC
1/2/2018	1	6339,238	7006,9	4202,53	3348,326
⋮	⋮	⋮	⋮	⋮	⋮
31/12/2020	755	Missing	12888,28	3673,63	3473,069
⋮	⋮	⋮	⋮	⋮	⋮
30/6/2023	1434	Missing	Missing	4096,26	Missing

Berdasarkan ilustrasi Tabel II di atas, variabel IHSG akan menjadi acuan untuk penyesuaian jumlah data variabel lainnya karena variabel IHSG menjadi output pada penelitian. Selanjutnya dilakukan penanganan *missing value* untuk mengatasi permasalahan data-data yang hilang, namun sangat dibutuhkan untuk mendukung keberlanjutan penelitian. Ilustrasi penanganan *missing value* disajikan pada Tabel III.

TABEL III
ILUSTRASI PENANGANAN MISSING VALUE

Periode (t)	Data	Hasil Penanganan Missing Value
t_1	100	100
t_2	300	300
t_3	missing	400
t_4	500	500
t_5	700	700

Berdasarkan ilustrasi Tabel III diatas, Dilakukan penanganan *missing value* 2 untuk memperoleh t_3 , yakni dengan menghitung rata-rata dari t_1, t_2, t_4 dan t_5 sehingga diperoleh nilai imputasi untuk *missing value* tersebut adalah 400. Setelah diperoleh data yang lengkap, maka dilakukan pembagian proporsi data *training* dan data *testing*, yaitu 80%:20% sesuai urutan waktu. Data *training* merupakan 80% dimulai dari awal periode sesuai urutan data, yaitu periode 2 Januari 2018 – 18 Mei 2022, sedangkan data *testing* merupakan 20% dimulai dari sisa keseluruhan data, yaitu periode 19 Mei 2022 – 27 Juni 2023.

Pada data *training* juga dilakukan normalisasi data untuk seluruh variabel menggunakan metode *Min-Max Normalization*. Pada penelitian ini, rentang data disesuaikan seperti yang dibutuhkan untuk pemodelan, yaitu rentang [-1,1]. Berikut merupakan perhitungan normalisasi data untuk Data IHSG periode 1 Februari 2018 (x_1).

$$x_1 = 6339,2378$$

$$New_{min} = -1$$

$$New_{max} = 1$$

$$x_{min} = 3937,632$$

$$x_{max} = 7318,016$$

$$x'_i = New_{min} + (New_{max} - New_{min}) \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

$$x'_1 = -1 + (1 - (-1)) \frac{6339,2378 - 3937,632}{7318,016 - 3937,632}$$

$$x'_1 = 0,420907027$$

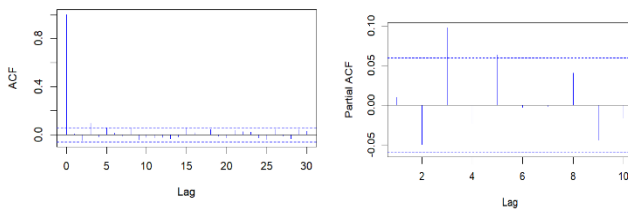
Dengan cara yang sama dilakukan untuk periode lainnya sehingga diperoleh hasil normalisasi data IHSG seperti pada Tabel IV.

TABEL IV
HASIL NORMALISASI DATA IHSG

Periode	Data	Hasil Normalisasi
1/2/2018	6339,2378	0,420907027
⋮	⋮	⋮
6/27/2023	6661,8789	0,611797239

B. Pemilihan Input

Peramalan membutuhkan input-input data variabel independen serta input berdasarkan lag-lag variabel dependen yang signifikan. Pemilihan lag yang dijadikan sebagai input dilakukan berdasarkan plot Partial Autocorrelation Function (PACF), untuk data yang telah stasioner dalam mean. Apabila pada lag tertentu, terdapat lag yang melebihi batas signifikansi, maka lag tersebut dijadikan sebagai input. Stasioneritas dideteksi menggunakan plot ACF, data hasil normalisasi untuk variabel IHSG tidak stasioner dalam mean, sehingga perlu dilakukan proses *differencing* orde pertama dan diperoleh hasil *plot ACF* dan *plot PACF* disajikan pada Gambar 5.



