Deep Learning Methods In Predicting Indonesia Composite Stock Price Index (IHSG)

Arief Fadhlurrahman Rasyid, Dewi Agushinta R.*, Dharma Tintri Ediraras
Magister of Information System Management, Information System, Economy
Gunadarma University
Depok, West Java, Indonesia
**Corresponding author's email: dewier [AT] staff.gunadarma.ac.id*

Abstract— **The stock price changes at any time within seconds. The stock price is a time series data. Thus, it is necessary to have the best analysis model in predicting the stock price to make decisions to avoid losses in investing. In this research, the method used two models Deep Learning namely Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) in predicting Indonesia Composite Stock Price Index (IHSG). The dataset used is historical data from the Jakarta Composite Index (^JKSE) stock price in 2013-2020 obtained through Yahoo Finance. The results suggest that Deep learning methods with LSTM and GRU models can predict Indonesia Composite Stock Price Index (IHSG). Based on the test results obtained RMSE value of 71.28959454502723 with an accuracy rate of 92.39% for LSTM models and obtained RMSE value of 70.61870739073838 with an accuracy rate of 96.77% on GRU models.**

Keywords- Composite Stock Price Index (IHSG); Deep learning; GRU; Jakarta Composite Index (^JKSE); LSTM; predicting; stock price

I. INTRODUCTION

Indonesia's investment and trading are growing quite rapidly, investing in stocks, forex or commodities is very tempting because we can get a considerable profit. Tens, even hundreds of percent in a short period (a day, a week, a month, depending on the conditions) we can get. The profit obtained can also be several times the profit of deposits. However, if there is a miscalculation, it can bankrupt investors in a short time as well.

In businessinsider.com, a stock is an investment that represents a unit of ownership in a company. Stocks are securities that prove an investor's ownership of a company. This is interpreted if a person buys stocks of a company. The number of stocks purchased has given capital to the company. The Company has the option to issue stocks to help fund the company. Nevertheless, stocks are a means of investment that has a high interest among investors because stocks can provide an attractive level of profit. Every investor who invests in a company usually has the purpose to get capital gain, namely the difference between the purchase price and the selling price of stocks and dividends in the form of company profits given to investors [1].

Stocks are the easiest asset to liquidate, when we need cash, we can get it quickly. Due to the momentary valuation by both buyers and sellers influenced by several factors, the stock price changes at any time within seconds.

Indonesia Stock Exchange (IDX) is one of the best-growing stock markets. The Indonesia Stock Exchange (IDX) is the capital market in Indonesia that has existed long before the Independence of Indonesia, and the first stock exchange in Indonesia was established in 1912. The Composite Stock Price Index (IHSG) is a tool to measure the performance of the Indonesian stock market.

Many factors can influence the movement of the Composite Stock Price Index (IHSG), one of which is macroeconomic and global economic conditions. Macroeconomic conditions may affect the movement of IHSG, such as the occurrence of inflation and exchange rates. Meanwhile, the global economy that can affect IHSG is the world gold price and the world oil price [2].

The stock price is a time-series data that has a large dimension and is not static, so developed an approach for non-linear data [3].

Popular theories indicate that stock markets are primarily a chance step, with certain rules on the preceding day's closing price. The majority of traditional time-series forecasting techniques are based on stationary trends. As a result, stock price forecasting is fraught with difficulty. Furthermore, due to the large number of variables involved, predicting stock prices is a difficult task in and of itself. The market behaves like a voting machine in the near term, but like a weighing machine in the end, therefore there is potential for forecasting market movements over a longer timescale [4].

The volatility of the stock market is complex and nonlinear. Relying only on a trader's personal experience and intuition for analysis and judgment is unreliable and inefficient. To lead stock trading, people require an intelligent, scientific, and successful study strategy. The application of deep learning in predicting stock prices has become a research hotspot with the rapid growth of artificial intelligence. Because of its good non-linear approximation abilities and adaptive self-learning, the neural network in deep learning has become a popular predictor [5].

Deep learning is known as the latest learning algorithm that has reliable performance and can build optimal models in data processing [6]. Deep learning is a subset of machine learning that employs multiple layers as a non-linear information-processing medium. Long Short Term Memory (LSTM), a development of the Recurrent Neural Network (RNN) algorithm, is one type of deep learning algorithm. LSTM solves the main problem of RNN by not processing sequential information over a long period, but rather processing data based on a time series [7].

Unlike RNN, LSTM is a cell with a series of gates (4 gates) that are quite complicated. With the large number of gates and network complexity in LSTM, of course, it will also affect the duration of the data processing process. Therefore, a model of the Gated Recurrent Unit (GRU) was developed. GRU is a cell containing only two gates with a simpler circuit compared to LSTM. GRU's capabilities are better than LSTM especially for small datasets [6].

The results of time-series data prediction of Bank BRI stocks by using LSTM learning machine seems successful enough to get a high accuracy value (RMSE = 227.470333244533), and visually the plotting results of the prediction data show a graph similar to the actual data [8].

