

Rancangan Sistem Prediksi Harga Saham dengan Menggunakan Metode LSTM dan ARMA klasik

Caesar Calendo Sumarga¹⁾, Dyah Erny Herwindiati²⁾, Janson Hendryli³⁾

^{1,2,3} Teknik Informatika, FTI, Universitas Tarumanagara
Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia

¹⁾E-mail: caesar.535190053@stu.untar.ac.id, ²⁾E-mail: dyahh@fti.untar.ac.id, ³⁾E-mail: jansonh@fti.untar.ac.id

ABSTRACT

Stocks are one of the types of assets that are currently popular with the wider community, just like gold and all other types of assets, the value of stocks tends to move up and down over time, therefore stock investors invest in stocks to achieve the desired profit (capital gain). Due to the movement of stocks that go up and down over time, it is difficult for investors to determine when to buy or sell stocks, therefore this study was conducted to compare the multivariate Long-Short Term Memory (LSTM) method, and the classic ARMA, then see which is suitable in forecasting stock prices, the comparison is seen from the results of the error evaluation metrics of the two methods.

Key words

ARMA, forecasting, LSTM, Stock, Time Series

1. Pendahuluan

Di Era pesatnya teknologi di dunia ini yang memberikan pengaruh, baik langsung maupun tidak langsung pada berbagai sektor, tidak terkecuali pada sektor pasar saham. Kerumitan indikator dan juga pergerakan yang cenderung fluktuatif pada pasar saham menjadi hal yang perlu menjadi perhatian untuk dapat memperoleh keuntungan dalam dunia pasar saham. Berbagai metode mulai diciptakan oleh manusia dengan kemampuan berfikirnya untuk memperkirakan kejadian di masa yang akan datang.

Berdasarkan data dari PT KSEI terdapat peningkatan jumlah investor pasar modal di Indonesia yang mencapai 7,48 juta pada tahun 2021 yang naik 92,99% dari tahun sebelumnya. Namun dibalik peningkatan kuantitas yang luar biasa belum dapat diimbangkan dengan kualitas investor dalam mempertimbangkan pemilihan saham yang dimiliki untuk itu dengan dibuatnya penelitian ini dapat membantu investor dalam memprediksi harga saham. Selain itu diharapkan dapat menjadi salah satu bahan referensi dan literatur dalam bidang teknologi informasi khususnya dalam hal prediksi saham[1]. Prediksi adalah

suatu proses perkiraan secara sistematis terhadap sesuatu yang dapat terjadi dimasa depan.

Di Indonesia terdapat beberapa indeks untuk mengukur kinerja dari saham yang terdaftar, salah satu indeksnya yaitu LQ45, LQ45 adalah representasi dari harga saham 45 emiten yang ada pada Bursa Efek Indonesia(BEI), LQ45 terpilih berdasarkan tingkat likuiditas dan kapitalis terbesar dengan kriteria yang sudah ditentukan, perusahaan pada saham LQ45 akan berubah setiap enam bulan sekali berdasarkan ketentuan dari pihak BEI.

Prediksi harga saham menjadi salah satu langkah yang dapat ditempuh oleh seorang investor untuk mengurangi kerugian dan memperbesar peluang laba dimasa mendatang. Untuk itu Melakukan prediksi harga saham diperlukan metode dan pendekatan dengan akurasi yang baik. Semakin tinggi akurasi yang dimiliki suatu metode prediksi yang dibuat juga akan meningkatkan peluang keuntungan dari saham yang dibeli.

