



PREDIKSI INDEKS HARGA SAHAM GABUNGAN (IHSG) MENGUNAKAN LONG SHORT-TERM MEMORY

Akhmad Yusuf

*Program studi Matematika FMIPA Universitas Lambung Mangkurat
Banjarbaru Kalimantan Selatan*

Email : yusuf@ulm.ac.id

ABSTRACT

The capital market is a meeting place for sellers and buyers and can be used as an indicator of a country's progress. The higher the movement or activity in the capital market, the higher the economic movement in a country. The Indonesia Stock Exchange (IDX) is a capital market located in Indonesia. The Composite Stock Price Index (CSPI) is the average overall stock price on the Indonesia Stock Exchange (IDX) and has high volatility, so a method is needed to predict the IHSG price movement which can be used as a reference for buyers (the surplus funds). Long Short-Term Memory (LSTM) is a forecasting method that can be used to predict time series data. In this study, the data used amounted to 1212 data at the time interval from February 16, 2017 to February 14, 2022 with a 1D time frame. The data is divided into 2 parts, namely training 970 data and testing 242 data. The LSTM parameter used is batch-size 25 and to test the success of these parameters, different epochs are used. Epoch of 50 is the best model to produce a smaller RMSE of 6.2335 with a predicted value of 6765.5103 and an actual value of 6807.50.

Keyword : *capital market ,Indonesia stock exchange, long short-term memory*

ABSTRAK

Pasar modal merupakan tempat bertemunya pihak penjual dan pembeli serta dapat dijadikan sebagai indikator kemajuan suatu negara. Semakin tinggi pergerakan atau aktifitas di pasar modal maka semakin tinggi pula pergerakan ekonomi di suatu Negara tersebut. Pasar modal yang terdapat di Indonesia yaitu Bursa Efek Indonesia (BEI). Index harga saham gabungan (IHSG) merupakan rata-rata harga saham keseluruhan yang berada di BEI dan memiliki volatilitas yang tinggi sehingga diperlukan suatu metode untuk memprediksi pergerakan harga IHSG tersebut yang dapat dijadikan sebagai acuan bagi para pembeli (pihak surplus dana). *Long Short-Term Memory* (LSTM) merupakan sebuah metode *forecasting* yang dapat digunakan untuk memprediksi data yang bersifat *time series*. Pada penelitian ini, data yang digunakan berjumlah 1212 data pada interval waktu 16 Februari 2017 sampai 14 Februari 2022 dengan *time frame* 1D. Data dibagi menjadi 2 bagian, yaitu data *training* 970 data dan data *testing* 242 data. Parameter LSTM yang digunakan *batch-size* 25 dan untuk menguji keberhasilan parameter tersebut digunakan epoch yang berbeda-beda. Epoch sejumlah 50 merupakan model terbaik menghasilkan RMSE lebih kecil yaitu 6.2335 dengan nilai prediksi 6765.5103 dan nilai aktual 6807.50.

Kata kunci: *Pasar modal, Bursa Efek Indonesia, long short-term memory*

1. LATARBELAKANG

Semakin berkembangnya teknologi dan didukung dengan perkembangan revolusi industry 4.0 menjadikan momentum yang tepat dalam perkembangan investasi digital (*digital investment*) terutama di instrument keuangan. Para pihak yang memiliki permodalan yang cukup dapat melakukan diversifikasi investasi di *financial asset*. Investasi memiliki tujuan agar memperoleh peningkatan nilai investasi atau peningkatan aset investasi selama periode tertentu dari sejumlah nilai investasi yang ditempatkan. (Husnan, 1994).

Pada era globalisasi saat ini, salah satu tanda perkembangan ekonomi dapat dilihat dari pasar modal yang dimiliki oleh negara tersebut. Fasilitas penyaluran modal dari unit yang memiliki surplus dana ke unit defisit dana dapat menggunakan pasar modal. Untuk melihat perkembangan suatu negara dapat dilihat dari perkembangan pasar modal yang dimiliki oleh negara tersebut, misalnya di Negara Jepang, Inggris, dan Amerika menjadikan pasar modal sebagai ukuran perkembangan suatu negara. Semakin baik perkembangan pasar modal di negara tersebut, maka dapat dipastikan perekonomian negara tersebut baik. (Utomo, 2010). Adapun yang diperjualbelikan di pasar modal berupa obligasi, sukuk, reksadana, atau saham.

