

## KEMAMPUAN LONG SHORT TERM MEMORY MACHINE LEARNING DALAM PROYEKSI SAHAM BANK BRI TBK

Widi Hastomo<sup>1</sup>, Aminudin<sup>2</sup> dan Adhitho Satyo Bayangkari Karno<sup>3</sup>

<sup>(1)</sup>STMIK Jakarta STI&K

Jl. BRI No.17, Radio Dalam, Kebayoran Baru, Jakarta Selatan 12140

<sup>(2)</sup>ITB Ahmad Dahlan Jakarta

Jl. Ir H. Juanda No.77, Cireundeu, Tangerang Selatan, Banten 15419

<sup>(3)</sup>Universitas Gunadarma

Jl. Margonda Raya No. 100, Depok, Jawa Barat 16424

{widie.has, adh1t10.2}@gmail.com

### ABSTRAK

Penelitian ini bertujuan untuk mengetahui proyeksi akurasi saham Bank BRI Tbk, menggunakan long short term memory [lstm] machine learning dan arima statistik, data yang digunakan sejumlah 3717 data harga penutupan, dalam rentang waktu dari tanggal 18 april 2005 sampai dengan 30 maret 2020 data tersebut dihimpun dari finance.yahoo.com. Hasil terbaik dari penelitian ini yaitu nilai yang dihasilkan lstm dengan root mean square error [rmse]1%, dibandingkan nilai rmse dari arima yaitu 2.46%. Hasil dari proses testing menghasilkan akurasi yang tinggi mencapai 99%. Pada penelitian selanjutnya bisa ditambahkan metode yang lain seperti polynomial, K-nearest neighbors (KNN) atau dapat menambah internal dan eksternal faktor, Isu-isu terkini semisal pandemik covid-19 yang lebih detail.

**Kata Kunci :** Long Short Term Memory Machine Learning

### PENDAHULUAN

Bank BRI termasuk salah satu bank milik pemerintah atau Badan Usaha Milik Negara (BUMN), pertama kali melantai di bursa efek pada tahun 2003 dan memutuskan untuk menjual 30% sahamnya, sehingga menjadi perusahaan publik dengan Nama PT. Bank Republik Indonesia Tbk, yang digunakan sampai saat ini [1]. Kini usianya mencapai 17 tahun melantai di bursa efek, awal initial public offering (IPO) harga saham di kisaran Rp. 875,-. Pada kinerja tahun 2018 Bank BRI Tbk membagikan dividen sebesar Rp.16,7 triliun hal ini menjadikan Bank BRI sebagai penyumbang dividen Bank BUMN terbesar [2]. Disaat pandemik covid-19 seperti sekarang ini saham Bank BRI Tbk mengalami koreksi yang cukup signifikan, dalam rentang waktu januari hingga april tahun 2020 nilai saham paling tinggi yaitu tanggal 24 januari 2020 sebesar Rp. 4.740 per lembar saham, sedangkan pada tanggal 9 April 2020 dikisaran Rp. 2.790. terlihat dalam gambar grafik 1 dibawah ini;



Gambar 1 Chart saham Bank BRI Tbk[3]

Untuk itu diperlukan sebuah alat atau metode untuk memproyeksinya. Peruntukan yaitu; (a). Untuk mendapatkan hasil proyeksi saham yang lebih akurat, (b). Sebagai dasar pengambil keputusan investor.

Dalam beberapa dekade terakhir ini ada terobosan terbaru dalam data sains, ditemukan bahwa dalam memprediksi data time series, pengamatan menggunakan Long Short Term Memory (LSTM)[4] merupakan solusi yang paling efektif. LSTM memiliki kelebihan dibandingkan dengan conventional feed-forward neural networks dan Recurrent Neural Network (RNN) [5].

Hochreiter & Schmidhuber pertama kali memperkenalkan teori LSTM pada tahun 1997. Model mesin learning LSTM dirancang untuk mengatasi kekurangan yang ada di model sebelumnya yaitu RNN

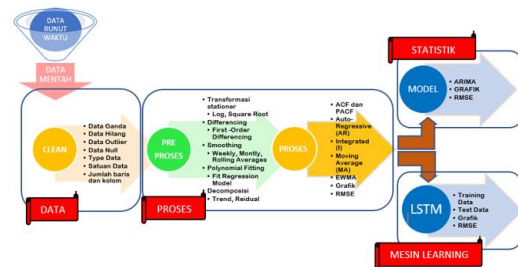
(Recurrent Neural Network). Masalah *vanishing gradient* yang ada di RNN disebabkan karena ketidak mampuan untuk mengatasi informasi yang panjang [6]. Selain menggunakan metode LSTM, juga digunakan metode pembandingan statistik Autoregresif Integrated Moving Average (ARIMA) [7].