Beberapa tahun ini telah banyak model yang digunakan untuk melakukan prediksi, beberapa model diantaranya yaitu model LSTM dengan pendekatan Deep Learning dan model AutoRegressive Moving Average (ARMA) yang merupakan model dengan pendekatan statistik. Untuk metode yang pertama yaitu metode LSTM atau Long Short Term Memory merupakan pengembangan dari metode RNN atau Recurrent Neural Network pengembangan tersebut bertujuan agar dapat memecahkan masalah terkait ketergantungan jangka panjang, dari beberapa literatur juga disebutkan bahwa metode ini lebih cocok untuk data yang bersifat sekuensial seperti data deret waktu[2], pada rancangan ini metode LSTM ini akan diimplementasikan kedalam kasus deret waktu khususnya untuk melakukan prediksi dari harga saham.

Untuk metode selanjutnya yaitu metode ARMA digunakan untuk dilakukannya perbandingan dengan metode LSTM, dengan menggunakan dataset yang sama nantinya akan dibandingkan kedua metode tersebut. Metode ARMA memiliki ketepatan dalam peramalan jangka pendek yang baik, dengan demikian metode cocok untuk bagi para trader saham jangka

pendek, tetapi untuk peramalan jangka panjang ketepatannya kurang baik sehingga tidak cocok untuk para investor saham jangka panjang.

2. Metodologi

Data yang digunakan pada penelitian ini adalah data harga penutupan dari saham Bank BCA dengan kode ticker BBKA, data tersebut diperoleh dari situs Yahoo Finance, jumlah data yang diambil adalah sebanyak 1304 data yang diperoleh yaitu harga pembukaan, tertinggi, terendah, dan penutupan, setelah data dikumpulkan akan dilakukan pra-pemrosesan data, kemudian perancangan model untuk tiap metodenya, dan terakhir yaitu untuk dilakukan evaluasi terhadap masing masing metode.

2.1 Pengumpulan data

	Open	High	Low	Close	Adj Close	Volume
Date						
2017-09-19	3770.0	3835.0	3765.0	3815.0	3523.369385	59579500.0
2017-09-20	3820.0	3820.0	3795.0	3810.0	3518.752197	76382000.0
2017-09-21	3810.0	3810.0	3810.0	3810.0	3518.752197	0.0
2017-09-22	3820.0	3995.0	3805.0	3975.0	3671.138916	125871000.0
2017-09-25	3950.0	3960.0	3925.0	3935.0	3634.196533	97549500.0
...
2022-11-22	8900.0	8950.0	8850.0	8900.0	8900.000000	84684200.0
2022-11-23	8900.0	8975.0	8875.0	8875.0	8875.000000	84642100.0
2022-11-24	8975.0	9000.0	8950.0	9000.0	9000.000000	53520800.0
2022-11-25	9000.0	9000.0	8900.0	8975.0	8975.000000	75048300.0
2022-11-28	9050.0	9100.0	8975.0	9025.0	9025.000000	96209700.0
1304 rows x 6 columns						

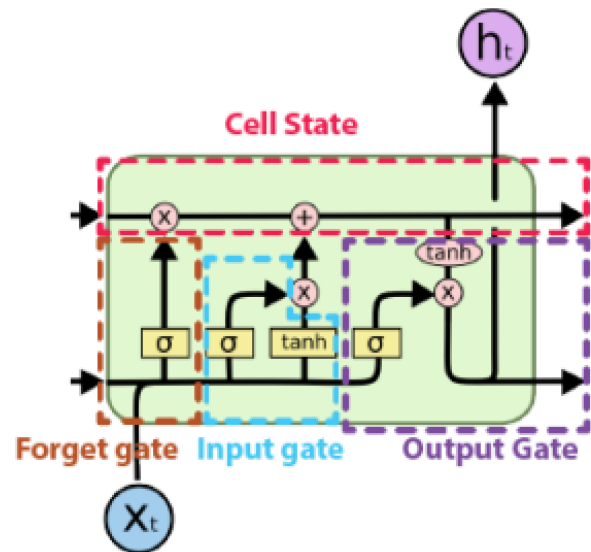
Gambar 1 Data harga saham Bank BCA

Date pada data tersebut mengacu pada tanggal dari data harga tersebut diterbitkan, Open adalah harga pembukaan pada hari itu, High adalah harga tertinggi hari itu, Low adalah harga terendah pada hari itu, close adalah harga penutupan pada hari itu, Adj Close adalah harga penutupan yang mempengaruhi pembagian dividen dan stock split, dan volume adalah jumlah lembar saham pada satu periode, pada penelitian ini, untuk metode LSTM digunakan empat fitur input yaitu fitur Open, High, Low, dan Close, sedangkan metode ARMA digunakan satu fitur input yaitu fitur Close.

