

# Perbandingan Penerapan Metode Bidirectional Long Short Term Memory dan Autoregressive Integrated Moving Average dalam Prediksi Harga Saham

Yoel Agustinus<sup>#1</sup>, Ventje J. L. Engel, S.T., M.T.<sup>\*2</sup>

<sup>#</sup>Program Studi Informatika, Institut Teknologi Harapan Bangsa  
Jalan Dipatiukur No. 80-84, Bandung, Indonesia, 40132

<sup>1</sup>yoelagustinus7@gmail.com

<sup>2</sup>ventje@ithb.ac.id

**Abstract**— The stock price is a very important factor and must be considered by investors in making investments because the stock price shows the performance of a company's issuer. This study compares the Long Short Term Memory (LSTM) model with the Autoregressive Integrated Moving Average (ARIMA), to find the best method for predicting stock prices. The results of this study ARIMA has a better average performance when compared to the LSTM model. This is because LSTM uses a larger number of features, not only Close but uses Open, High, Low, Volume for observation and prediction, while ARIMA only uses the Close feature for observation and prediction. The average values of RMSE, MAE, MAPE ARIMA are 363.11, 266.75, 2.06 respectively. While the average LSTM error values are 621.98, 525.51, 3.83. PSDN it is more suitable to make predictions using the method in this study. Because the RMSE, MAE, MAPE obtained are much better than Blue Chip companies and PSDN prices are cheaper (hundreds) compared to GGRM or UNVR stock prices which cost thousands and tens of thousands.

**Keywords**— Stocks Price, Long Short Term Memory, Autoregressive Integrated Moving Average, Neural Network, Prediction, Epoch, Unit, (p,d,q)

**Abstrak**— Harga saham merupakan faktor yang sangat penting dan harus diperhatikan oleh investor dalam melakukan investasi karena harga saham menunjukkan prestasi emiten sebuah perusahaan. Penelitian ini membandingkan model Long Short Term Memory (LSTM) dengan Autoregressive Integrated Moving Average (ARIMA), untuk menemukan metode terbaik bagi prediksi harga saham. Hasil dari penelitian ini ARIMA memiliki performa rata-rata yang lebih baik jika dibandingkan dengan model LSTM. Hal ini disebabkan, LSTM menggunakan jumlah fitur yang lebih banyak, tidak hanya Close tetapi menggunakan Open, High, Low, Volume untuk observasi dan melakukan prediksi, sementara ARIMA hanya menggunakan fitur Close untuk observasi dan prediksi. Nilai rata-rata RMSE, MAE, MAPE ARIMA berturut-turut adalah 363.11, 266.75, 2.06. Sementara nilai rata-rata error LSTM ialah 621.98, 525.51, 3.83. PSDN lebih cocok untuk melakukan prediksi dengan menggunakan metode pada penelitian ini. Disebabkan RMSE, MAE, MAPE yang didapatkan sangat jauh lebih baik daripada perusahaan Blue Chip dan harga PSDN lebih murah (ratusan) dibandingkan harga saham GGRM atau UNVR yang harganya ribuan dan puluhan ribu.

**Kata Kunci**— Harga saham, Long Short Term Memory, Autoregressive Integrated Moving Average, Neural Network, Prediksi, Epoch, Unit, (p,d,q)

## I. PENDAHULUAN

Perekonomian Indonesia tidak lepas dari pasar modal. Banyak perusahaan yang menawarkan kepemilikan perusahaan dalam bentuk saham kepada publik [1]. Harga saham merupakan faktor yang sangat penting dan harus diperhatikan oleh investor dalam melakukan investasi karena harga saham menunjukkan prestasi emiten sebuah perusahaan. Terdapat empat komponen utama dalam Harga saham dipasar modal, yaitu: harga tertinggi (*high price*), harga terendah (*low price*), harga penutupan (*close price*), harga penutupan yang sudah disesuaikan (*adjusted close*). Harga tertinggi atau terendah merupakan harga yang paling tinggi atau paling rendah yang terjadi pada satu hari bursa, harga penutupan merupakan harga yang terjadi terakhir pada saat akhir jam bursa, harga penutupan yang disesuaikan merupakan harga penutupan yang telah disesuaikan dengan aksi korporasi seperti *right issue*, *stock split* atau *stock reverse*. Oleh karena itu, diperlukan sistem yang dapat memprediksi harga saham dengan akurat

Di beberapa penelitian sebelumnya mengenai prediksi harga saham, menggunakan metode ARIMA dan mendapatkan nilai RMSE 38.03 [2]. Kemudian ada yang menggunakan berbagai macam metode *Machine Learning* yang berbeda, diantaranya: *ANN*, *Random Forest*, *SVM*, *LSTM*, namun, yang mendapatkan nilai RMSE terendah yaitu metode LSTM dengan nilai RMSE 0.306543[4]. Dan ada yang menggunakan LSTM dan GRU dengan mendapatkan nilai RMSE terendah yaitu 71.658. Lalu, ada penelitian yang membandingkan metode ARIMA dan LSTM, hasil RMSE LSTM lebih baik dari pada ARIMA dengan nilai 64.445 dan 511.381. Objek penelitian dengan menggunakan dataset PT. Unilever Indonesia (UNVR), PT. Gudang Garam (GGRM), PT. Prasadha Aneka Niaga (PSDN). Pengujian dilakukan dengan mengkombinasikan nilai *epoch*, *unit* pada LSTM dan nilai ordo p,d,q pada ARIMA.