Beberapa perusahaan yang ingin mengembangkan perusahaan menerbitkan *Initial Public Offering* (IPO), dengan melakukan IPO perusahaan menerbitkan saham yang dapat dimiliki oleh setiap orang. Saham merupakan surat berharga yang berbentuk selembar kertas yang menyatakan bahwa kepemilikan atas perusahaan yang dibeli dari pemilik perusahaan yang menerbitkan surat tersebut (Darmadji & Fakhruddin, 2012).

Di Indonesia untuk melihat pergerakan harga saham rata-rata di dalam negeri dapat dilihat dari Indeks Harga Saham Gabungan (IHSG) (Anoraga, 2003). IHSG adalah *average index* gabungan seluruh saham yang diperjualbelikan dan jenis index di Bursa Efek Indonesia (BEI). IHSG digunakan untuk melihat kinerja total saham yang berada di Bursa Efek dengan menjumlahkan seluruh harga saham dan di bagi dengan total jumlah saham tersebut yang tercatat di Bursa Efek Indonesia.

Menurut (Tandelin, 2010) Indeks Harga Saham Gabungan adalah gabungan dari beberapa saham yang ada di Indonesia dan tercatat di bursa efek Indonesia yang dihitung nilai rata-rata dari beberapa saham tersebut. Informasi yang diberikan IHSG merupakan data historis dari pergerakan rata-rata saham tersebut dan dapat dijadikan sebagai kinerja atau performance dari saham-saham.

Dari IHSG dapat dilihat situasi pasar modal yang terdapat di Negara karena IHSG menggambarkan gabungan dari seluruh harga-harga saham yang berada pada pasar modal. IHSG memiliki volatilitas yang tinggi sehingga diperlukan suatu metode untuk melakukan forecasting agar dapat membantu investor dalam mengambil keputusan.

Long Short Term Memory Neural Network (LSTM) merupakan salah satu jenis Jaringan Syaraf Berulang/*Recurrent Neural Network* (RNN). LSTM memiliki kemampuan untuk menyimpan pola data yang digunakan untuk melakukan prediksi. LSTM memiliki sensitifitas untuk menyeleksi dan mempelajari mengenai pola data sehingga dapat menyimpulkan data yang akan di eliminasi dan data mana saja yang akan dipertahankan karena pada masing-masing neuron di LSTM untuk mengatur memori pada masing-masing neuron menggunakan beberapa gates. Penggunaan metode LSTM dapat digunakan di berbagai bidang terapan seperti Estimasi, *Forecasting*, *Classification*, dan *Clustering*. Penggunaan parameter-parameter LSTM yaitu maksimum epoch, jumlah layer tersembunyi, jumlah pola deret waktu, dan komposisi data dalam memprediksi nilai Cryptocurrency dalam hal ini Bitcoin didapatkan hasil terbaik dengan parameter komposisi data training 70% dan data testing 30% dengan 25 lapisan layer tersembunyi, maksimum epoch adalah 100 memiliki akurasi rata-rata sebesar 95.36% untuk data training dan 93.5% untuk data testing (Aldi, 2018). LSTM memiliki akurasi yang baik dalam melakukan estimation dan forecasting untuk data deret waktu (A. Yadav, C. K. Jha, 2020). Penelitian ini bertujuan melakukan prediksi harga pada Index Harga Saham Gabungan (IHSG) menggunakan metode *Long Short-Term Memory* (LSTM).