Gambar 5. Plot ACF dan PACF differencing orde 1

Berdasarkan *plot ACF* pada Gambar 5 di atas, nilai lag secara cepat menuju nilai nol sehingga dapat disimpulkan data IHSG telah memenuhi syarat stasioneritas. Berdasarkan Gambar 5 di atas, diperoleh bahwa nilai lag PACF yang melewati batas signifikansi adalah *lag 3* dan *lag 5*. Dengan demikian, kombinasi input yang digunakan adalah *lag 3*, *lag 5*, serta *lag 3* dan *5*. Namun, dikarenakan data IHSG merupakan data *differencing*, maka input variabel yang digunakan dapat disajikan pada Tabel V.

TABEL V
KOMBINASI INPUT VARIABEL

Lag PACF Signifikan	Input Variabel
3,5	$Z_{t-1}, Z_{t-3}, Z_{t-4}, Z_{t-5}, Z_{t-6}$
5	$Z_{t-1}, Z_{t-5}, Z_{t-6}$
3	$Z_{t-1}, Z_{t-3}, Z_{t-4}$

C. Model LSTM dan Bi-LSTM

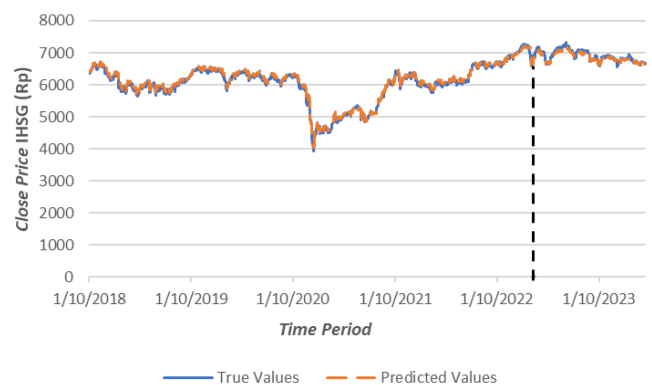
Pemodelan dilakukan terhadap seluruh variabel indeks global untuk masing-masing kombinasi *input* yang telah diperoleh seperti pada Tabel V menggunakan metode LSTM dan *Bi-LSTM*. Pada model dalam penelitian ini. Jumlah *neuron* pada *hidden layer* akan diseleksi berdasarkan hasil pengujian terbaik dari 15 kali Selain itu, penelitian ini akan menggunakan *RMSprop Optimization* untuk memperbarui bobot dalam pemodelan. Iterasi atau *epoch* yang dilakukan adalah sebanyak 100 kali dengan nilai *dropout rate* sebesar

0,5. Setiap pemodelan akan menghasilkan satu model *training* yang diukur kinerjanya terhadap data *testing*. Dilakukan pemodelan menggunakan metode LSTM dan didapatkan MAPE *testing* terbaik untuk setiap variasi jumlah *hidden neuron* yang disajikan pada Tabel VI.

TABEL VI
NILAI MAPE TERBAIK MODEL LSTM SELURUH VARIABEL

Variabel	Kombinasi input lag IHSG	Jumlah Hidden Neuron	MAPE (%)
IHSG – IXIC	1, 3, 4, 5, 6	10	0,81771
IHSG – FTSE	1, 5, 6	10	0,74326
IHSG – SSEC	1, 3, 4, 5, 6	9	0,82538
IHSG – IXIC, FTSE, SSEC	1, 3, 4, 5, 6	9	0,86416

Pemilihan arsitektur LSTM dilakukan berdasarkan nilai MAPE terkecil untuk data *testing*. Berdasarkan Tabel VI, diketahui bahwa model terbaik IHSG dengan input FTSE dan lag 1, 5, 6 memiliki nilai MAPE terkecil di antara model terbaik variabel lainnya, yaitu 0,74326%. Dengan demikian, dapat disimpulkan bahwa Model LSTM dengan evaluasi kinerja model terbaik berasal dari model terbaik IHSG – FTSE dengan kombinasi *input* berasal dari lag 5 dan jumlah *hidden neuron* 10. Kombinasi *input neuron* terdiri dari FTSE All Share, Z_{t-1}, Z_{t-5} , dan Z_{t-6} . Hasil Prediksi IHSG berdasarkan model terbaik dari metode LSTM disajikan melalui Gambar 6.