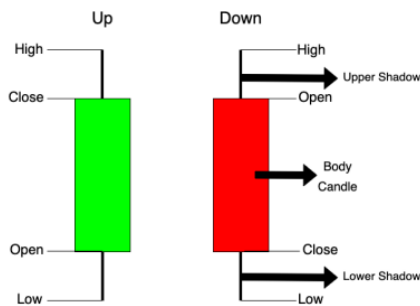
## II. METODOLOGI

### A. Harga Saham

Saham adalah surat yang menjadi bukti seseorang memiliki bagian modal suatu perusahaan. Saham (stock) merupakan salah satu instrumen pasar keuangan yang paling populer. Menerbitkan saham merupakan salah satu pilihan perusahaan ketika memutuskan untuk pendanaan perusahaan [10]. Pada sisi yang lain, saham merupakan instrumen investasi yang banyak dipilih para investor karena saham mampu memberikan tingkat keuntungan yang menarik. Seorang investor umumnya dituntut untuk selalu mengikuti perkembangan informasi pasar dan harga pasar. Terdapat empat komponen utama dalam harga saham dipasar modal, yaitu: harga tertinggi (high price), harga terendah (low price), harga penutupan (close price), harga penutupan yang sudah disesuaikan (adjusted close). Harga tertinggi atau terendah merupakan harga yang paling tinggi atau paling rendah yang terjadi pada satu hari bursa, harga penutupan merupakan harga yang terjadi terakhir pada saat akhir jam bursa, harga penutupan yang disesuaikan merupakan harga penutupan yang telah disesuaikan

#### 1) Candlestick

*Candlestick* adalah salah satu jenis grafik harga saham yang digunakan dalam analisis teknikal yang menunjukkan harga tertinggi, terendah, pembukaan dan penutupan dari suatu saham pada periode waktu tertentu. *Candlestick* yang saat ini digunakan sebagai analisis teknikal di pasar saham memiliki dua komponen utama, yaitu: *Body Candle*, *Shadow Candle*. *Body Candle* adalah bagian dari *candlestick* yang menunjukkan harga pembukaan dan harga penutupan pada titik waktu tertentu yang ditunjukkan dari bentuk persegi empat berwarna merah atau hijau dapat dilihat pada Gambar 2.1. *Shadow Candle* adalah bagian yang menunjukkan harga tertinggi dan harga terendah saham pada titik waktu tertentu yang ditunjukkan dari garis lurus yang membentang di atas dan di bawah *body candlestick* umumnya *shadow candle* itu warnanya sama seperti *body candle*.



Gambar 1 Candlestick [11]

Jika *Close* letaknya di atas *Open*, dapat dikatakan *candlestick* lagi mengalami *bullish* atau saham sedang naik, pada warna hijau. Jika *Close* letaknya di bawah *Open*, maka dapat dikatakan *candlestick* lagi mengalami *bearish* atau saham sedang turun.

### 2) Saham Blue Chip dan Gorengan

Dalam berinvestasi pada suatu perusahaan, terdapat istilah-istilah untuk mendeskripsikan suatu perusahaan untuk membedakan mana perusahaan yang mempunyai kinerja yang baik, atau kinerja yang buruk. Maka terdapat 2 macam jenis saham:

#### a) Saham Blue Chip

Sering diindikasikan mempunyai kinerja perusahaan dan saham yang berkualitas baik. Perusahaan ini biasanya menjual produk-produk yang berkualitas dan digunakan oleh banyak orang. Contoh daftar perusahaan blue chip: Bank BCA(BBCA), PT. Telkom(TLKM), PT. Unilever(UNVR), PT. Astra(ASII), PT. Gudang Garam(GGRM).

#### b) Saham Gorengan

Sering diindikasikan saham gorengan adalah saham suatu perusahaan yang mengalami kenaikan tidak wajar akibat adanya rekayasa oleh pelaku pasar yang bertujuan untuk mendapatkan keuntungan jangka pendek. Contoh daftar perusahaan gorengan: PT. FKS Multi Agro Tbk(FISH), PT. Prsidha Aneka Niaga(PSDN).

### B. Autoregressive Integrated Moving Average (ARIMA)

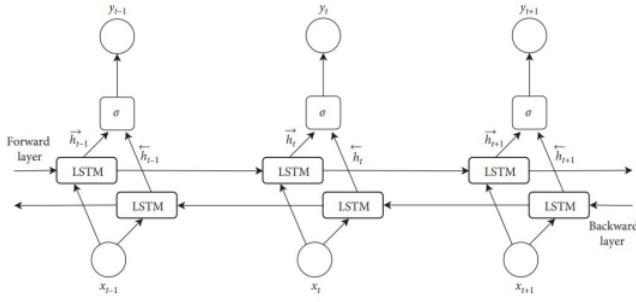
*Autoregressive Integrated Moving Average* atau yang biasa disingkat menjadi ARIMA atau juga dikenal sebagai peramalan *Box and Jenkins*, merupakan metode yang menghasilkan ramalan berdasarkan sintesis dari pola data secara historis [2]. Metode ARIMA sangat baik untuk memprediksi data time series karena berdasarkan asumsi bahwa data time series tersebut stationer yang berarti rata-rata dan varian suatu data time series konstan. Metode ARIMA dibagi dalam 3 unsur, yaitu: Autoregressive(AR), Moving Average(MA), dan Integrated(I). Ketiga unsur ini dimodifikasi secara digabungkan dan menjadi ARIMA(p,d,q), p dinyatakan sebagai ordo AR, d dinyatakan sebagai ordo Integrated atau difference, dan q dinyatakan sebagai ordo MA. Apabila ada nilai parameter p,d,q berturut-turut ialah ARIMA(1,0,0) akan menjadi model yang menjalankan Autoregressive, dan nilai parameter p,d,q berturut-turut ialah ARIMA(0,0,1) akan menjadi model yang menjalankan Moving Average

$$Y_t = c + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \Phi_1 \varepsilon_{t-1} + \dots + \Phi_q \varepsilon_{t-q} + e_t \quad (1)$$