2. METODOLOGI PENELITIAN

A. Collecting data

Pada penelitian ini menggunakan data IHSG yang bersumber dari yahoo finance, periode pengambilan data 16 Februari 2017 sampai 14 Februari 2022 dengan time frame 1D dan jumlah data sebanyak 1212 data. Berikut merupakan tabel data

Tabel 1. Data IHSG

	Date	Open	High	Low	Close	Adj Close	Volume
0	2/16/2017	5389.068848	5395.278809	5357.711914	5377.998047	5377.998047	218393700
1	2/17/2017	5366.854004	5370.331055	5340.857910	5350.932129	5350.932129	179271900
2	2/20/2017	5358.551758	5373.091797	5351.018066	5359.288086	5359.288086	107921200
3	2/21/2017	5363.613770	5375.695801	5336.834961	5340.990234	5340.990234	209661500
4	2/22/2017	5353.106934	5360.414063	5331.431152	5358.683105	5358.683105	189512900
...
1207	2/8/2022	6813.184082	6860.750000	6780.416992	6789.521973	6789.521973	255862800
1208	2/9/2022	6811.753906	6840.394043	6796.632813	6834.605957	6834.605957	253458500
1209	2/10/2022	6859.333984	6874.351074	6814.210938	6823.642090	6823.642090	279507200
1210	2/11/2022	6813.607910	6825.062012	6773.537109	6815.606934	6815.606934	223574800
1211	2/14/2022	6791.315918	6796.711914	6698.508789	6734.485840	6734.485840	202907800

1212 rows × 7 columns

B. Praprosesing Data

Data yang digunakan dibagi menjadi 2 bagian yaitu data training sejumlah 80% dan data testing sejumlah 20%. Data yang diperoleh dipersiapkan terlebih dahulu untuk memastikan apakah data telah siap atau tidak. Jika terdapat *missing value* pada data maka dapat dilakukan pengisian data dengan menggunakan interpolasi. *Interpolation* adalah metode untuk mengisi nilai yang kosong (*missing value*) diantara 2 nilai data atau lebih dengan tujuan mendapatkan nilai yang tepat. *Interpolation* dapat digunakan untuk mengisi *missing value* pada data *time series*. Persamaan *interpolation* diuraikan pada persamaan berikut :

$$y = y_1 + (x - x_1) \frac{(y_2 - y_1)}{(x_2 - x_1)} \quad (1)$$

Keterangan :

y = hasil nilai interpolasi

x = variable bebas

x_1, y_1 = nilai pada titik tertentu

x_2, y_2 = nilai pada titik lainnya

Setelah praprosesing data selesai dilakukan, selanjutnya data dilakukan normalisasi pada interval (0,1) dengan menggunakan metode Minmaxscaler. Rumus normalisasi data dapat dilihat pada persamaan berikut :

$$y = \frac{(x - x_{min})}{x_{max} - x_{min}} \quad (2)$$

Keterangan :