LSTM obtained the results of forecasting predicted that the stock price of SSMS will increase up to the price of IDR 966.57 per stock on July 25, 2019. Similarly, the stock prices of LSIP and SIMP are expected to increase by IDR 1,138.12 per stock and IDR 358.87 per stock on July 25, 2019, respectively [9].

The LSTM method in time series prediction is very appropriate to use because LSTM can remember and store data history either short term or long term [10].

Thus, it is necessary to have the best analysis model in predicting the stock price to make decisions to avoid losses in investing (Indonesia as a case). In this research, the method used is Deep Learning by using two models namely LSTM and GRU in predicting the Composite Stock Price Index (IHSG). The dataset used is historical data from the Jakarta Composite Index (^JKSE) stock price or better known as the Composite Stock Price Index (IHSG) in 2013-2020 obtained through Yahoo Finance.

II. Literature Review

A. Stocks

Stocks are the capital of a person or business entity in a company where the party has a claim on the company's income, assets, and the party has the right to attend the General Meeting of Shareholders (GMS) [11].

In general, stock indices gained from high-market investment stock prices and often provide an estimate of economic status in individual countries. Findings show, for instance, that stock market capitalization has a positive impact on economic growth in countries. Investments risk for investors, the nature of the stock values movement is ambiguous. Furthermore, the market status for governments is usually difficult to detect. Indeed, stock values are frequently dynamic, non-parametric, and non-linear, resulting in poor statistical model performance and the inability to anticipate precise values and movements [4].

Stocks are the easiest assets to liquidate; we can get them quickly if we need cash.

B. Stock Price

The stock price is the present value of income to receive by investors in the future. The stock price shows the achievements of companies that move in the direction of the company's performance. Companies that have excellent achievements can improve their company performance that reflected in the company's financial statements so investors will be interested in investing in the company. An increase in investor demand for the company will cause the price of the company's shares to tend to increase as well. The company's financial statements can be a reference for investors in making investment decisions, such as selling, buying, or investing in shares. The stock price can change at any time, depending on the amount of supply, and demand of investors [12].

C. Composite Stock Price Index (IHSG)

Indonesia Stock Exchange (IDX) is one of the best-growing stock markets that has existed long before the Independence of Indonesia, established in 1912. The Composite Stock Price Index (IHSG) is a tool used to measure the performance of the IDX.

Several things can affect the movement of the Composite Stock Price Index (IHSG), one of which is the macroeconomic and global economic conditions. Macroeconomic conditions may affect the movement of IHSG, such as the occurrence of inflation and exchange rates. Meanwhile, the global economy that can affect IHSG is the world gold price and the world oil price [2].

D. Deep Learning

Deep learning is the most recent learning technique that can develop optimal models in data processing and has a high level of reliability. Long Short Term Memory (LSTM) model, which is a derivation of Recurrent Neural Network (RNN) algorithm is in the Deep learning algorithms. LSTM solves the main problem of RNN by processing data based on a time series, not processing sequential information over a long period [7].

In contrast to RNN, LSTM is a cell with a gate arrangement (4 gates) that is quite complicated. With the large number of gates and network complexity in LSTM, of course, it will also affect the duration of the data processing process.

Therefore, a model was developed, namely Gated Recurrent Unit (GRU). GRU is a cell containing only two gates with a simpler arrangement compared to LSTM. GRU's capabilities are better than LSTM especially for small datasets [6].

1) Long Short Term Memory (LSTM): a special type of RNN that works better in practice, due to the update of the equation and the presence of backpropagation dynamics in it. Long Short Term Memory network (LSTM) architecture, as seen in Fig. 1.

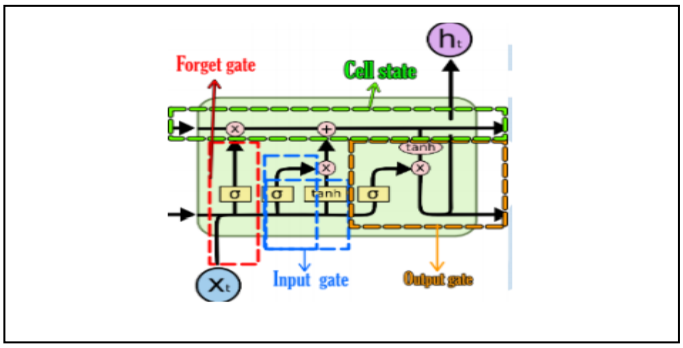


Figure 1. Unit of LSTM.

The architecture has a gate that serves to delete and add information that is [13]:

a) Forget Gate: conditionally decide whether or not information should be removed from the processing

b) Input Gate: conditionally to determine an input will be added to the cell state memory at that time or not [14].

c) Cell state: this serves as a memory or memory for a layer. The value of a cell state is manipulated using a gateway system.

d) Output Gate: conditionally decide what will be generated based on input and block memory.