### C. Long Short Term Memory (LSTM)

*Long Short Term Memory* (LSTM) merupakan arsitektur *Recurrent Neural Network* (RNN) yang digunakan dalam bidang *deep learning* untuk melakukan pembelajaran jangka panjang dengan memanfaatkan sebuah timestep memori yang disebut cell [6]. Arsitektur yang dimiliki oleh LSTM memungkinkan informasi dari memori sebuah *timestep* diteruskan tidak hanya pada *timestep* berikutnya, melainkan dapat terus memberikan kontribusi pada timestep selanjutnya. LSTM didesain untuk mengatasi *vanishing gradient problem* yang terdapat pada RNN dengan memiliki *Constant Error Carousel* yang memungkinkan *error* untuk melakukan *backpropagation* tanpa terjadi *vanishing gradient* atau disebut *backpropagation through time* [11]. *Constant Error Carousel* mempertahankan aktivasi internal yang disebut sebagai (state)

dengan *recurrent connection* dengan bobot tetap yang bernilai 1.0, yang dapat diatur ulang oleh forget gate [25].

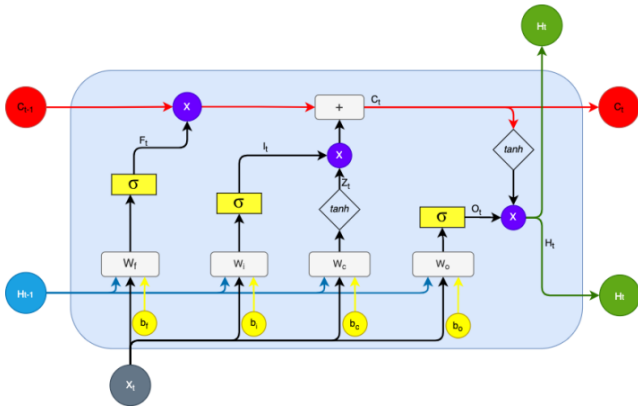


Gambar 2 Bidirectional LSTM [18]

Pada Gambar 2 merupakan gambar dengan 2 arah dari LSTM dengan menggunakan Forward Layer dan Backward Layer. Maka, pada sub bab selanjutnya akan dijelaskan secara detail mengenai *Forward* dan *Backward Layer* pada *Long Short Term Memory*.

#### 1) Forward Layer LSTM

Merupakan salah satu proses belajar dalam perhitungan LSTM yang alurnya maju. Dimana proses ini akan melakukan perhitungan LSTM yang variabel dan rumusnya sudah terdapat dalam masing-masing gate.



Gambar 3 LSTM Unit [7]

Unit LSTM terbagi menjadi 4 gate yang berinteraksi untuk menambahkan atau mengeluarkan informasi ke dalam memori setiap unit. Gambar 2.8 menjelaskan setiap timestep pada jaringan LSTM menerima 2 input, yaitu  $X_t$  timestep input dan  $H_{t-1}$  output dari timestep sebelumnya. Langkah pertama, LSTM akan memutuskan seberapa penting sebuah informasi untuk disimpan. Langkah ini dilakukan oleh layer sigmoid "forget gate" dengan Persamaan 2:

$$F_t = \sigma(W_f * H_{t-1} + U_f * X_t + b_f) \quad (2)$$

Fungsi aktivasi sigmoid akan menggunakan nilai  $H_{t-1}$  dan  $X_t$  menghasilkan nilai antara 0 dan 1, dimana hasil 1 menunjukkan bahwa nilai memori akan disimpan seluruhnya sedangkan hasil 0 dihapus selamanya [11].

$$I_t = \sigma(W_i * H_{t-1} + U_i * X_t + b_i) \quad (3)$$

$$Z_t = \sigma(W_c * H_{t-1} + U_c * X_t + b_c) \quad (4)$$

$$C_t = F_t * C_{t-1} + I_t * Z_t \quad (5)$$

Langkah berikutnya adalah melakukan perubahan terhadap nilai dari memori timestep yang terbagi menjadi 3 tahap. Pertama, input gate akan menentukan berapa besar informasi dengan melakukan fungsi aktivasi sigmoid terhadap input dari timestep yang akan dijadikan memori dengan menggunakan Persamaan 3. Tahap kedua, memory gate akan melakukan fungsi aktivasi hyperbolic tangent terhadap input timestep untuk menghasilkan nilai memori yang akan digunakan dengan Persamaan 4.

Tahap terakhir adalah untuk menghasilkan nilai memori baru yang diperoleh dari penjumlahan dari hasil perkalian forget gate (hasil dari Persamaan 2) terhadap nilai memori timestep sebelumnya (hasil dari Persamaan 3) dengan hasil perkalian input gate (hasil dari Persamaan 4) dan cell state menghasilkan nilai memori timestep saat ini yang akan digunakan untuk perhitungan selanjutnya pada Persamaan 5.

$$O_t = \sigma(W_o * H_{t-1} + U_o * X_t + b_o) \quad (6)$$

$$Z_t = O_t * \tanh(C_t) \quad (7)$$

Langkah terakhir adalah menghasilkan output  $H_t$  untuk timestep saat ini ( $t$ ). Output yang dihasilkan merupakan nilai dari *memory cell* yang telah diseleksi. Pertama, *output gate* akan menentukan jumlah informasi yang akan dijadikan output  $H_t$  dengan menggunakan fungsi aktivasi sigmoid terhadap input timestep saat ini dan output dari timestep sebelumnya (Persamaan 6). Kemudian nilai dari output gate akan dikalikan dengan hasil fungsi aktivasi hyperbolic tangent terhadap nilai memori sesuai Persamaan 7, menghasilkan output untuk timestep saat ini.