2.2 Metode Long Short Term Memory(LSTM)

Long Short-Term Memory atau LSTM ini merupakan bentuk model dari Recurrent Neural Network atau RNN, LSTM muncul untuk menghindari permasalahan dari RNN yaitu masalah ketergantungan informasi jangka panjang. Pada LSTM, neuron pada

Hidden layer RNN digantikan oleh sel memori LSTM guna menjaga informasi sebelumnya



Gambar 2 Arsitektur LSTM

Pada gambar diatas dapat dilihat arsitektur LSTM terdiri dari empat gate yaitu forget gate (f_t), input gate (i_t), cell state (C_t), dan output gate (O_t).

1. Forget Gate

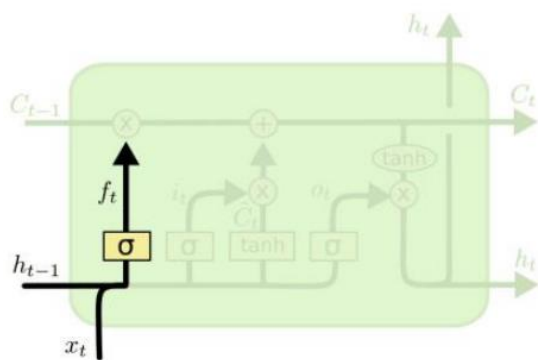
Pada komponen forget Gate (f_t) pada tahap ini bertujuan untuk menentukan apakah informasi pada cell state sebelumnya itu diperlukan atau tidak, untuk forget gate dapat didefinisikan dengan persamaan berikut

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

Keterangan :

f_t = Forget Gate layer orde ke-t
 σ = Fungsi sigmoid
 W_f = Bobot pada Forget Gate Layer
 h_{t-1} = hidden state pada orde t-1
 X_t = input pada orde ke-t
 b_f = bias pada forget gate layer

Berikut Forget Gate Pada LSTM bisa dilihat pada Gambar 3.



Gambar 3 Forget Gate Pada LSTM

2. Input Gate dan candidate gate

Tahap selanjutnya yaitu Gerbang Input, pada tahapan ini informasi akan diolah dan ditentukan banyaknya input yang nantinya akan dimasukkan kedalam cell state, Input gate (i_t) dan candidate gate (\hat{C}_t) dapat didefinisikan sebagai persamaan berikut:

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

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (3)$$

Keterangan :

i_t = Input Gate layer orde ke-t

\hat{C}_t = Kandidat baru untuk dimasukan ke cell state

σ = fungsi sigmoid

\tanh = fungsi tanh

W_i = Bobot pada Input Gate

W_c = bobot pada candidate Gate

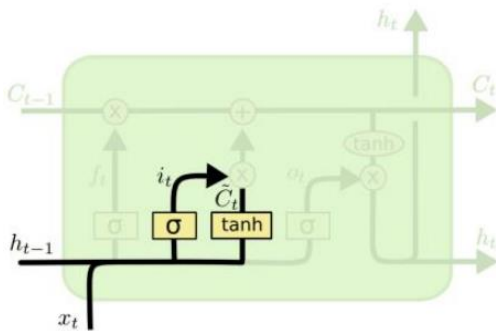
h_{t-1} = hidden state pada orde t-1

X_t = input pada orde ke-t

b_i = bias pada Input gate

b_c = bias pada candidate gate

Berikut untuk Input Gate pada LSTM dapat dilihat pada Gambar 4.