y = nilai hasil normalisasi

x = nilai data yang di normalisasi

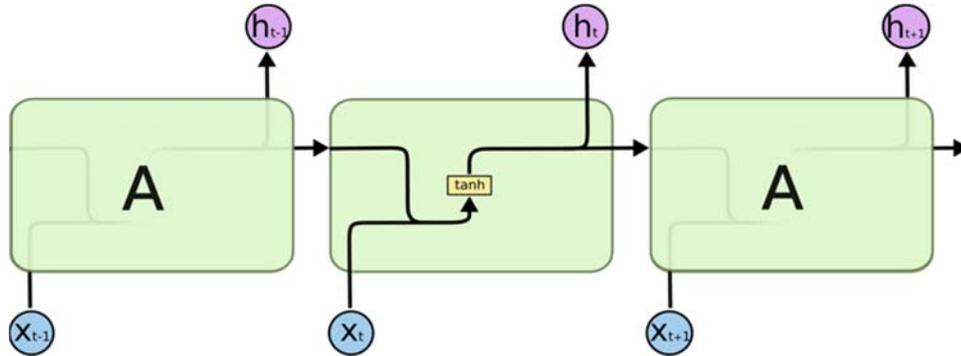
x_{min} = nilai minimum data

x_{max} = nilai maksimum data

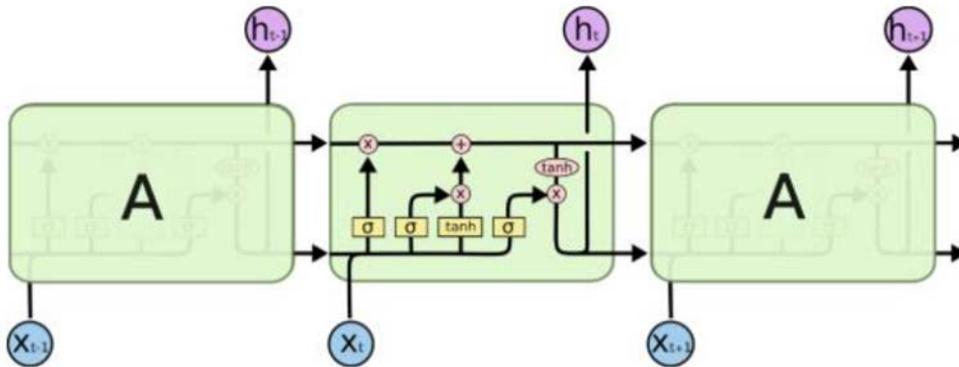
C. Metode Prediksi/*Forecasting*

Proses *forecasting* dilakukan dengan menggunakan aplikasi python minianaconda dan terminal Jupyter Notebook dengan metode *Long Short-Term Memory* (LSTM). *Recurrent Neural Network* (RNN) merupakan Jaringan Syaraf (NN) berulang. Kelemahan pada RNN mengenai ketergantungan jangka panjang, hal ini dapat diselesaikan dengan metode LSTM. RNN pada perulangannya hanya memiliki satu layer tunggal (single layer), yaitu layer tanh yang dapat dilihat pada Gambar 1. Pada prosesnya diperlukan konversi nilai berada pada interval (-1,1) dan dilakukan pada layer tanh. x_{t-1} adalah *input* sebelumnya, h_{t-1} adalah *output* sebelumnya yang akan dimasukkan sebagai *input* bersamaan dengan *input* baru. h_{t+1} adalah *output* setelah orde t dan x_{t+1} adalah *input* setelah orde t. *LSTM* memiliki empat *layer*, yaitu gerbang lupa (1), gerbang masuk (2),

new cell state candidate (3), dan gerbang keluar (4) pada perulangan modelnya seperti pada Gambar 2



Gambar 1. Model Recurrent Neural Network (RNN)



Gambar 2. Model LSTM dengan 4 layer (Olah, 2015)

Menurut (Hochreiter, S., & Schmidhuber, 1997) persamaan metode LSTM dapat diuraikan pada persamaan berikut :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\bar{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \quad (6)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

dengan

- f_t : gate lupa
- σ : sigmoid function
- W_f : nilai bobot untuk *gate* lupa
- h_{t-1} : nilai keluaran sebelum orde ke-t
- x_t : nilai masukan pada orde ke-t
- b_f : nilai bias pada *gate* lupa
- i_t : masukan gate
- W_i : nilai bobot untuk *gate* masukan
- b_i : nilai bias pada *gate* masukan

- \bar{C}_t : nilai baru yang dapat ditambahkan ke *cell state*
- \tanh : *tanh function*
- W_c : nilai bobot untuk *cell state*
- C_{t-1} : *Cell state* sebelum orde ke-t
- o_t : output gate
- σ : *sigmoid function*
- W_o : nilai bobot untuk *gate* keluaran
- b_o : nilai bias pada *gate* keluaran
- h_t : nilai keluaran orde ke-t
- C_t : *Cell state*

D. Penentuan *error* dan akurasi prediksi

Pada penelitian ini untuk melihat *error* prediksi dari metode LSTM digunakan Root Mean Square Error (RMSE). RMSE adalah penentuan error dengan menghitung *standar deviation* (std) dari titik-titik data prediksi yang menjauhi garis regresi (data aktual). Persamaan RMSE menurut (Barnston, 1992) diuraikan pada persamaan 9.

$$RMSE_{f_0} = \left[\sum_{i=1}^N \frac{(z_{f_i} - z_{o_i})^2}{N} \right]^{1/2} \quad (9)$$

Dimana:

- z_{f_i} : nilai prediksi
- z_{o_i} : nilai aktual
- N : jumlah pengamatan pada data

3. HASIL DAN PEMBAHASAN

Hasil dan pembahasan adalah sebagai berikut :