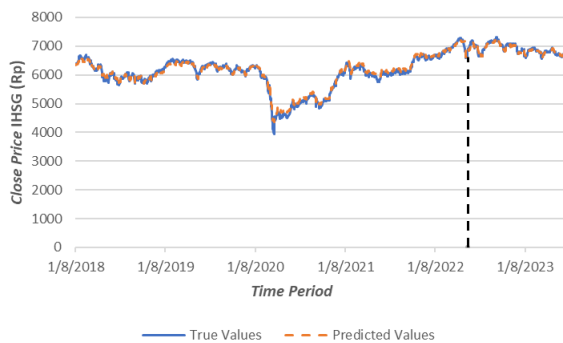
Gambar 6. Hasil Prediksi Model Terbaik LSTM

Berdasarkan Gambar 6 di atas, hasil dari prediksi yang telah dilakukan memberikan gambaran bahwa prediksi harga penutupan IHSG telah sesuai dan menyerupai harga aslinya dengan garis putus-putus sebagai pemisah atau pembagi antara data *training* dan data *testing*. Dilakukan pemodelan menggunakan metode *Bi-LSTM* dan didapatkan MAPE *testing* terbaik untuk setiap variasi jumlah *hidden neuron* yang disajikan pada Tabel VII.

TABEL VII
NILAI MAPE TERBAIK MODEL BI-LSTM SELURUH VARIABEL

Variabel	Kombinasi input lag IHSB	Jumlah Hidden Neuron	MAPE (%)
IHSB – IXIC	1, 3, 4	9	0,584451
IHSB – FTSE	1, 3, 4	9	0,580571
IHSB – SSEC	1, 3, 4	10	0,628203
IHSB – IXIC, FTSE, SSEC	1, 3, 4	9	0,572314

Berdasarkan Tabel VII di atas, diketahui bahwa model terbaik IHSB – IXIC, FTSE, SSEC memiliki nilai MAPE terkecil di antara model terbaik variabel lainnya, yaitu 0,572314%. Dengan demikian, dapat disimpulkan bahwa Model *Bi-LSTM* dengan evaluasi kinerja model terbaik berasal dari model terbaik IHSB – IXIC, FTSE, SSEC dengan kombinasi *input input neuron* terdiri dari IXIC, FTSE All Share, SSEC, Z_{t-1} , Z_{t-3} , dan Z_{t-4} . dan jumlah *hidden neuron* 9. Hasil Prediksi IHSB berdasarkan model terbaik dari metode *Bi-LSTM* disajikan melalui Gambar 7.



Gambar 7. Hasil Prediksi Model Terbaik *Bi-LSTM*

Berdasarkan Gambar 7, hasil dari prediksi yang telah dilakukan memberikan gambaran bahwa prediksi harga penutupan IHSB telah sesuai dan menyerupai harga aslinya dengan garis putus-putus sebagai pemisah atau pembagi antara data *training* dan data *testing*.

D. Perbandingan Model LSTM dan *Bi-LSTM*

Setelah diperoleh model terbaik dari masing-masing metode, dilakukan perbandingan untuk menyeleksi metode yang terbaik untuk peramalan IHSB 10 periode kedepan. Pemilihan metode terbaik dilakukan berdasarkan MAPE terkecil *training* dan *testing* antara model terbaik LSTM dan model terbaik dari *Bi-LSTM*. Berdasarkan hasil pemodelan LSTM dan *bi-LSTM* pada section sebelumnya, diperoleh bahwa arsitektur model terbaik LSTM memiliki 4 input (FTSE, lag 1 IHSB, lag 5 IHSB, dan lag 6 IHSB), dengan banyaknya *neuron* di *hidden layer* adalah 10, dan 1 *output layer*, sehingga dapat dituliskan sebagai LSTM (4-10-1). Pada model *bi-LSTM*, arsitektur terbaiknya memiliki 6 *input* (IXIC, FTSE, SSEC, lag 1 IHSB, lag 3 IHSB, dan lag 4 IHSB), dengan banyaknya *neuron* di *hidden layer* adalah 9, dan 1 *output layer*, sehingga dapat dituliskan sebagai *bi-LSTM* (6-9-1) Perbandingan nilai MAPE masing-masing metode disajikan pada Tabel VIII.