Gambar 4 Input Gate Pada LSTM

3. Cell state

Tahap selanjutnya yaitu tahap cell state, pada tahap ini ,informasi dari gate sebelumnya akan melalui dua tahapan perhitungan yaitu forget Gate (f_t) dan Input gate(i_t), untuk persamaannya adalah sebagai berikut :

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (4)$$

Keterangan :

C_t = cell state orde ke-t

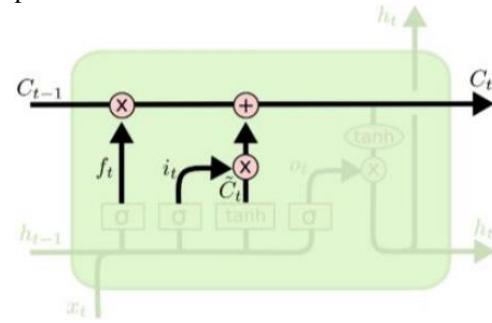
f_t = Forget Gate layer orde ke-t

C_{t-1} = cell state orde ke t-1

i_t = Input Gate layer orde ke-t

\hat{C}_t = Kandidat baru untuk dimasukan ke cell state

Berikut untuk Cell State pada LSTM dapat dilihat pada Gambar 5.



Gambar 5 Cell State Pada LSTM

4. Output Gate

Untuk menentukan output pada Hidden State dengan menjalankan sigmoid kemudian menempatkan cell state pada tanh. . Untuk proses dari output gate dapat didefinisikan sebagai persamaan berikut :

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

$$h_t = O_t * \tanh(C_t) \quad (6)$$

Keterangan :

O_t = Input Gate layer orde ke-t

σ = fungsi sigmoid

\tanh = fungsi tanh

W_o = Bobot pada Output Gate

h_{t-1} = hidden state pada orde t-1

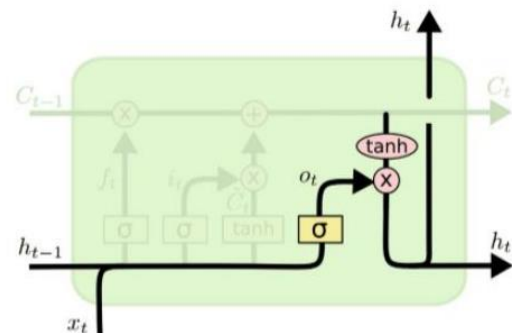
X_t = input pada orde ke-t

b_o = bias pada Output gate

h_t = hasil perhitungan hidden state ke-t

C_t = cell state orde ke-t

Berikut untuk Output Gate pada LSTM dapat dilihat pada Gambar 6.



Gambar 6 Output Gate Pada LSTM

2.3 Metode ARMA

ARIMA disebut juga metode deret waktu Box-Jenkins, ARIMA membutuhkan data historis untuk

melakukan peramalan karena cara kerja ARIMA yaitu dengan mencari pola-pola data dari sekumpulan data yang ada, maka dari itu ARIMA memiliki ketepatan yang baik dalam peramalan jangka pendek, sedangkan untuk peramalan jangka panjang, hasil yang didapatkan akan cenderung konstan, ARMA merupakan salah satu dari kelompok model ARIMA, tepatnya ARMA adalah ARIMA dengan orde $(p,0,q)$, model ARMA merupakan gabungan dari model AR dan MA, didefinisikan dengan rumus :

$$x_t = \mu' + \phi_1 x_{t-1} + e_t - \theta_1 e_{t-1} \quad (7)$$

Keterangan :