#### 2) Bidirectional Long Short Term Memory

RNN merupakan model deep learning yang baik dalam mengolah data time series karena keadaan internal RNN dapat menampilkan fitur temporal yang dinamis [27]. Pada Gambar 2 merupakan arsitektur Bidirectional LSTM adalah varian dari jaringan saraf berulang, yang memecahkan ketergantungan jangka panjang dari RNN dan LSTM. Ini menggabungkan LSTM dalam dua arah yang berbeda dan mengekstrak data informasi maju (forward) dan mundur (backward) secara bersamaan [28].

*Backward Layer* mempunyai gate dan perhitungan yang hampir sama dengan *forward*. *Backward Layer* ada untuk meningkatkan akurasi prediksi dalam machine learning. Masing-masing dari perhitungan pada *Backward Layer* akan menghasilkan nilai perubahan bobot yang terlibat di setiap perhitungan.

$$\Delta Y_t = \Delta E + W_c * \Delta Z_{t+1} + W_i * \Delta I_t + W_f * \Delta F_{t+1} + W_o * \Delta O_{t+1}$$

$$\Delta C_t = \Delta Y_t * O_t * (1 - \tanh^2(C_t)) + \Delta C_{t+1} * F_{t+1} \quad (9)$$

$$\Delta O_t = \Delta Y_t * \tanh(C_t) * \sigma(O_t) \quad (10)$$

$$\Delta F_t = \Delta C_t * C_{t-1} * (1 - \sigma^2(F_t)) \quad (11)$$

#### D. Evaluasi Forecasting

##### 1) Root Mean Square Error (RMSE)

RMSE adalah aturan penilaian kuadrat yang mengukur besarnya rata-rata kesalahan. Persamaan untuk RMSE diberikan di kedua referensi. Atau penjelasan lainnya yaitu,

RMSE adalah perbedaan antara nilai forecasting dan nilai yang diamati yang dikuadratkan dan kemudian dirata-ratakan pada sampel. Lalu, akar kuadrat dari rata-rata diambil. Karena kesalahan dikuadratkan sebelum dirata-ratakan, RMSE memberikan bobot yang relatif tinggi untuk nilai kesalahan yang besar. Ini berarti RMSE paling berguna ketika nilai kesalahan yang besar sangat tidak diinginkan.

$$RMSE = \sqrt{\frac{1}{N} * \sum_{i=1}^N ((y_p - y_i)^2)}$$

### 2) Mean Absolute Error (MAE)

MAE mengukur besarnya rata-rata kesalahan dalam saat melakukan forecasting, tanpa mempertimbangkan arahnya. MAE mengukur akurasi untuk variabel yang continue. Atau penjelasan lainnya yaitu, MAE adalah rata-rata atas sampel verifikasi dari nilai absolut dari perbedaan antara nilai forecasting dan nilai observasi. MAE adalah skor linier yang berarti bahwa semua perbedaan individu diberi bobot yang sama dalam rata-rata.

$$MAE = \frac{1}{N} * \sum_{i=1}^N |y_p - y_i|$$

### 3) Mean Absolute Percentage Error (MAPE)

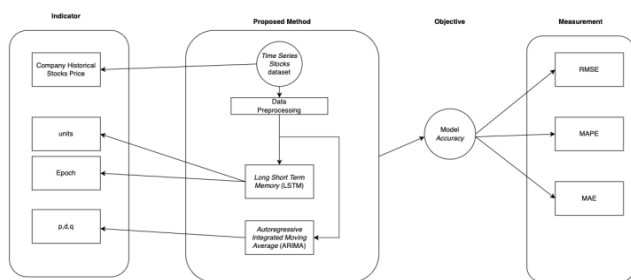
MAPE merupakan teknik pengukuran seberapa akurat sistem forecasting yang sudah dibuat [16]. Pengukuran ini mengukur akurasi sebagai persentase, dan dapat dihitung sebagai Mean Absolute Error dalam bentuk persentase untuk setiap nilai prediksi periode waktu dikurangi nilai aktual dibagi dengan nilai aktual. MAPE adalah ukuran paling umum yang digunakan untuk memperkirakan kesalahan, dan bekerja paling baik jika tidak ada nilai ekstrem pada data.

$$MAE = 100\% * \frac{1}{N} * \sum_{i=1}^N \frac{|y_p - y_i|}{y_i}$$

## III. ANALAISIS DAN PERANCANGAN SISTEM

### A. Kerangka Pemikiran

Berikut ini adalah kerangka pemikiran dari metode yang diusulkan untuk membangun model prediksi harga saham.



Gambar 4 Kerangka Pemikiran

- **Indicators** adalah variabel-variabel yang digunakan dan akan mempengaruhi hasil akhir.