- a. Visualisasi data awal untuk melihat pola pergerakan harga IHSG dapat ditampilkan dalam bentuk grafik pada Gambar 3



Gambar 3. Grafik pergerakan harga IHSG

- b. Normalisasi Data

Berdasarkan pada persamaan (2) maka hasil normalisasi data harga penutupan (close) IHSG adalah sebagai berikut :

```
array([[0.49719674],
       [0.48785392],
       [0.49073829],
       ...,
       [0.99621541],
       [0.99344177],
       [0.96543976]])
```

c. *Long Short-Term Memory* (LSTM)

Tabel 2 merupakan pemberian nilai dasar untuk melakukan perhitungan langkah-langkah dari *Long Short-Term Memory*

Tabel 2. Nilai awal untuk contoh perhitungan

Nilai IHSG	0,4972
Weight	0,0287
Bias Forget (b_f)	1
Bias Input (b_i)	0,5
Bias C_t	0
Bias Output (b_o)	0,1

Nilai *Weight* $W = \left(-\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\right)$, $W = \left(-\frac{1}{\sqrt{1212}}, \frac{1}{\sqrt{1212}}\right)$ akan dinisialisasikan dengan jangkauan dari (-0.0287, 0.0287)

Dimana :

W : weight (bobot)

N : jumlah data

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ &= \sigma([0,0287 \times 0,4972 + 0,0287 \times 0,4972] + 1) \\ &= \sigma(1.028539) \\ &= 0.7366 \end{aligned}$$

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ &= \sigma([0,0287 \times 0,4972 + 0,0287 \times 0,4972] + 0,5) \\ &= \sigma(0.528539) \\ &= 0.6291 \end{aligned}$$

$$\begin{aligned} \bar{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ &= \tanh([0,0287 \times 0,4972 + 0,0287 \times 0,4972] + 0) \\ &= \tanh(0.028539) \\ &= 0.02853 \end{aligned}$$

$$\begin{aligned} C_t &= f_t \times C_{t-1} + i_t \times \bar{C}_t \\ &= 0.7366 \times 0 + 0.6291 \times 0.02853 \\ &= 0.017948 \end{aligned}$$

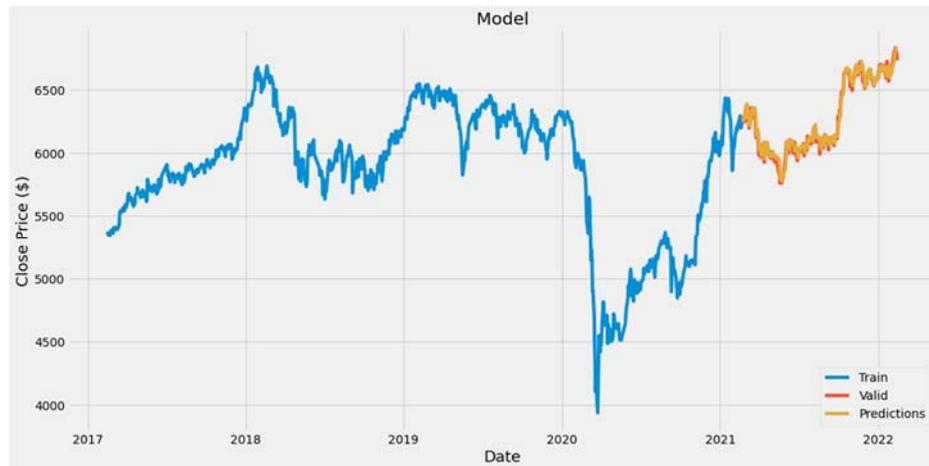
$$\begin{aligned} o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ &= \text{sigmoid}([0,0287 \times 0,4972 + 0,0287 \times 0,4972] + 0.1) \\ &= \text{sigmoid}(0.128539) \\ &= 0.53209 \end{aligned}$$

$$\begin{aligned} h_t &= o_t * \tanh(C_t) \\ &= 0.53209 \times \tanh(0.017948) \\ &= 0.009548 \end{aligned}$$