Equations formed from forget gate, input gate, output gate, tanh layer create a new candidate value vector, cell state, and hidden layer are:

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

$$i_t = \sigma\ (W_i\ .[h_{t-1}\ ,x_t\ ] + b_i) \qquad (2)$$

$$\tilde{C}_t = \tanh(W_c\ .[h_{t-1}\ ,x_t\ ] + b_c) \qquad (3)$$

$$C_t = f_t * C_{t-1} + i_t * C_t \qquad (4)$$

$$O_t = \sigma\ (W_o\ .[h_{t-1}\ ,x_t\ ] + b_o) \qquad (5)$$

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

Where:

f_t = forget gate.

i_t= input gate.

O_t= output gate.

C_t= cell state on timestamp (t).

$\tilde{C}_t$= candidate for cell state on timestamp (t).

h_t= the final output of the LSTM block at timestamp (t).

h_{t-1}= output from the previous LSTM block at timestamp (t-1).

W_x= weight for each neuron gate(x).

b_x= bias for each gate(x).

2) Gated Recurrent Unit (GRU): a variation of the simpler LSTM. GRU is a carefully designed repeat structure that makes a good trade-off between performance and speed. Therefore, GRU has been commonly used in academic and industrial fields. GRU was first introduced [15]. The main purpose of loading GRU is to create each recurrent unit to be able to capture dependencies in different timescales adaptively. As an analogy, humans do not have to use all the information in the past to be able to make decisions now. In GRU, the information flow regulatory component is referred to as gate and GRU has two gates, gate reset, and gate update. A gate reset on the GRU will determine how to combine the new input with the past information, and the gate update will determine how much past information should remain stored. The architecture of the GRU can be seen in Fig. 2.

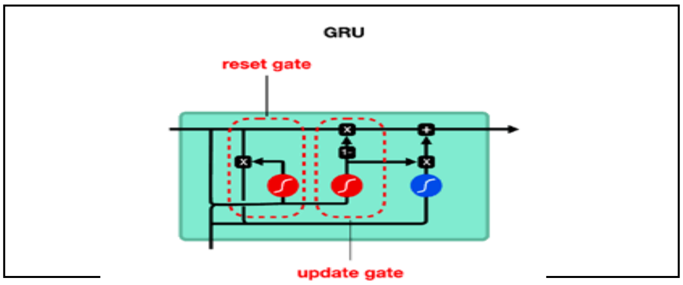


Figure 2. GRU architecture.

The following formulas of Gated Recurrent Unit (GRU) are [15]:

$$z_t = \sigma\ (W_z \cdot [h_{t-1}, x_t]) \qquad (7)$$

$$r_t = \sigma\ (W_r \cdot [h_{t-1}, x_t]) \qquad (8)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \qquad (9)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \qquad (10)$$

Where:

z_t= update gate.

r_t= reset gate.

$\tilde{h}_t$= activation candidate.

h_t= vector output.

x_t= vector input.

E. Optimizer

Optimizer aims to find the optimal weight, minimize errors and maximize accuracy. During the training process, the model parameters (weights) changed to minimize the loss function, to be able to predict as accurately as possible. In simpler terms, optimizers form the model that belongs to it into the most accurate form by utilizing its weight [16]. The various optimizers are as follows [8]:

1) Adam (Adaptive Moment Estimation): a method that calculates the adaptive learning rate for each parameter. The recommended parameter values are $\beta 1 = 0.9$, $\beta 2 = 0.999$, and $\varepsilon = 10\text{-}8$. With $\beta 1 = \beta 2$ = exponential decrease rate and ε = epsilon value for parameter update.

2) Adamx: a type of Adam that has infinite norms. The recommended parameter values are $\eta = 0.002$, $\beta 1 = 0.9$, and $\beta 2 = 0.999$. With $\beta 1 = \beta 2$ = exponential rate of decline and η = learning rate with bias-correction.

3) RMSprop: a method that maintains the moving average of the square gradient for each weight. The recommended parameter value is $\eta = 0.001$ with η = learning rate with bias-correction.

F. Deep Learning with Keras

Keras is a deep learning framework for Python that provides an easy way to define and train almost any type of deep learning model. Keras originally developed for researchers, to allow experiments to run faster. Today, Keras is Kaggle's most popular framework and the world's deep learning competition. This framework supports two of the most popular data science languages, R and Python. The advantages offered are the emergence of features such as making code run smoothly on the CPU or GPU, a user-friendly API that makes it easy to prototype deep learning, and supports RNN models [17].

G. TensorFlow Execution Model

TensorFlow represent all computation and state in a machine learning algorithm using a single dataflow graph, including mathematical operations, parameters and their update rules, and input preprocessing. The dataflow graph explicitly expresses the communication between sub computations, making it simple to run independent computations in parallel and partition computations across multiple devices. TensorFlow is distinct from batch dataflow systems in two ways [18]:

- Multiple concurrent executions on overlapping subgraphs of the overall graph are supported by the model.
- Individual vertices may have a mutable state that can be stored between graph executions.