- 1) **Company Historical Stocks Price** merupakan berbagai macam dataset historis harga saham perusahaan yang dari sektor FMCG, yang bersumber dari Yahoo

Finance. Rentang waktu yang digunakan dalam prediksi harga saham:

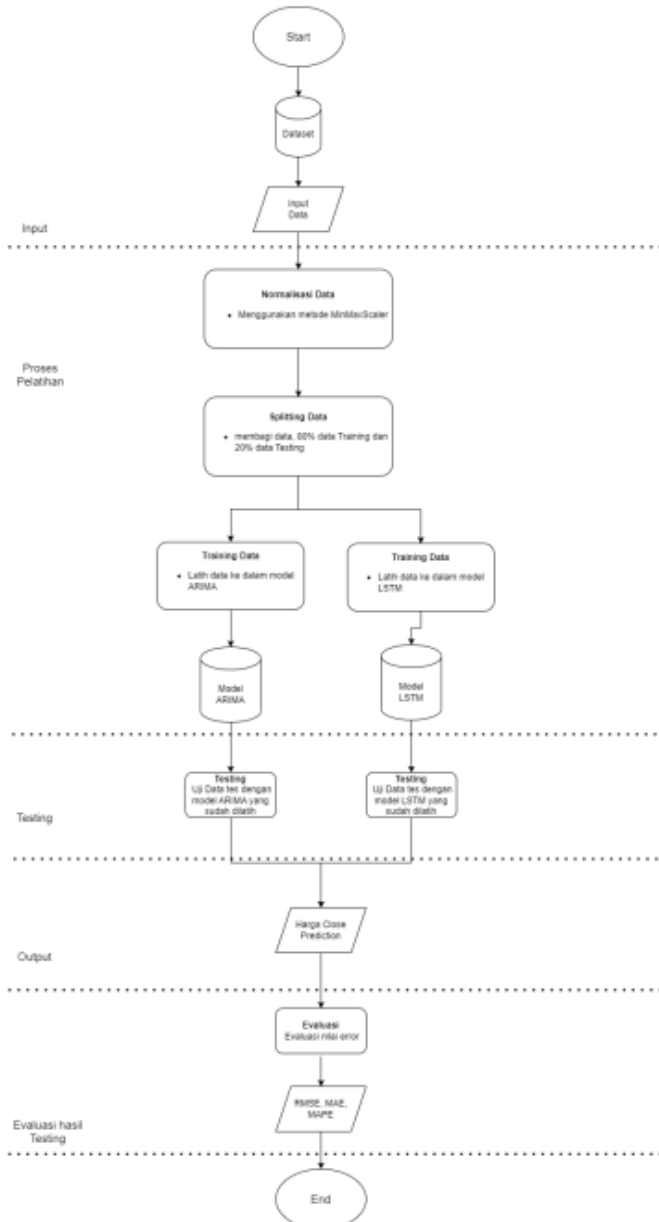
- a) **Short-term forecasting** menggunakan 3 bulan dari keseluruhan data. Rentang waktu yang digunakan dari 1 Januari 2017 sampai 31 Maret 2017, dengan jumlah 62 data.
  - b) **Mid-term forecasting** menggunakan 1 tahun dari keseluruhan data [2]. Rentang waktu yang digunakan dari 1 Januari 2017 sampai 31 Desember 2017, dengan jumlah 253 data
  - c) **Long-term forecasting** menggunakan 5 tahun dari keseluruhan data [6]. Rentang waktu yang digunakan 1 Januari 2017 sampai 31 Desember 2021, dengan jumlah 1261 data.
- 2) **Units** merupakan jumlah atau ukuran dimensi dari *hidden state* atau *output*, atau yang bisa dianggap sebagai *neuron* yang berada di *hidden layer*. Pada penelitian ini akan menggunakan jumlah sebesar 10, 50, 128.
  - 3) **Epoch** merupakan iterasi 1 siklus program selesai dijalankan. Semakin besar epoch, semakin bisa meningkatkan akurasi, namun akan semakin lama prosesnya berjalan. Epoch yang akan digunakan dalam penelitian ini adalah sebesar 10, 100, 1000.
  - 4) **Ordo p** merupakan nilai parameter untuk Autoregressive(AR), ordo d merupakan nilai parameter untuk Integrated (I), dan ordo q merupakan nilai parameter untuk Moving Average (MA). Pada penelitian ini menggunakan nilai ordo d adalah 1, dan menggunakan nilai ordo p dan q dengan rentang 1 sampai 2.

- **Proposed Method** adalah bagian yang menjelaskan proses penelitian dari awal hingga akhir. Dengan melakukan *preprocessing* dengan menghapus kolom *Volume* dan melakukan *MinMaxScaler*. Setelah itu membuat model ARIMA dan Bi-LSTM.
- **Objectives** adalah bagian yang menjelaskan acuan pengukuran. Penelitian ini menggunakan acuan performa dari model yang sudah dibuat.
- **Measurement** adalah bagian yang menjelaskan ukuran yang dipakai pada bagian objectives, Penelitian ini menggunakan RMSE, MAPE, dan MAE.

### B. Urutan Proses Global

Pada Gambar 5, dijelaskan bahwa penelitian ini diawali dengan input data lalu melakukan *preprocessing*, setelah itu dibangun model ARIMA dan LSTM. Model tersebut

dilakukan prediksi lalu dievaluasi dengan menggunakan RMSE,



Gambar 5 Urutan Proses Global

#### IV. HASIL DAN PEMBAHASAN

##### A. Hasil Pengujian Long Term Prediction

TABEL 1  
HASIL PENGUJIAN LONG TERM PREDICTION ARIMA

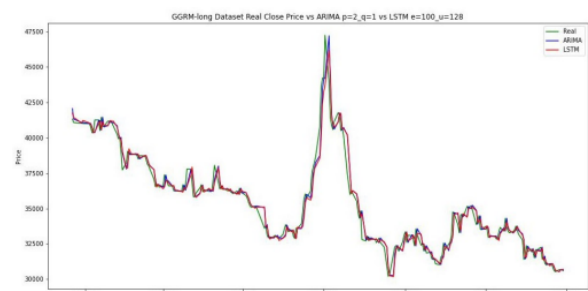
Dataset	p	d	q	RMSE	MAE	MAPE
GGRM	2	1	1	663.47	426.03	1.18
UNVR	2	1	1	110.93	82.74	1.57
PSDN	1	1	2	7.54	5.12	3.1

Pada Tabel 1 hasil dari model *Autoregressive Integrated Moving Average* untuk memprediksi harga saham dari masing-masing perusahaan dengan jangka waktu panjang. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 663.473, 426.0, 1.18. Perusahaan UNVR memiliki nilai *error* 110.93, 82.74, 1.57. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 7.54, 5.12, 3.1.