Penelitian yang dilakukan [8] prediksi harga saham Standard and Poor's (S&P) 500 menggunakan Support Vector Machine Regressor dan kombinasi CNN-LSTM, menghasilkan peningkatan kinerja mencapai 9% dengan model CNN dan LSTM. [9] berpendapat bahwa komparasi data time series dengan metode Arima, LSTM dan Gated Recurrent Unit (GRU) menghasilkan ARIMA memberikan hasil terbaik masing-masing sebesar 2,76% dan 302,53 untuk MAPE dan RMSE. Namun Gated Recurrent Unit (GRU) berkinerja lebih baik daripada Long Short-term Memory (LSTM), dengan masing-masing 3,97% dan 381,34 MAPE dan RMSE. Penelitian [10] dengan metode Support Vector Machine epoch 10 rerata akurasi 66,9823 dengan standar deviasi 0,05256 dan LSTM epoch 10 rerata akurasi 68,51635 dengan standar deviasi 0,71779. Hasil penelitian [4] dengan metode Arima dan Lstm, yang menghasilkan tingkat akurasi mencapai 99% (RMSE 1%) dengan metode LSTM.

### METODE PENELITIAN

Alur dari penelitian ini yaitu proses cleansing data, melalui data ganda, data hilang, data outlier, data null, type data, satuan data dan jumlah baris dan kolom, selanjutnya melalui pre proses yaitu transformasi stasioner, differencing, smoothing, polynomial fitting, dekomposisi, setelah melalui tahapan proses data kemudian dibandingkan antara hasil RMSE Arima dan LSTM. Data yang digunakan dalam eksperimen ini yaitu data sekunder yang diunduh dari finance.yahoo.com. Rentang waktu yang digunakan dari 18 April 2005 sampai 30 April 2020, sejumlah 3717 baris data. Bahasa pemrograman yang dipergunakan yaitu bahasa pemrograman Python, merupakan bahasa pemrograman open source, menjadikan bahasa ini cepat

berkembang dengan banyak paket dan library yang memudahkan dalam pemrograman ML.



Gambar 2 Skenario Penelitian proyeksi saham dengan LSTM

Model statistik arima dijadikan pembandingan dalam eksperimen ini karena mempunyai sifat tetap dan nilai rerata yang tidak bergeser terhadap waktu. Saat menentukan model yang paling sesuai untuk data runtun waktu dengan cara membandingkan penyaluran autokorelasi antara data dengan teoritis model, metode ini bersifat rekursif. Arima membentuk ikatan antara prosedur autoregressive (AR) dan moving average (MA). Model ARIMA dituliskan sebagai (a,b,c) dengan model matematis sebagai berikut [11].

$$\phi(B)(1 - B)^d Z_t = \theta(B)a_t \quad \dots\dots\dots(1)$$

Dimana (a, b, c) ; orde AR (a), orde differencing (b), orde MA (c) untuk pola non musiman;  $\phi(B)$  : koefisien AR dengan derajat a;  $\theta(B)$  : koefisien MA dengan derajat c.

Dalam pengujian ini digunakan nilai RMSE yang paling kecil antara Arima dan LSTM. Secara garis besar nilai RMSE dipahami untuk mengetahui seberapa jauh nilai kesalahan dari suatu model terhadap garis regresi linier [11]. RMSE berhubungan dengan variasi sebaran frekuensi (frequency distribution) dari besar kesalahan yang diperoleh, tapi tidak dengan variasi kesalahan. Jadi dapat dilihat bahwa RMSE mampu meredam perubahan kesalahan yang besar.