III. Research Methods

There are 6 stages of the research method, consisting of the data collection stage, preprocessing data, then initializing the parameters, going to the training stage, followed by the testing stage, and in the end, will get the best prediction results. The research flow is illustrated in Fig. 3.

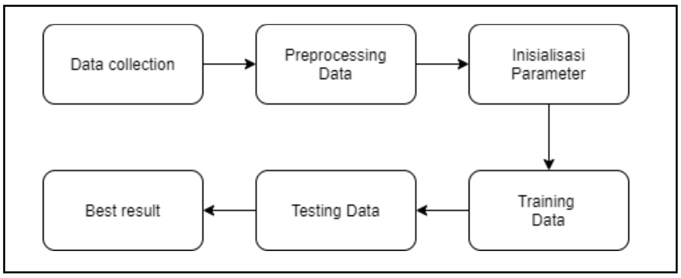


Figure 3. Research Stage.

This research was created using Python programming language with the help of several modules including pandas, NumPy, matplotlib, TensorFlow, and Keras. The stages of the research will explain as follows:

A. Data Collection

The data used in this research is retrieved from Jakarta Composite Index (^JKSE) stock data or better known as Composite Stock Price Index (IHSG) which can be accessed online from https://finance.yahoo.com/ website. Datasets used from January 1, 2013, to December 28, 2020, with 1941 total data. This dataset consists of six column data features, namely *High, Low, Open, Close, Volume, Adj. Close*. The *Close* feature will be used as a dataset for the next process.

The *Close* feature or closing price was chosen because it is a useful marker for investors to use in assessing stock price changes over time. The closing price on a day compared to the closing price of the previous day, or 30 days earlier or a year earlier, to measure changes in market sentiment against the stock.

Jakarta Composite Index (^JKSE) stock data or better known as Composite Stock Price Index (IHSG) can be seen in Table I.

TABLE I. JAKARTA COMPOSITE INDEX (^JKSE) STOCK DATA

Date	Open	High	Low	Close	Adj.Close	Volume
2013-01-02	4,322.58	4,364.67	4,316.42	4,346.48	4,346.48	32,794,300
2013-01-03	4,356.41	4,401.33	4,356.41	4,399.26	4,399.26	43,267,700
2013-01-04	4,406.93	4,418.48	4,389.72	4,410.02	4,410.02	36,418,100
2013-01-07	4,407.73	4,427.65	4,387.49	4,392.38	4,392.38	40,179,400
2013-01-08	4,394.76	4,413.94	4,381.65	4,397.54	4,397.54	48,360,000
…..	…..	…..	…..	…..	…..	…..
2020-12-21	6,131.62	6,195.15	6,119.92	6,165.62	6,165.62	238,612,700
2020-12-22	6,146.65	6,174.46	6,010.21	6,023.29	6,023.29	250,085,200
2020-12-23	6,061.85	6,104.35	5,853.26	6,008.71	6,008.71	203,528,500
2020-12-28	6,067.00	6,095.01	5,979.90	6,093.55	6,093.55	180,855,700
2020-12-29	6,112.72	6,143.87	6,026.71	6,036.17	6,036.17	175,135,000

The *Close* feature or closing price used as a dataset for the next process is shown in Fig. 4.

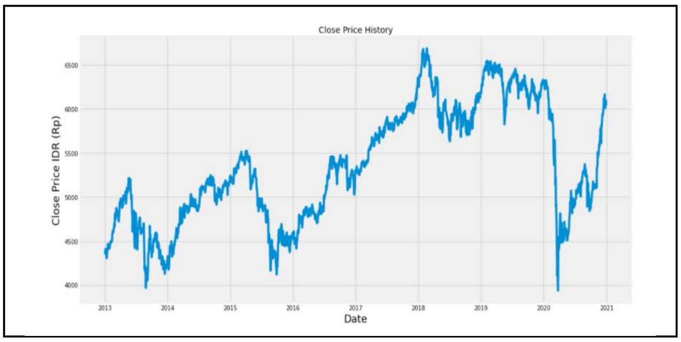


Figure 4. *Close* feature of Jakarta Composite Index (^JKSE) stock.

B. *Preprocessing Data*

1) Splitting Data: the dataset will be divided into training and testing data. Training data amounting to 80% of the dataset used for the training process. After obtaining the best model from the training data, then the model is tested on the test data amounting to 20% of the remaining dataset to see if the resulting model will have high accuracy or not.

2) Data Normalization: to reduce errors, normalize the dataset by converting the raw data to a value in the [0, 1] interval. The min-max scaling approach was employed to normalize the data. The normalization formula for min-max scaling is as follows [19]:

$$x' = \frac{x-\min(x)}{\max(x)-\min(x)} \qquad (11)$$

Where:

x' : normalization results.

x : data normalized.

$\min(x)$: the minimum value of the entire data.

$\max(x)$: the maximum value of the entire data.