TABEL 2  
HASIL PENGUJIAN LONG TERM PREDICTION Bi-LSTM

Dataset	Epoch	Units	RMSE	MAE	MAPE
GGRM	100	128	733.23	506.21	1.41
UNVR	100	128	215.92	187.06	3.89
PSDN	1000	50	7.89	5.22	3.16

Pada Tabel 2 hasil dari model *Bidirectional Long Short Term Memory* untuk memprediksi harga saham dari masing-masing perusahaan dengan jangka waktu panjang. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 733.23, 506.21, 1.41. Perusahaan UNVR memiliki RMSE, MAE, MAPE yang paling rendah adalah 215.9, 187.06, 3.89. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 7.89, 5.22, 3.16.



Gambar 6 Perbandingan Chart data GGRM untuk Long Term

Gambar 6 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli untuk *long term prediction*. Sementara, Gambar 7 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli. Lalu, Gambar 8 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM



Gambar 7 Perbandingan Chart data UNVR untuk Long Term



Gambar 8 Perbandingan Chart data PSDN untuk Long Term

### B. Hasil Pengujian Mid Term Prediction

TABEL 3  
HASIL PENGUJIAN MID TERM PREDICTION ARIMA

Dataset	p	d	q	RMSE	MAE	MAPE
GGRM	2	1	1	1604.70	1212.40	1.6
UNVR	1	1	2	107.10	76.80	0.75
PSDN	1	1	2	29.77	18.77	6.29

Pada Tabel 3 hasil dari model *Autoregressive Integrated Moving Average* untuk memprediksi harga saham dari masing-masing perusahaan dengan jangka waktu menengah. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 1604.70, 1212.40, 1.6. Perusahaan UNVR memiliki nilai *error* 107.10, 76.80, 0.75. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 29.77, 18.77, 6.29.

TABEL 4  
HASIL PENGUJIAN LONG TERM PREDICTION Bi-LSTM

Dataset	Epoch	Units	RMSE	MAE	MAPE
GGRM	100	50	1628.43	1296.74	1.7
UNVR	1000	128	111.63	83.16	0.81
PSDN	1000	10	30.78	19.83	6.62

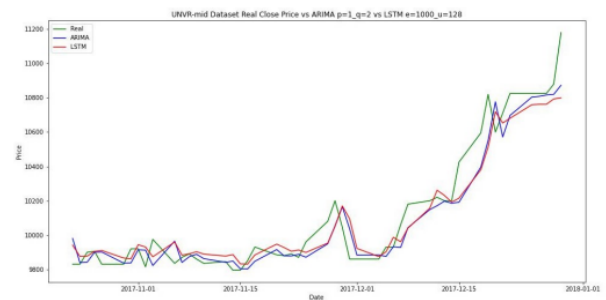
Pada Tabel 4 hasil dari model *Bidirectional Long Short Term Memory* untuk memprediksi harga saham dari

masing-masing perusahaan dengan jangka waktu menengah. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 1628.43, 1296.74, 1.7,. Perusahaan UNVR memiliki RMSE, MAE, MAPE yang paling rendah adalah 111.63, 83.16, 0.81. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 30.78, 19.83, 6.62.

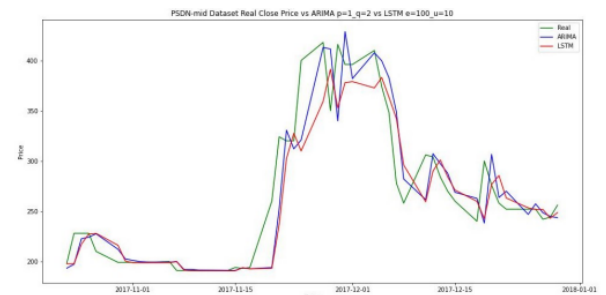


Gambar 9 Perbandingan Chart data GGRM untuk Mid Term

Gambar 9 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli untuk *Mid term prediction*. Sementara, Gambar 10 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli. Lalu, Gambar 11 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM



Gambar 10 Perbandingan Chart data UNVR untuk Mid Term



Gambar 11 Perbandingan Chart data PSDN untuk Mid Term



### C. Hasil Pengujian Short Term Prediction

TABEL 5  
HASIL PENGUJIAN SHORT TERM PREDICTION ARIMA

Dataset	p	d	q	RMSE	MAE	MAPE
GGRM	1	1	1	606.12	482.77	0.74
UNVR	1	1	2	111.80	75.07	0.87
PSDN	1	1	1	5.36	3.32	2.2

Pada Tabel 1 hasil dari model *Autoregressive Integrated Moving Average* untuk memprediksi harga saham dari masing-masing perusahaan dengan jangka waktu pendek. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 1604.70, 1212.40, 1.6. Perusahaan UNVR memiliki nilai *error* 111.80, 75.07, 0.87. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 5.36, 3.32, 2.2.