μ' = suatu konstanta

θ_1 sampai θ_q = parameter moving average

e_t = nilai kesalahan pada saat t

e_{t-1} = nilai kesalahan pada saat t-1

ϕ_p = parameter autoregresif ke p

X_{t-1} = nilai X pada keadaan t-1

2.3 Metode Evaluasi

Metode Evaluasi disini untuk menilai keakuratan dari model yang dibuat. Berikut Beberapa Metode yang digunakan untuk melakukan Evaluasi Terhadap Model yang digunakan yaitu :

1. Mean Square Error (MSE)

Mean Square Error(MSE) adalah hasil penjumlahan kuadrat dari kesalahan diantara nilai aktual dengan nilai peramalan dan dibagi dengan jumlah waktu peramalan, secara perhitungan MSE adalah sebagai berikut [4]:

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (8)$$

Keterangan :

\hat{Y}_i = nilai peramalan

Y_i = nilai aktual

n = jumlah periode data

2. Mean Absolute Percentage Error(MAPE)

adalah rata rata diferensiasi absolut antara nilai peramalan dan nilai aktual yang didefinisikan dalam bentuk persentase. Berikut untuk rumus perhitungan MAPE:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \quad (9)$$

Keterangan :

\hat{Y}_i = nilai peramalan

Y_i = nilai aktual

n = jumlah periode data

Untuk kriteria MAPE yang baik yaitu MAPE dengan nilai terkecil atau semakin kecil nilainya semakin baik akurasi, untuk tabel Kriteria MAPE Dapat dilihat pada tabel 1 berikut :

Tabel 1 Kriteria MAPE

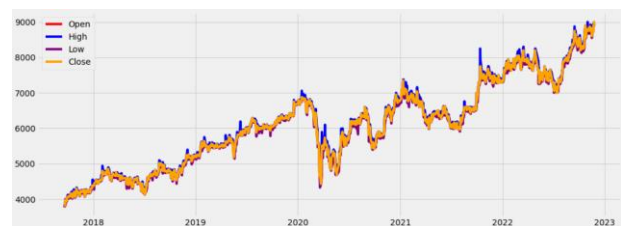
Range Mape	Arti Nilai
< 10%	Kemampuan model peramalan sangat baik
10-20%	Kemampuan model peramalan baik
20-50%	Kemampuan model peramalan cukup
>50%	Kemampuan model peramalan buruk

3. Hasil Percobaan

Hasil Percobaan yang dilakukan dengan kedua metode yaitu metode LSTM dan ARMA, dan nantinya dari kedua metode tersebut akan dibandingkan hasilnya, untuk metode LSTM akan mempertimbangkan empat fitur(multivariate) variable yaitu fitur open, high, low, dan close untuk memprediksi harga closenya, sedangkan untuk metode ARMA akan menggunakan satu fitur(univariate) variable yaitu fitur close.

3.1 Hasil Percobaan LSTM

Pertama akan diambil dulu data dari website yahoo finance dengan menggunakan library, data diambil dari tanggal 2017-09-19 sampai dengan tanggal 2022-11-27, untuk grafik data saham bank BCA dapat dilihat pada Gambar 7.



Gambar 7 Grafik Saham Bank BCA

Sebelum dilakukan Training Data akan terlebih dahulu dipisahkan untuk data training dan data testingnya, untuk data training digunakan sebanyak 80% dari data, sedangkan data testing menggunakan sisanya yaitu 20%, pembagian data dapat dilihat pada Gambar 8.



Gambar 8 Data Training dan Data Testing

Pada Percobaan LSTM dilakukan beberapa skenario untuk menentukan arsitektur LSTM yang tepat seperti perbandingan neuron dan epochnya, skenario percobaan dapat dilihat pada Tabel 2.