C. *Initialization of Parameters*

After preprocessing, the data is then used to determine the initialization of basic parameters needed, such as:

- Number of hidden layers
- Number of neuron on the hidden layer
- Mean Square Error (MSE) target error value
- Maximum Epoch.

D. *Training*

1) The training process in LSTM conducted as follows.

a) Count all function gate units on each neuron: By sorting the function of the gate to be calculated are forgotten gates, input and cell gates function, and output gates function.

b) The training has repeated as many epochs as it has specified: If not, Adam's optimization will be used to update the system’s weight and bias.

2) The training process in GRU is as follows.

a) In GRU: we first calculate the C̃t context candidate as the tanh of the old context combination Ct-1 and input xt. Then a sigmoid gate called gate update ut determines how much of the new context Ct comes from the candidate C̃t and how much of the old context Ct-1.

b) For complete GRU: computing added one more gate that is rt called reset gate or relevance gate to calculate C̃t.

c) The training has repeated as many epochs as it has specified: If not, Adam’s optimization will be utilized to update the system’s weight and bias.

E. *Testing*

With accuracy, methods used using MSE, RMSE, and MAPE, the model obtained in the training process tested using data testing obtained from preprocessing data.

1) Mean Square Error (MSE): the sum of squared errors or differences between the actual value and the predicted value, then divides the summation result by the amount of forecasting time. MSE can see small error changes. The equation is as follows [7]:

$$MSE = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n} \qquad (12)$$

Where:

$\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n$ = predicted value.

$Y_1, Y_2, \ldots, Y_n$ = observed value.

n = amount of data observed.

2) Root Mean Square Error (RMSE): the sum of the squares of errors or the difference between the actual value and the predictive value, and divides the summation result by the amount of forecasting time, and draws the root. The RMSE value is understood to know how far the error value of a model is against the linear regression line [20]. RMSE was able to muffle a major change. The equation is [7]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}} \qquad (13)$$

Where:

$\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n$ = predicted value.

$Y_1, Y_2, \ldots, Y_n$ = observed value.

n = amount of data observed.

3) Mean Absolute Percentage Error (MAPE): the absolute value of the percentage of data errors against the mean. The equation is as follows [21]:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| * 100 \qquad (14)$$

Where:

A_t= actual value.

F_t= forecast value.

n= sample size.

t= time-step.

IV. RESULT AND DISCUSSION

A descriptive analysis of Jakarta Composite Index (^JKSE) stock data or better known as the Indonesia Composite Stock Price Index (IHSG) was conducted to find out an overview of the Composite Stock Price Index (IHSG) from January 01, 2013, to December 28, 2020.

Based on the Jakarta Composite Index (^JKSE) stock data in Fig. 5, it can be seen that the price movement of Jakarta Composite Index (^JKSE) stocks tends to have an uptrend or increase. The highest jump was in 2018 that reached prices of more than 6,500. However, a drastic decrease in 2020 reaches the price of 4,500. The highest price of Jakarta Composite Index (^JKSE) shares was at the beginning of 2018 because it reached 6,605. While the lowest price was 4,195 in mid-2013.

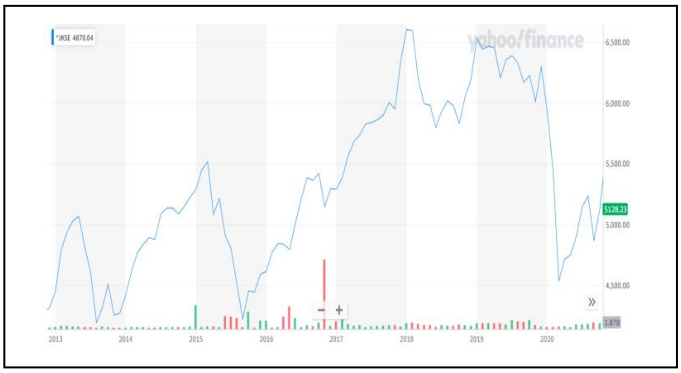


Figure 5. Jakarta Composite Index stock price chart (^JKSE).

A. *Data Training and Data Testing Sharing*

Research data is separated into training data and testing data, with an 80% training data ratio of up to 1553 and a 20% testing data ratio of up to 389, as seen in Table II.

TABLE II. DATA TRAINING AND DATA TESTING

Variable	Data Training	Data Testing	Total
JKSE	1553	388	1941

The use of more data training is done so that the learning machine or learning algorithms are better trained with data patterns from data training. The training data trained using LSTM and GRU methods. The training will produce a model. This model with the test data performs against data testing. This process continues until it gets the model with the best accuracy value. The model will be used for the prediction process if the best model is completed.

B. *Normalization*

At this stage, the dataset will be converted to a value with a range interval of [0, 1] using min-max scaling. The formula for calculating min-max scaling is found in (11). The following is the normalization result indicated by Table III.

TABLE III. NORMALIZATION RESULTS (X')

No.	x'
1.	0.14858077
2.	0.16776294
3.	0.17167411
…	…
1939.	0.7526659
1940.	0.78350015
1941.	0.76264711

C. *Parameter Initialization*

After preprocessing, the data then dictates how the basic settings should be set up. Here are the required parameters on the LSTM and GRU models.