TABEL 6  
HASIL PENGUJIAN SHORT TERM PREDICTION Bi-LSTM

Dataset	Epoch	Units	RMSE	MAE	MAPE
GGRM	10	128	759.84	593.17	0.91
UNVR	1000	128	126.76	105.62	1.21
PSDN	100	128	6.4	3.75	2.49

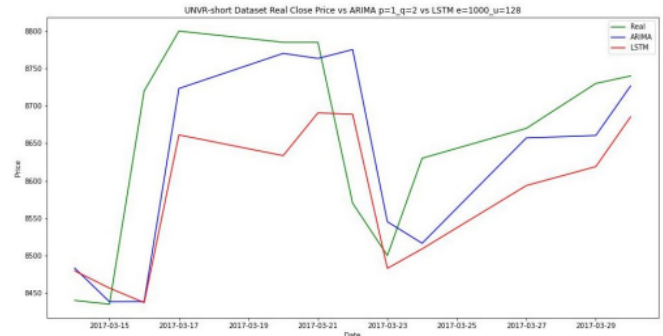
Pada Tabel 2 hasil dari model *Bidirectional Long Short Term Memory* untuk memprediksi harga saham dari masing-masing perusahaan dengan jangka waktu pendek. Perusahaan GGRM memiliki nilai *error* yang paling rendah adalah 759.84, 593.17, 0.91. Perusahaan UNVR memiliki RMSE, MAE, MAPE yang paling rendah adalah 126.76, 105.62, 1.21. Perusahaan PSDN memiliki pengukuran *error* yang paling rendah adalah 6.4, 3.75, 2.49.



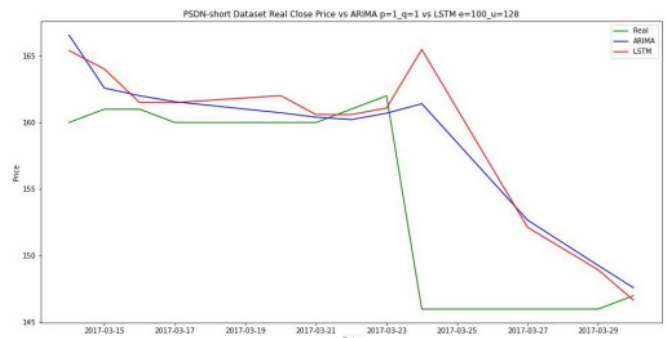
Gambar 12 Perbandingan Chart data GGRM untuk Long Term

Gambar 12 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli untuk *Short term prediction*. Sementara, Gambar 13 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM dan harga asli. Lalu, Gambar

14 memperlihatkan perbandingan dataset GGRM dari model terbaik ARIMA dengan LSTM



Gambar 13 Perbandingan Chart data UNVR untuk Short Term



Gambar 14 Perbandingan Chart data PSDN untuk Short Term

### V. SIMPULAN

Kesimpulan dari metode *Autoregressive Integrated Moving Average* dan *Bidirectional Long Short Term Memory* pada prediksi harga saham adalah:

- Nilai rata-rata RMSE, MAE, MAPE ARIMA berturut-turut ialah 363.11, 266.75, 2.06. Lalu, nilai rata-rata RMSE, MAE, MAPE LSTM ialah 621.98, 525.51, 3.83.
- Pada data saham untuk prediksi jangka waktu panjang, menengah, pendek (long, mid, short term), secara rata-rata nilai error ARIMA lebih rendah daripada rata-rata nilai error LSTM. Hal itu disebabkan LSTM menggunakan jumlah fitur yang lebih banyak tidak hanya *Close* tetapi menggunakan *Open*, *High*, *Low*, *Volume* untuk observasi dan melakukan prediksi, sementara ARIMA hanya menggunakan fitur *Close* untuk observasi dan prediksi.
- Jumlah *hyperparameter* epoch dan unit pada metode LSTM sangat berpengaruh pada prediksi harga saham. Walaupun pada saham PSDN dengan prediksi mid term dengan menggunakan jumlah unit 10 mendapatkan nilai error yang lebih rendah.
- Pada metode ARIMA, jumlah ordo p dan q tidak terlalu berpengaruh saat melakukan prediksi harga saham. Karena, banyak jumlah ordo p dan q bernilai

1 yang lebih baik daripada jumlah ordo p dan q yang bernilai 2

- Pada pengujian yang menggunakan data Blue chip pada perusahaan PT. Gudang Garam (GGRM), PT. Unilever Indonesia (UNVR) dan Gorengan pada perusahaan PT. Prasadha Aneka Niaga (PSDN) di setiap jangka waktu yang diprediksi. Saham dari perusahaan PT. Prasadha Aneka Niaga (PSDN) lebih cocok untuk melakukan prediksi dengan menggunakan metode pada penelitian ini. Disebabkan RMSE, MAE, MAPE yang didapatkan sangat jauh lebih baik daripada perusahaan Blue Chip dan harga PSDN lebih murah (ratusan) dibandingkan harga saham GGRM atau UNVR yang harganya ribuan dan puluhan ribu.

## VI. SARAN

Saran dari peneliti untuk pengembangan model prediksi harga saham di masa mendatang adalah:

- Menggunakan beberapa dataset saham dari berbagai sektor yang blue chip dan gorengan, sehingga lebih mengetahui metode apa yang lebih cocok pada setiap sektor
- Menambahkan *spektrum hyperparameter tuning* pada LSTM, supaya hasil RMSE, MAE, MAPE yang didapatkan lebih baik.
- Menerapkan metode lanjutan dari *Recurrent Neural Network* (RNN) seperti GRU atau metode *Machine Learning* lainnya.