Tabel 2 Skenario Menentukan Parameter LSTM

Percobaan	Neuron	Epoch
1	32	50
2		100
3		150
4	64	50
5		100
6		150
7	128	50
8		100
9		150

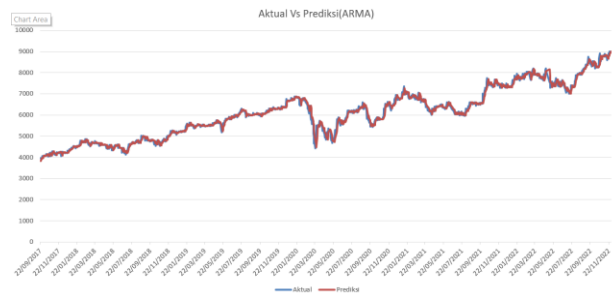
Percobaan dilakukan training sebanyak satu kali dan hasilnya didapatkan parameter dengan nilai Evaluasi terkecil adalah dengan neuron sebanyak 128, epoch sebesar 100 dengan nilai MAPE Testing sebesar 6.7889%, untuk Grafik Perbandingan dari nilai Actual dengan nilai prediksi dapat dilihat pada Gambar 9.



Gambar 9 Actual Data vs Predicted Data(LSTM)

3.2 Hasil Percobaan ARMA

Pada Percobaan ARMA pertama ditentukan untuk Orde yang akan digunakan adalah ARMA berorde $p = 1$, dan $q = 1$ sehingga untuk model ARMA yang digunakan adalah model $AR_{(1)}MA_{(1)}$. Pada Perhitungan model $AR_{(1)}MA_{(1)}$ didapatkan Hasil dari percobaan dengan Model $AR_{(1)}MA_{(1)}$ mendapatkan nilai MAPE sebesar 1,485%, untuk grafik perbandingan nilai aktual dengan nilai prediksi metode ARMA dapat dilihat pada Gambar 10.



Gambar 10 Actual Data vs Predicted Data(ARMA)

3.3 Pengujian Program

Pada Pengujian Program disini dilakukan Prediksi terhadap data 2 hari kedepan yaitu pada tanggal 28-11-2022 sampai tanggal 29-11-2022 menggunakan kedua metode tersebut kemudian dibandingkan untuk hasil prediksi dari kedua model tersebut dengan nilai aktual dari harga close saham bank BCA, untuk model LSTM akan dilakukan lima kali iterasi untuk diambil rata-ratanya, training dari 5 iterasi tersebut dapat dilihat pada Tabel 3.

Tabel 3 Skenario Menentukan Parameter LSTM

Iterasi	Prediksi dua hari kedepan	MAPE
1	8190	6,790%
	8220	
2	8310	5,936%
	8350	
3	8470	4,059%
	8510	
4	8150	8,517%
	8190	
5	8350	5,372%
	8390	

Dari lima iterasi tersebut didapatkan untuk rata-rata MAPE testing-nya adalah sebesar 6,13%, dengan rata-rata prediksi pada tanggal 28-11-2022 adalah 8294, dan untuk tanggal 29-11-2022 adalah 8332, setelah didapatkan rata-ratanya kemudian dibandingkan dengan Metode ARMA dan juga nilai aktual pada tanggal 28 dan 29.

Tabel 4 Perbandingan Hasil Metode LSTM dan ARMA dengan Nilai Aktualnya

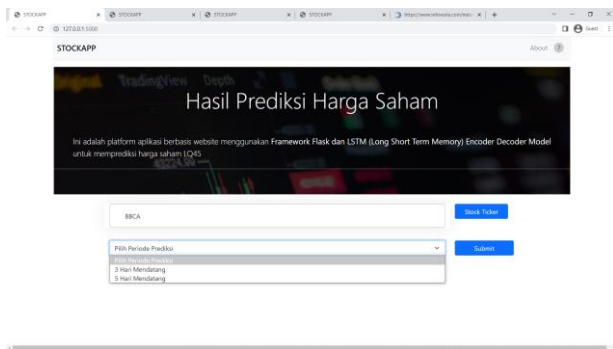
	Tanggal	Nilai prediksi	Nilai Aktual	MAPE
LSTM	28/11/2022	8294	9025	8,10%

	29/11/2022	8332	8975	7,16%
ARMA	28/11/2022	8935	9025	1,00%
	29/11/2022	8997	8975	0,25%

3.4 Tampilan AntarMuka WebApp

Tampilan AntarMuka WebApp Prediksi harga saham memiliki 5 modul yaitu modul Beranda, modul list Ticker, Modul Prediksi, modul About, dan modul Help.