Initialization of parameters in both LSTM and GRU models consists of one input layer with 25 neurons, 2 hidden layers with 50 neurons in each, and 1 output layer with one neuron, as in Table IV.

TABLE IV. MODEL PARAMETERS

Parameters	Amount
Total layers	4 layers
Input layer	1 layer (25 neurons)
Hidden layer	2 layers (50 neurons)
Output layer	1 layer (1 neuron)

D. *Determining the Neurons and Epochs Number*

Determining the neurons optimal number on hidden layers and epochs is an important task in LSTM and GRU. Networks with multiple neurons on the hidden layer may not be able to capture the complex relationship between the target variable and the input. However, if too many neurons assigned in a hidden part of the network it will create poor predictive capabilities from invisible data due to over parameterization. Therefore, several neurons in a hidden layer should be examined, and architectural performance should be assessed after each run with several neurons in a hidden layer.

In this research, the experiment employed a network with one input layer, two hidden layers, and one output layer with the number of neurons on the hidden layer being 10, 20, 30, 40, and 50. 50, 100, and 150 will be utilized as epochs. The least loss value, where the loss value is the MSE, can be used to determine the exact number of neurons and epoch. As well as using Adam's optimization. Adam optimization is used because it uses bias correction techniques. There are no rules in determining the neurons and epochs number so they obtained through experiments until it achieves the highest level of accuracy in predicting time series. Table V shows the result of a test with a certain number of neurons and epoch.

TABLE V. TEST RESULTS WITH DIFFERENT NUMBER OF NEURONS AND EPOCHS

Number of neurons	Number of epochs	MSE
10	50	3.4009e-04
	100	3.3210e-04
	150	3.9101e-04
20	50	6.8051e-04
	100	3.5997e-04
	150	3.4002e-04
30	50	4.6234e-04
	100	3.5139e-04
	150	3.4863e-04
40	50	5.1377e-04
	100	3.2982e-04
	150	3.7939e-04
50	50	5.0785e-04
	100	3.5148e-04
	150	3.6484e-04

Table V showed that the number of neurons 40 and 100 epoch (iterations) produced the smallest MSE value of 3.2982e-04 compared to other neurons and epochs. The excessively small number of epochs results in a very general web, meaning that the network's ability to recognize patterns is too little or nothing. While too many epochs can result in the network experiencing overfit conditions or the network is too specific to the training data.

E. Data Training

Data training was carried out by measuring the MSE value for each epoch after getting the maximum number of neurons and epochs. Fig. 6 shows some of the training.

F. Data Prediction

After obtaining the model in the training process, then the model tested using the test data obtained in the distribution of training data and data testing, with the accuracy method used using RMSE to see whether or not the results of the model formed.

In this research, data predictions using 2 models, LSTM and GRU. The following are the predicted results of the data tested with both models.

1) LSTM model data prediction results: The following are the predicted results of the data using the LSTM model can be seen from the comparison graph of prediction data and actual data shown in Fig. 6.

We can see the formed LSTM model can produce the appropriate output in Fig. 7. This can be seen from the pattern of prediction data that follows the pattern formed by the actual data. Actual data is represented by blue and prediction data is represented by red.

```
Epoch 1/100

47/47 [==============================] - 5s 42ms/step - loss:
0.0577

Epoch 2/100

47/47 [==============================] - 2s 43ms/step - loss:
0.0018

Epoch 3/100

47/47 [==============================] - 2s 43ms/step - loss:
0.0016

Epoch 4/100

47/47 [==============================] - 2s 43ms/step - loss:
0.0015

Epoch 5/100

. . .

Epoch 97/100

47/47 [==============================] - 2s 45ms/step - loss:
3.8037e-04

Epoch 98/100

47/47 [==============================] - 2s 44ms/step - loss:
3.5645e-04

Epoch 99/100

47/47 [==============================] - 2s 45ms/step - loss:
3.8875e-04

Epoch 100/100

47/47 [==============================] - 2s 44ms/step - loss:
3.7830e-04

<tensorflow.python.keras.callbacks.History at 0x7f5f4a7d8190>
```

Figure 6. Data Training.

Based on the testing of the model, a fairly good result was obtained with an RMSE value of 71.28959454502723.