## DAFTAR REFERENSI

- [1] M. Taufiq, W. R. Dharmawan, N. Jannah, "Peran dan Kontribusi Pasar Modal terhadap Pertumbuhan Ekonomi Indonesia," TRIANGLE : Journal of Management, Accounting, Economic and Business, vol 2, no. 4, 2021.
- [2] W. Y. Rusyida, V. Y. Pratama, "Prediksi Harga Saham Garuda Indonesia di Tengah Pandemi COVID-19 Menggunakan Metode ARIMA," SQUARE : Journal of Mathematics and Mathematics Education vol 2, no. 1, pp. 73-81, 2020.
- [3] S. A. Wadi, M. Almasarweh, A. A. Alsaraireh, "Predicting Close Price Time Series Data Using ARIMA Model," Modern Applied Science, Published by Canadian Center of Science and Education, vol. 12, no. 11, 2018.
- [4] M. Nikou, G. Mansourfar, J. Bagherzadeh, "Stock Price Prediction using Deep Learning Algorithm and its Comparison with Machine Learning Algorithm," Intelligent Systems in Accounting Finance & Management, December 2019.
- [5] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, Shahab S, "Deep Learning for Stock Market Prediction," Entropy, vol. 22, 2020.
- [6] W. Hastomo, A. S. B. Karno, N. Kalbuana, E. Nisfiani, E. T. P. Lussiana, "Optimasi Deep Learning untuk Prediksi Saham di Masa Pandemi COVID19," Jurnal Edukasi dan Penelitian Informatika, vol. 7, 2021.
- [7] J. Qiu, B. Wang, C. Zhou, "Forecasting Stock Prices with Long-Short Term Memory Neural Network Based on Attention Mechanism," Journal Plos One, 3 Januari 2020.
- [8] S. S. Namini, N. Tavakoli, A. S. Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," 17th IEEE International Conference on Machine Learning and Applications, 2018.
- [9] H. Roondiwala, H. Patel, S. Varma, "Predicting Stock Prices Using LSTM," International Journal of Science and Research, vol. 6, April 2017.
- [10] A. Geron, Hands On Machine Learning with ScikitLearn, Keras & Tensorflow, O'Reilly Media, Inc, 2019.
- [11] M. Homma. The Candle Stick Trading Bible. Independently Published by Munehisa Homma, April 2020.
- [12] A. Zhang, Z. C. Lipton, M. Li, and Alexander, Dive into Deep learning, Release 0.17.0, 2021.
- [13] F. R. Pratama, "Implementasi Algoritme Random Forest untuk Prediksi Pergerakan Harga Saham," Dept. Informatika ITHB, 2017.
- [14] J. D. Schwager. Getting Started in Technical Analysis. John Wiley & Sons, 1999.
- [15] Ian G., Yoshua B., Aaron C., Deep Learning. MIT Press, 2016
- [16] Towardsdatascience. "Forecast KPIs: RMSE, MAE, MAPE & Bias". [Online]. Available at: <https://towardsdatascience.com/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d> [Accessed: Feb. 20, 2022]
- [17] A Graves, Supervised Sequence Labelling with Recurrent Neural Networks, Illustrated Edition. Springer, 2012
- [18] D. Stifanic, J. Musulin, A. Miocevic, S. B. Segota, R. Subic. Z. Car, "Impact of COVID-19 on Forecasting Stock Prices: An Integration of Stationary Wavelet Transform and Bidirectional Long Short-Term Memory," Hindawi Complexity, 2020.
- [19] Yahoo Finance, 2022. "Historical Price Data Indonesia Equity". [Online]. Available: <https://finance.yahoo.com/> [Accessed: Jan. 31, 2022]
- [20] R. J. Hyndman, Forecasting: Principles & Practice, Illustrated Edition. Springer, 2014
- [21] MBA Knowledge Base, 2021. Time Horizon in Forecasting. [Online]. Available: <https://www.mbaknol.com/managerial-economics/time-horizon-in-forecasting/> [Accessed: March. 28, 2022]
- [22] M. Pradana. Saham Blue Chip dan Saham Gorengan . Available: <https://investbro.id/saham-gorengan/> [Accessed: Apr. 12, 2022]
- [23] F Bre, J. M. Gimenez, V. D. Fachinotti. Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks. [Online]. Available: researchgate [Accessed: Feb. 22, 2022]
- [24] Geeksforgeeks, 2019. Implement sigmoid function using Numpy. [Online]. Available: <https://www.geeksforgeeks.org/implement-sigmoid-function-using-numpy/> [Accessed: Feb. 24, 2022]
- [25] X. Wang, Y. Liu, C. Sun, B. Wang and X. Wang, "Predicting Polarities of Tweets by Composing Word Embeddings with Long Short-Term Memory," Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics, pages 1343–1353, Beijing, China, July 26-31, 2015.
- [26] Towardsdatascience, 2019. The Most Intuitive and Easiest Guide for Recurrent Neural Network. [Online]. Available: <https://towardsdatascience.com/recurrent-neural-networks-a-guide-for-beginners/> [Accessed: Mar. 24, 2022]
- [27] Q. Chen, W. Zhang, Y. Lou, "Forecasting Stock Prices Using a Hybrid Deep Learning Model Integrating Attention Mechanism Multi-Later Perceptron, and Bidirectional Long-Short Term Memory Neural Network," School of Information Management and Artificial Intelligence, Zhejiang University of Finance and Economics, China, Beijing, China, vol. 8, 2020.
- [28] M. Jia, J. Huang, L. Pang, Q. Zhao, "Analysis and Research on Stock Price of LSTM and Bidirectional LSTM Neural Network," Advances in Computer Science Research, vol. 90, 2019
- [29] Tensorflow Keras, 1997. Long Short Term Memory. [Online]. Available: <https://keras.io/lstm/> [Accessed: Des. 09, 2022]