TABEL VIII
PERBANDINGAN MAPE TERBAIK LSTM DAN *Bi-LSTM*

	LSTM (4-10-1)	<i>Bi-LSTM</i> (6-9-1)
<i>Training MAPE</i> (%)	0,99380	1,17390
<i>Testing MAPE</i> (%)	0,74326	0,57231

Berdasarkan Tabel VIII, diperoleh bahwa metode *Bi-LSTM* memiliki model terbaik dengan MAPE yang lebih kecil dibandingkan dengan model LSTM. Dengan demikian, peramalan IHSB 10 periode berikutnya akan dilakukan menggunakan metode *Bi-LSTM* (6-9-1). Namun, LSTM memiliki model *training* yang lebih baik dibandingkan dengan *Bi-LSTM*. Hal ini kemungkinan disebabkan oleh adanya *overfitting* yang terjadi pada model LSTM. Selanjutnya akan dilakukan peramalan IHSB 10 periode berikutnya dengan metode *Bi-LSTM* dimana variabel input adalah indeks global IXIC, FTSE All Share, dan SSEC serta input dari lag IHSB yang signifikan.

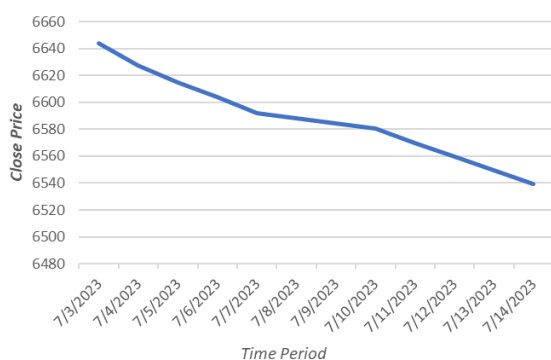
E. Peramalan IHSB Hasil Metode Terbaik

Berdasarkan perbandingan metode terbaik yang telah dilakukan sebelumnya, diperoleh bahwa metode *Bi-LSTM* yang terdiri atas 6 *input neuron*, 9 *hidden neuron*, dan 1 *output layer* merupakan metode dan model terbaik. Setelah memperoleh metode terbaik, dilakukan peramalan IHSB menggunakan model terbaik dari metode *Bi-LSTM* untuk mengetahui nilai IHSB 10 periode berikutnya terhitung mulai tanggal 3 Juli 2023 – 14 Juli 2023. Input variabel yang dibutuhkan, yaitu IXIC, FTSE All Share, SSEC, Z_{t-1} , Z_{t-3} , dan Z_{t-4} . Peramalan IHSB juga akan melibatkan peramalan pada variabel-variabel indeks global yang terpilih sehingga perlu dilakukan peramalan pada variabel IXIC, FTSE All Share, dan SSEC 10 periode berikutnya. Peramalan variabel indeks global ini akan melalui proses yang sama seperti pada variabel IHSB, mulai dari identifikasi stationeritas dan PACF yang signifikan hingga peramalan berdasarkan model yang diperoleh. Model terbaik *univariate* variabel indeks global disajikan pada Tabel IX.

TABEL IX
MODEL TERBAIK PERAMALAN UNIVARIATE VARIABEL INDEKS GLOBAL

Variabel	Input Lag Signifikan	Jumlah Hidden Neuron	MAPE (%)
IXIC	X_{1t-1}	10	1,380479
	X_{1t-2}		
	X_{1t-6}		
	X_{1t-7}		
	X_{1t-8}		
	X_{1t-9}		
FTSE	X_{2t-1}	9	0,671195
	X_{2t-6}		
	X_{2t-7}		
	X_{2t-8}		
SSEC	X_{3t-1}	10	0,646514
	X_{3t-4}		
	X_{3t-5}		

Model *univariate* pada Tabel IX akan digunakan untuk memperoleh peramalan nilai harga penutupan harian masing-masing indeks global 10 periode berikutnya. Peramalan dilakukan satu per satu mengingat variabel *input* berasal dari lag signifikan indeks itu sendiri. Hal yang sama juga dilakukan pada peramalan IHSG karena untuk melakukan peramalan selanjutnya dibutuhkan pembaruan atau penambahan data hasil peramalan ke dalam dataset sehingga program dapat menghitung dengan benar dan akurat. Sebagai contoh, peramalan nilai SSEC pada $t = 13$ akan membutuhkan nilai IXIC pada $t = 12$, $t = 11$, dan sebagainya. Hal tersebut juga sama untuk peramalan nilai IHSG pada tanggal 3 Juli 2023 akan membutuhkan nilai indeks global IXIC, FTSE, dan SSEC tanggal 3 Juli 2023, IHSG tanggal 27 Juni 2023, 25 Juni 2023, 24 Juni 2023, 23 Juni 2023, dan 22 Juni 2023. Hasil peramalan *univariate* ini akan berfungsi sebagai input untuk peramalan IHSG. Hasil peramalan IHSG 10 periode berikutnya dapat divisualisasikan pada Gambar 8.