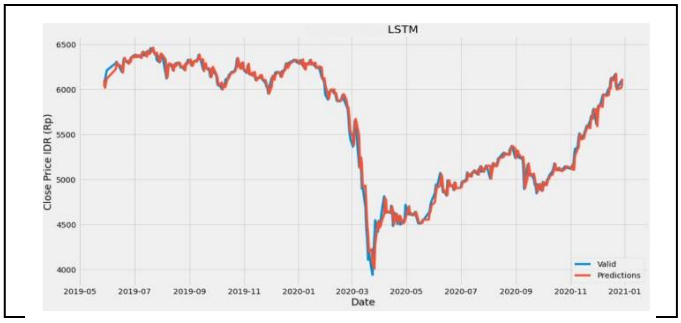


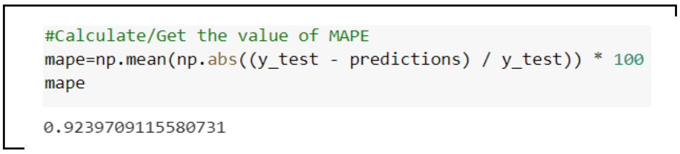
Figure 7. Comparison graphs of LSTM prediction data and actual data.

Table VI will be displayed the closing price prediction or *Close* feature on Jakarta Composite Index (^JKSE) shares with the LSTM model.

TABLE VI. JKSE SHARE PRICE PREDICTION WITH LSTM MODEL

Period	Actual Price	Prediction Price
2019-05-28	6033.142090	6097.926270
2019-05-29	6104.105957	6016.520996
2019-05-31	6209.117188	6117.272949
2019-06-10	6289.609863	6221.883301
2019-06-11	6305.992188	6289.487793
...	...	...
2020-12-21	6165.624023	6096.337891
2020-12-22	6023.289062	6172.948242
2020-12-23	6008.708984	5995.841797
2020-12-28	6093.554199	6017.269531
2020-12-29	6036.173828	6116.111816

Table VI showed that the results of the predicted prices are close to the actual prices. Hence, it shows that this LSTM model can predict the closing price or the *Close* feature on the Jakarta Composite Index (^JKSE). Mean Absolute Percentage Error (MAPE) measured the results of the trained LSTM model. MAPE values resulting from prediction are in Fig. 8.



```
#Calculate/Get the value of MAPE
mape=np.mean(np.abs((y_test - predictions) / y_test)) * 100
mape

0.9239709115580731
```

Figure 8. MAPE value of LSTM model.

From the MAPE value, it showed that the predicted results using the LSTM model have an accuracy value of 92.39%.

2) GRU model data prediction: the result of predicting data using the GRU model can be seen from the comparison graph of prediction data and actual data as in Fig. 9.

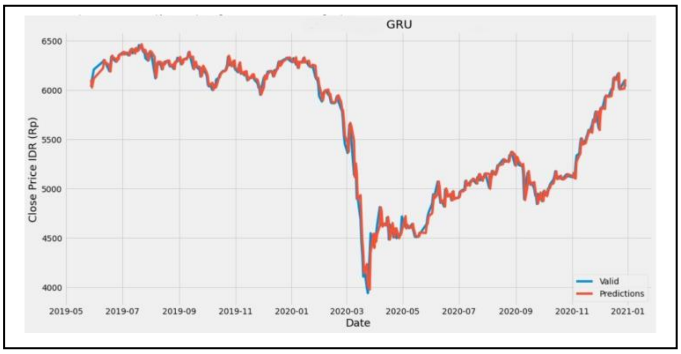


Figure 9. Comparison graphs of prediction data and actual GRU data.

Fig. 9 showed that the formed GRU model could produce the appropriate output. The actual data formed the pattern of prediction data. Blue lines represent the actual data and red lines represent the prediction data as well. Based on the model testing, a good result with an RMSE value of 70.61870739073838.

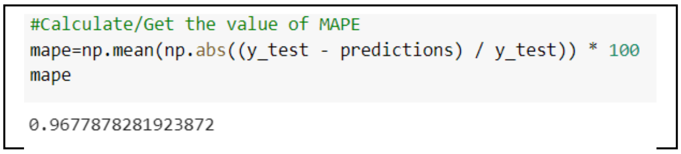
Table VII displayed closing price prediction or *Close* feature on Jakarta Composite Index (^JKSE) stock with GRU model.

TABLE VII. JKSE STOCK PRICE PREDICTION WITH GRU MODEL

Period	Actual Price	Prediction Price
2019-05-28	6033.142090	6104.620117
2019-05-29	6104.105957	6026.433105
2019-05-31	6209.117188	6117.701660
2019-06-10	6289.609863	6221.676758
2019-06-11	6305.992188	6294.505859
...	...	...
2020-12-21	6165.624023	6102.475098
2020-12-22	6023.289062	6173.863770
2020-12-23	6008.708984	6006.526855
2020-12-28	6093.554199	6016.301270
2020-12-29	6036.173828	6110.760254

Table VII showed that the results of the predicted prices are close to the actual prices. Therefore, it shows that this GRU model can predict the closing price or the *Close* feature on the Jakarta Composite Index (^JKSE).

Mean Absolute Percentage Error (MAPE) measured the results of the GRU model trained. The MAPE value resulting from the prediction is in Fig. 10.



```
#Calculate/Get the value of MAPE
mape=np.mean(np.abs((y_test - predictions) / y_test)) * 100
mape

0.9677878281923872
```

Figure 10. MAPE value of GRU model.

From the MAPE value, it showed that the predicted results using the GRU model have an accuracy value of 96.77%.