Hasil yang didapatkan $h_t = 0.009548$ dan $C_t = 0.017948$ dan akan dilanjutkan sebagai input pada proses selanjutnya sampai pada jumlah epoch yang ditentukan.

d. *Performance Long Short-Term Memory (LSTM)*

Uraian pada poin (c) di atas merupakan contoh perhitungan Long short-term memory dan untuk selanjutnya menggunakan komputasi dengan bahasa pemrograman Python. ditampilkan dalam bentuk grafik pada Gambar 4



Gambar 4. Grafik Training dan Testing model LSTM

Untuk melihat performance secara visual dapat dilihat pada Gambar 4 di atas. Pada Gambar 4 diperlihatkan performance model LSTM untuk data training berjumlah 80% yang diperlihatkan pada grafik warna biru dan performance testing sejumlah 20% dapat dilihat pada grafik berwarna merah. Sedangkan performance untuk grafik prediksi dapat dilihat pada grafik berwarna orange. Secara visual grafik berwarna orange dan merah saling berhimpitan dan mengindikasikan memiliki error yang relatif kecil.

Tabel 3. Hasil model LSTM dengan variasi epoch

Epoch	Batch-Size	RMSE	Harga ke-t	Predict (t+1)	Actual (t+1)
1	25	99.13156	6734.485840	6494.7056	6807.50
25	25	6.415767	6734.485840	6837.057	6807.50
50	25	6.2335	6734.485840	6765.5103	6807.50
75	25	6.581738	6734.485840	6725.724	6807.50
100	25	44.531806	6734.485840	6656.799	6807.50

Pada Tabel 3 dapat dilihat bahwa pada saat epoch 1 dengan batch-size 25 menghasilkan nilai RMSE 99.13156 dan nilai prediksi 6494.7056, Epoch 25 dengan batch-size 25 menghasilkan nilai RMSE 6.415767 dan nilai prediksi 6837.057, Epoch 50 dengan batch-size 25 menghasilkan nilai RMSE 6.2335 dan nilai prediksi 6765.5103, Epoch 75 dengan batch-size 25 menghasilkan nilai RMSE 6.581738 dan nilai prediksi 6725.724, Epoch 100 dengan batch-size 25 menghasilkan nilai RMSE 44.531806 dan nilai prediksi 6656.799.

4. KESIMPULAN

Hasil dari penelitian ini menggunakan data pada periode 16 Februari 2017 sampai 14 Februari 2022 memberikan informasi bahwasanya Long short-term memory dapat digunakan untuk melakukan prediksi terhadap pergerakan nilai IHSG dengan performance yang diperoleh dapat dilihat dari nilai RMSE yaitu sebesar 6.2335 pada epoch 50 dengan nilai prediksi 6765.5103 dan nilai actual sebesar 6807.50

REFERENSI

- A. Yadav, C. K. Jha, dan A. S. (2020). Optimizing LSTM for prediction in Indian Stock market. *Procedia Comput. Sci*, 2091–2100.
- Aldi, M. W. P. A. (2018). Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin. *E-Proceeding of Engineering*, 5.
- Anoraga, P. (2003). *Pengantar Pasar Modal*. Rineka Cipta.
- Barnston, A. G. (1992). Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score. *Amet Soc*, 7(4), 699–709. [https://doi.org/https://doi.org/10.1175/1520-0434\(1992\)007%3C0699:CATCRA%3E2.0.CO;2](https://doi.org/https://doi.org/10.1175/1520-0434(1992)007%3C0699:CATCRA%3E2.0.CO;2)
- Darmadji, T., & Fakhruddin, H. M. (2012). Pasar Modal Di Indonesia. In *Salemba Empat*. Salemba Empat.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9, 1735–1780. <https://doi.org/https://doi.org/10.1162/neco.1997.9.8.1735>
- Husnan, S. (1994). *Dasar-dasar Teori dan Analisis Sekuritas*. YKPN.
- Olah, C. (2015). *Understanding LSTM Networks*.
- Tandelin, E. (2010). *Portofolio dan Investasi*. Kanisius.
- Utomo, P. (2010). *Peluang dan Tantangan Pertumbuhan Reksadana*. PT. Minna Padi Aset Manajemen.