Rumus RMSE;

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_1 - \hat{y}_1)^2}{n}}$$

$\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n = \text{nilai yang diprediksi}$   
 $y_1, y_2, \dots, y_n = \text{nilai yang diamati}$   
 $n = \text{jumlah data yang diamati}$   
 .....(2)

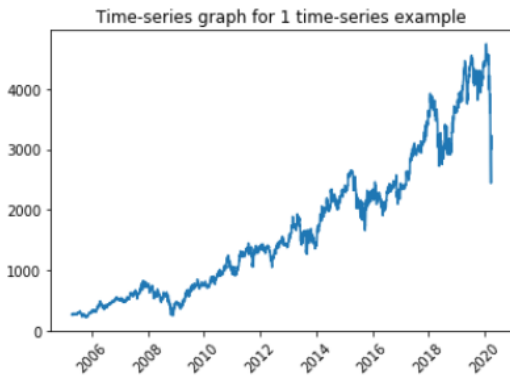
**HASIL DAN PEMBAHASAN**

Proses pertama sebelum mendapatkan pemodelan terbaik guna proyeksi data saham Bank BRI yaitu melalui transformasi, *smoothing*, dan differensing, *cleansing*, proses transformasi log dan moving average (*ts\_log\_moving\_avg*). Pada tabel 1 yaitu sumber data saham Bank BRI Tbk (2005-2020) sejumlah 3717 data, melalui proses *cleansingdata* sebagai berikut :

**Tabel 1** Data saham Bank BRI Tbk (2005-2020)

listing_id	listing_id
Date	Date
2005-04-18 267.5	2020-03-23 2620.0
2005-04-19 277.5	2020-03-24 2440.0
2005-04-20 285.0	2020-03-26 2940.0
2005-04-21 272.5	2020-03-27 3230.0
2005-04-25 265.0	2020-03-30 3010.0

(3717, 1)



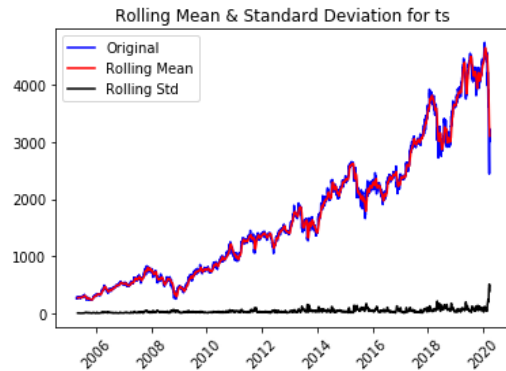
**Grafik 2** Fluktuatif data saham Bank BRI Tbk (2005-2020)

Terlihat pada grafik 2 nilai data saham Bank BRI Tbk secara annual berfluktuatif positif.

**Tabel 2** Data Statistik saham Bank BRI Tbk (2005-2020)

listing_id	
count	3717.000000
mean	1716.767554
std	1179.077456
min	232.500000

Hasil mean dan std tidak stasioner (karena sangat besar) maka perlu diuji menggunakan Dickey Fuller Test (grafik 3), karena p value > 0,5 data tersebut tidak stasioner.

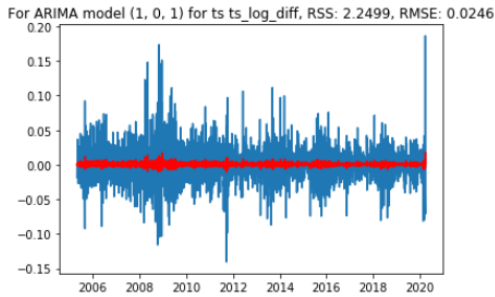


Results of Dickey-Fuller Test:  
 Test Statistic -1.082253  
 p-value 0.722097

**Grafik 3** Dickey Fuller Test

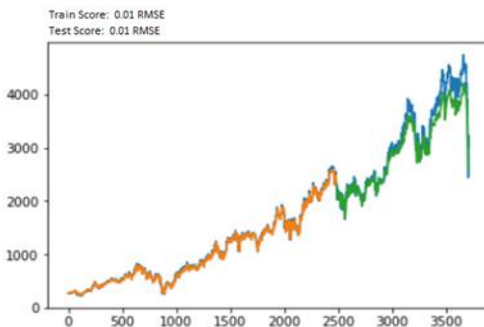
Pengolahan data pada statistik harus stasioner. Dengan melalui proses seperti transformasi, *smoothing* dan differensing. Penghitungan Dickey Fuller dilakukan untuk delapan jenis proses test dan grafik mean dan standar deviasinya.