V. CONCLUSION

Based on the results of determining the optimal number of neurons in the hidden layer and epoch on the LSTM and GRU models in predicting the Composite Stock Price Index (IHSG), it showed that the optimal number of neurons 40 and 100 epoch (iteration) produces the smallest MSE value of 3.2982e-04.

Deep learning methods with LSTM and GRU models can predict the Indonesia Composite Stock Price Index (IHSG). The best IHSG prediction result is to use the GRU model because it has RMSE value and better accuracy compared to the LSTM model.

This research could be a method reference chosen for the development of other research and enhance to produce accurate methods, especially in predicting the Composite Stock Price Index (IHSG). This research entailed having the best analysis model to predict stock prices to help investors in making decisions to avoid a loss in investment. Further extend research is to use other optimizations e.g. AdaMax, RMSprop, etc. so that more accurate predictions are obtained.

References

[1] L. P. Putri, "Pengaruh Profitabilitas Terhadap Harga Saham Pada Perusahaan Pertambangan Batubara di Indonesia," Jurnal Ilmiah Manajemen dan Bisnis, 16(2), 2015, pp. 49-59.

[2] V. Harfikawati, "Pengaruh Tingkat Inflasi Nilai Tukar Rupiah Terhadap USD, dan Indeks Dow Jones Terhadap IHSG Di Bursa Efek Indonesia Tahun 2011-2015," Journal Eksekutif, 13(2), 2016.

[3] L. E. Siahaan, Prediksi Indeks Harga Saham dengan Metode Gabungan Support Vector Regression dan Jaringan Syaraf Tiruan. Indonesia Journal on Computing (Indo-JC), 2(1), 2017, pp. 21-30.

[4] Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., & Salwana, E. "Deep learning for stock market prediction," Entropy, 22(8), 2020, p. 840.

[5] J. Qiu, B. Wang, and C. Zhou, "Forecasting stock prices with long-short term memory neural network based on attention mechanism." PLoS ONE 15(1): e0227222. 2020, https://doi.org/10.1371/journal.pone.0227222.

[6] L. Deng, and D. Yu, "Deep Learning: Methods and Applications," Foundations and Trends in Signal Processing, 7(3–4), 2014, pp. 197–387.

[7] Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, 9(8), 1997, pp. 1735–1780.

[8] A. S. B. Karno, and W. Hastomo, "Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long Short Term Memory)," Journal of Informatic and Information Security, 1(1), 2020, pp. 1-8.

[9] R. Yotenka, and F. F. El Huda, "Implementasi Long Short-Term Memory Pada Harga Saham Perusahaan Perkebunan Di Indonesia," Unisda Journal of Mathematics and Computer Science (UJMC), 6(01), 2020, pp. 9-18.

[10] S. Zahara, and M. B. Ilmiddafiq, "Prediksi Indeks Harga Konsumen Menggunakan Metode Long Short Term Memory (LSTM) Berbasis Cloud Computing," Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 3(3), 2019, pp. 357-363.

[11] E. Mahpudin, and N. A. Mahmud, "Analysis of The Contribution of Tax Increases In Changes of Stock Prices In Indonesia and The United States," Journal of Critical Reviews, 7(11), 2020, pp. 2194-2197.

[12] M. R. Hutauruk, and I. Ghozali, "Overview of return on investment on cigarette companies registered in Indonesia stock exchange," International Journal of Scientific and Technology Research, 9(03), 2020, pp. 4633-4637.

[13] J. Brownlee, "Develop Your First Neural Network in Python With Keras Step-By-Step," Machine Learning Mastery, 2016.

[14] L. Zaman, S. Sumpeno, and M. Hariadi, "Analisis Kinerja LSTM dan GRU sebagai Model Generatif untuk Tari Remo," Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI), 8(2), 2019, pp. 142-150.

[15] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in NIPS 2014 Workshop on Deep Learning, December 2014.

[16] N. Donges, "Towards Data Science," Gradient Descent in a Nutshell. https://towardsdatascience.com/gradient-descent-in-a-nutshell eaf8c18212f0, 2018.

[17] F. Chollet, and J. J. Allaire, Deep Learning with R, Vol. 1, Manning Publications Co. NY, 2018.

[18] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean,... and X. Zheng, "Tensorflow: A system for large-scale machine learning," in Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation (OSDI 16), 2016, pp. 265-283.

[19] M. Mada, A. Farmadi, I. Budiman, M. R. Faisal, and M. I. Mazdadi, "GRU, AdaGrad, RMSprop, Adam Implementation of GRU and Adam Optimization Method for Stock Price Prediction," Journal of Data Science and Software Engineering, 2(01), 2021, pp. 36-45.

[20] L. J. Tashman, Out-of-sample tests of forecasting accuracy: an analysis and review. International journal of forecasting, 16(4), 2000, pp. 437-450.

[21] A. De Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean absolute percentage error for regression models," Neurocomputing, 192, 2016, pp. 38-48.