Gambar 8. Hasil Prediksi Model Terbaik *Bi-LSTM*

Berdasarkan visualisasi Gambar 8, dapat disimpulkan bahwa nilai IHSG 10 periode berikutnya diperkirakan akan cenderung mengalami penurunan walaupun tidak signifikan.

IV. KESIMPULAN

Pada data IHSG terdapat beberapa *missing value* sehingga dilakukan imputasi berdasarkan *central moving average*. Penelitian ini membandingkan penggunaan metode LSTM dan *bi-LSTM* dalam meramalkan IHSG. Model terbaik dalam peramalan Indeks Harga Saham Gabungan berdasarkan

indeks global adalah model terbaik *Bi-LSTM* (6-9-1), dengan 6 *input* (IXIC, FTSE All Share, SSEC, lag 1 IHSG, lag 3 IHSG, dan lag 4), 9 *neuron* di *hidden layer*, dan 1 *output*. Nilai MAPE yang diperoleh sebesar 0.572314%. Nilai MAPE yang sangat kecil menunjukkan bahwa model *bi-LSTM* dapat meramalkan data IHSG dengan sangat baik. Hasil peramalan berdasarkan model dari metode terbaik menunjukkan bahwa IHSG 10 periode berikutnya cenderung mengalami penurunan. Hal ini dapat menjadi pertimbangan investor karena berdampak pada risiko finansial.

DAFTAR PUSTAKA

- [1] D. Bery, "Pengaruh Indeks Harga Saham Global Terhadap Indeks Harga Saham Gabungan (IHSG)," Universitas Brawijaya, 2018.
- [2] S. D. Prahesti, "Pengaruh Indeks SSEC, N225, STI, dan Faktor Makroekonomi terhadap IHSG," *J. Ilmu Manaj. Vol.*, vol. 8, no. 3, pp. 878–893, 2020.
- [3] M. Jonathan, "Pengaruh Pergerakan Indeks Global Terhadap Pergerakan IHSG Periode 2015-2021," Universitas Atma Jaya Yogyakarta, 2022.
- [4] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," 2019.
- [5] K. Moharm, M. Eltahan, and E. Elsaadany, "Wind speed forecast using LSTM and Bi-LSTM algorithms over gabal el-zayt wind farm," *Proc. - 2020 Int. Conf. Smart Grids Energy Syst. SGENS 2020*, pp. 922–927, 2020, doi: 10.1109/SGENS51519.2020.00169.
- [6] A. Nilsen, "Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45.," *J. Stat. dan Apl.*, vol. 6, no. 1, pp. 137–147, 2022.
- [7] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, pp. 1–23, 2018, doi: 10.3390/en11051255.
- [8] S. Gangwar, V. Bali, and A. Kumar, "o n Scalable Information Systems EAI Endorsed Transactions Comparative Analysis of Wind Speed Forecasting Using LSTM and SVM," vol. 7, no. 25, pp. 1–9, 2020.
- [9] S. T. Lammoreno, "Peramalan IHSG Berdasarkan Faktor Indeks Global Menggunakan Metode Long Short-Term Memory dan Gated Recurrent Unit," Institut Teknologi Sepuluh Nopember, 2023.
- [10] S. Zaheer *et al.*, "A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model," *Mathematics*, vol. 11, no. 3, pp. 1–24, 2023, doi: 10.3390/math11030590.
- [11] C. Choi, "Time Series Forecasting with Recurrent Neural Networks in Presence of Missing Data," UiT The Arctic University of Norway, 2018.
- [12] D. I. Puteri, "Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah," *EULER J. Ilm. Mat. Sains dan Teknol.*, vol. 11, no. 1, pp. 35–43, 2023.