**Tabel 2** Proses transformasi, *smoothing* dan differensing (Lampiran)



Grafik 4-a. *ts\_log\_moving\_avg\_diff* ARIMA

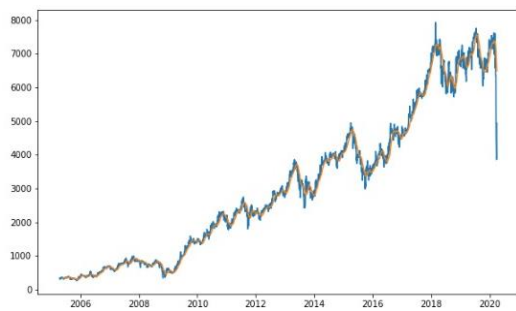
Menghasilkan transformasi terbaik pada Arima (1,0,1) dan RMSE 0,0246.



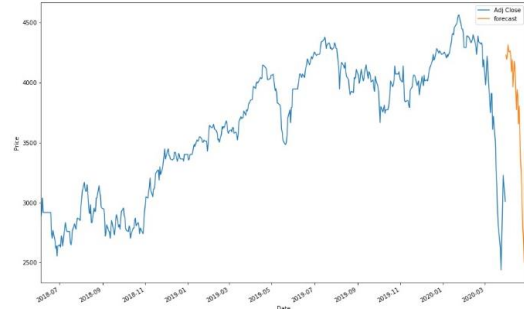
Grafik 4-b. *ts\_log\_moving\_avg\_diff* LSTM

Menghasilkan transformasi terbaik pada LSTM nilai train 0,01 dan RMSE tes 0,01.

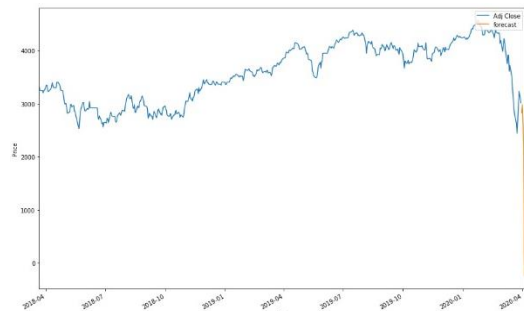
Paket PROPHET[12] digunakan untuk proyeksi grafik harian, mingguan, bulanan dan tahunan sebagai berikut :



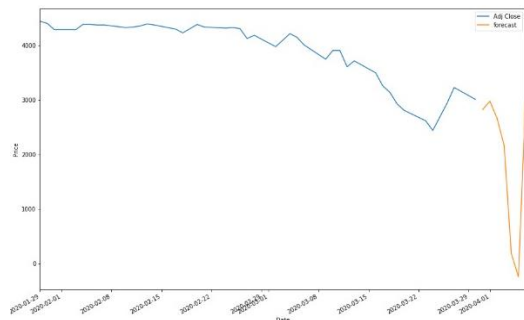
Grafik 5-a. Proyeksi Tahunan Saham Bank BRI Tbk



Grafik 5-b. Proyeksi Bulanan Saham Bank BRI Tbk



Grafik 5-c. Proyeksi Mingguan Saham Bank BRI Tbk



Grafik 5-d. Proyeksi Mingguan Saham Bank BRI Tbk

## PENUTUP

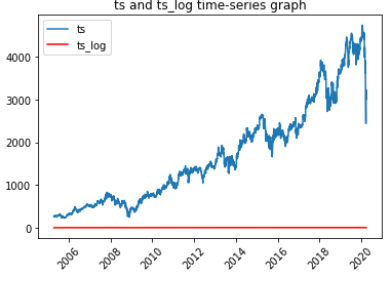
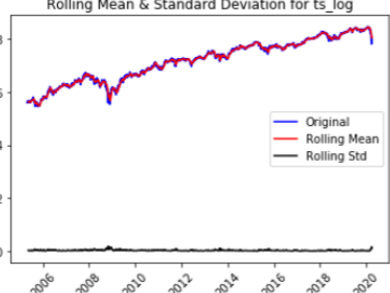

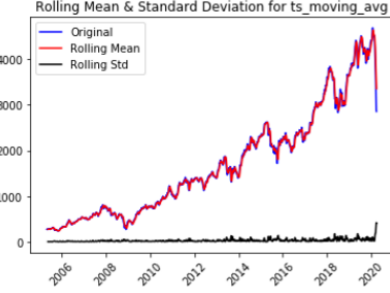
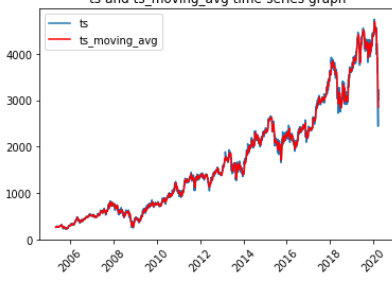