Modul yang pertama yaitu Modul Beranda, pada tampilan beranda terdapat kolom untuk mengisi nama kode emiten yang ingin di prediksi, untuk list tickernya bisa dilihat pada modul ticker list yang bisa diakses dengan mengklik Stock Ticker, selanjutnya terdapat pilihan untuk memilih periode prediksi dari 3 hari kedepan dan 5 hari kedepan. modul ini dapat dilihat pada Gambar 11.



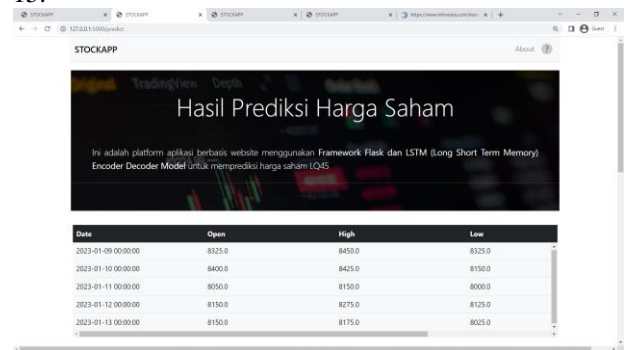
Gambar 11 Modul Beranda

Selanjutnya ada modul ticker list, berisi sekumpulan saham emiten LQ45 salah satunya BBKA, modul ini dapat dilihat pada Gambar 12.

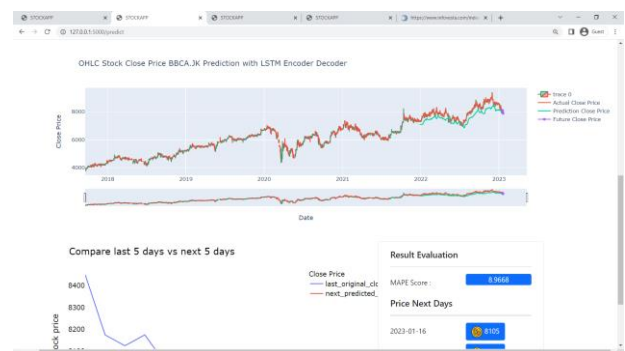
Kode	Open	High	Low	Close	1 Hr	1 Bln	1 Tln
ADRO	3,120	3,150	3,110	3,140	1,62	-16,04	37,72
AMRT	2,720	2,770	2,700	2,760	1,84	6,15	152,05
ANTM	2,150	2,170	2,090	2,150	0,94	11,4	9,41
ARTO	3,250	3,290	3,010	3,050	-4,09	-25,97	-62,45
ASII	5,375	5,500	5,300	5,450	1,87	-4,44	-5,63
BBKA	8,150	8,175	8,025	8,050	-1,53	-7,47	4,55
BBRI	8,675	8,675	8,575	8,675	0,87	-10,1	23,49
BBTN	4,510	4,510	4,420	4,480	-0,44	-9,90	6,82
BETH	1,240	1,260	1,225	1,240	-0,8	-11,43	-27,06
BPIN	1,100	1,105	1,080	1,090	-0,40	0,46	-20,73
BWRI	9,125	9,225	9,050	9,175	-0,27	-7,32	28,77