Setelah melalui proses dari transformasi, smoothing, dan differencing menghasilkan; (1) Perbedaan nilai RMSE dari statistik ARIMA dan LSTM, (2) Hasil statistik terbaik ARIMA (1,0,1) setelah melalui tahap transformasi *ts\_log\_moving\_avg\_diff* (3) Hasil ML LSTM terbaik dengan nilai RMSE = 0,01 setelah melalui tahap transformasi *ts\_log\_moving\_avg*, (4) Hasil dari eksperimen ini yaitu LSTM lebih prioritas dibanding dengan statistik ARIMA, (5) Input model statistik untuk LSTM berhasil meningkatkan nilai error 1%, hal tersebut

dapat disimpulkan mendekati data aslinya (99%).

#### DAFTAR PUSTAKA

- [1] BRI, "Sejarah Bank BRI." [Online]. Available: <https://bri.co.id/tentang-bri>. [Accessed: 20-Aug-2020].
- [2] CNN Indonesia, "BUMN bagi dividen 624 T," 2019. [Online]. Available: <https://www.cnnindonesia.com/ekonomi/20190527125901-92-398708/13-bumn-bagi-dividen-rp624-triliun>.
- [3] Yahoo finance, "BBRI\_JK History." [Online]. Available: <https://finance.yahoo.com/quote/BBRI.JK/history?period1=946659600&period2=1009818000&interval=1d&filter=history&frequency=1d>.
- [4] A. Satyo and B. Karno, "Analisis Data Time Series Menggunakan LSTM ( Long Short Term Memory ) dan ARIMA ( Autocorrelation Integrated Moving Average ) dalam Bahasa Python," *Ultim. InfoSys*, vol. XI, no. 1, pp. 1–7, 2020.
- [5] T. Fischer and C. Krauss, "Networks for Financial Market Predictions," *FAU Discuss. Pap. Econ. No. 11/2017*, Friedrich-Alexander-Universität Erlangen-Nürnberg, Inst. Econ. Erlangen, pp. 1–34, 2017.
- [6] S. Hochreiter and J. Urgan Schmidhuber, "Long Short Term Memory," *Neural Comput.*, vol. 9, no. 8, p. 17351780, 1997.
- [7] W. Hastomo and A. Satyo, "Long Short Term Memory Machine Learning Untuk Memprediksi Akurasi Nilai Tukar IDR Terhadap USD," vol. 3, 2019.
- [8] J. Eapen, D. Bein, and A. Verma, "Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction," *2019 IEEE 9th Annu. Comput. Commun. Work. Conf. CCWC 2019*, pp. 264–270, 2019.
- [9] P. T. Yamak, L. Yujian, and P. K. Gadosey, "A comparison between ARIMA, LSTM, and GRU for time series forecasting," *ACM Int. Conf. Proceeding Ser.*, pp. 49–55, 2019.
- [10] D. Karmiani, R. Kazi, A. Nambisan, A. Shah, and V. Kamble, "Comparison of Predictive Algorithms: Backpropagation, SVM, LSTM and Kalman Filter for Stock Market," *Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019*, pp. 228–234, 2019.
- [11] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: an analysis and review," *Int. J. Forecast. Elsevier*, vol. 16, no. 4, pp. 437–450, 2000.
- [12] Facebook, "Prophet forecasting at scale," *Menlo park, California*. [Online]. Available: <https://github.com/facebook/>.

LAMPIRAN

	TRANSFORMASI, SMOOTHING, DIFFERENSING	MEAN, STANDAR DEVIASI, DICKEY FULLER TEST
ts_log	 <p>ts and ts_log time-series graph</p>	 <p>Rolling Mean &amp; Standard Deviation for ts_log</p> <p>Results of Dickey-Fuller Test:                      Test Statistic -1.736508                      p-value 0.412407</p>
ts_log_moving_avg	 <p>ts_log and ts_log_moving_avg time-series graph</p>	 <p>Rolling Mean &amp; Standard Deviation for ts_log_moving_avg</p> <p>Results of Dickey-Fuller Test:                      Test Statistic -1.184813                      p-value 0.680094</p>
ts_moving_avg	 <p>ts and ts_moving_avg time-series graph</p>	 <p>Rolling Mean &amp; Standard Deviation for ts_moving_avg</p> <p>Results of Dickey-Fuller Test :                      Test statistic -0.499464                      p-value 0.892121</p>