Gambar 12 Modul Ticker List

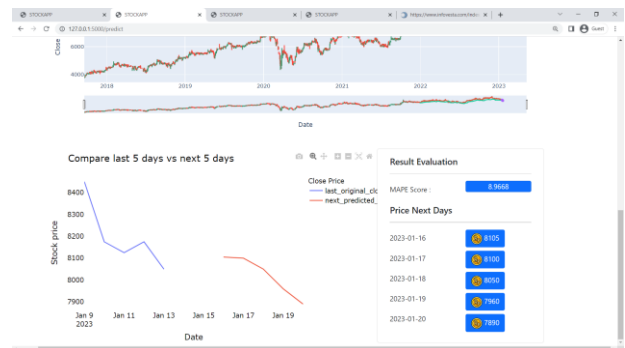
Modul Prediksi berisi chart-chart dan juga tabel hasil dari prediksi yang dijalankan program, modul ini dapat dilihat pada Gambar 13, Gambar 14, dan Gambar 15.



Gambar 13 Modul Prediksi



Gambar 14 Modul Prediksi(Lanjutan 1)



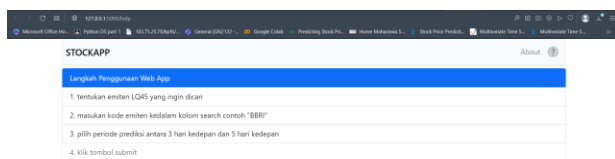
Gambar 15 Modul Prediksi(Lanjutan 2)

Modul About berisi tentang detail dari pengembang dari webapp ini, untuk modul About dapat dilihat pada Gambar 16.



Gambar 16 Modul About

Terakhir adalah Modul Help, pada modul help berisi tentang panduan untuk menjalankan webapp ini, untuk halaman modul help dapat dilihat pada Gambar 17.



Gambar 17 Modul Help

4. Kesimpulan

Dari Penelitian ini diketahui bahwa dapat di mungkinkan untuk dilakukannya prediksi terhadap harga saham, yang menjadi permasalahannya adalah bagaimana untuk mendapatkan model prediksi dengan akurasi yang mendekati nilai aktual. Beberapa Hal yang dapat disimpulkan dari penelitian ini yaitu :

- Prediksi Harga Saham Menggunakan Metode LSTM Multivariate memiliki ketepatan yang baik dalam memprediksi harga saham data Bank BCA hal ini dapat dilihat untuk nilai MAPE testing rata rata dari 5 iterasi
- menunjukkan nilai 6,13% yang dimana mengacu pada tabel kriteria MAPE merupakan Kriteria MAPE sangat baik.
- Penggunaan Metode ARMA dengan model $AR_{(1)}MA_{(1)}$ Memiliki ketepatan yang sangat baik dengan akurasi MAPE sebesar 1,485%
- Metode ARMA dengan model $AR_{(1)}MA_{(1)}$ Memiliki Akurasi yang lebih baik dalam memprediksi Harga saham Bank BCA dimana dilihat pada prediksi dua hari kedepan MAPE antara Nilai prediksi dengan Nilai aktual metode ARMA lebih kecil yaitu jika dirata rata adalah sebesar 0,62% dibandingkan dengan

MAPE rata-rata metode LSTM yaitu sebesar 7,63%

REFERENSI

- [1] PT KSEL. Statistik Pasar Modal Indonesia. https://www.ksei.co.id/files/Statistik_Publik_Februari_2021.pdf. tanggal akses 12 September 2022.
- [2] Kurniawan, Febryo; Herwindiati, Dyah Erny; Lauro, Manatap Dolok.” Raw Paper Material Stock Forecasting with Long Short-Term Memory”. 9th International Conference on Information and Communication Technology (ICoICT). 2021..
- [3] Mao, Yaobin., Chen, Guanrong., 2003, “Chaos-Based Image Encryption”, Handbook of Computational Geometry for Pattern Recognition, Computer Vision, Neural Computing and Robotics, Springer, Berlin.
- [4] Hajjah, Alyauma dan Marlim, Yulvia Nora. “Analisis Error Terhadap Peramalan Data Penjualan”. Techno.COM. Vol. 20, No. 1, 2021.