International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Comparative Study of LSTM Algorithm Performance (Long Short Term Memory and CNN (Convolutional Neural Network) in Making Machine Learning Model in Google Inc Share

Mulia Estuari[1], Dr. Lintang Yuniar Banowosari[2]
[1]Master of Information Systems Management Department, Business Information System, Gunadarma University, Depok, West Java, Indonesia-16424
[2]Master of Information Systems Management Department, Business Information System, Gunadarma University, Depok, West Java, Indonesia-16424
Email Address: [1]muliaestuari @ gmail.com, [2]lintang @ staff.gunadarma.ac.id

Abstract— *Stocks are one of the tools used in the capital market which is very attractive to investors because they have attractive advantages. Shares can be defined as a sign or proof of someone's ownership in a company or limited liability company. The general characteristic of stock prices is that they have a degree of uncertainty. This uncertainty is in terms of stock price movements in the short term, or in the long term. This characteristic is not liked by investors because it poses a risk to their investment. This uncertainty is also unavoidable in investing. To deal with uncertainty in price movements, what investors can do is reduce this uncertainty. One of the tools / tools to reduce uncertainty is to predict stock prices. To predict future stock prices, a method that is capable of predicting stock prices is needed. With the existence of various prediction methods and the development of prediction methods with historical data that are fast enough so that there are many choices of methods that can be used in predicting stock movement patterns as needed, it is necessary to compare methods. One another so that the prediction results are obtained with high accuracy.*

Keywords— *Stock predictions,machine learning.*

I. INTRODUCTION

An Stocks are one of the tools used in the capital market which is very attractive to investors because they have attractive advantages. Shares can be defined as a sign or proof of someone's ownership in a company or limited liability company. The shape and form of the shares itself is a sheet of paper which states that the holder or owner of the company is the owner of the company that issues the securities itself. The proportion of ownership is determined by how much investment is planted in the company (Eka Patriya, 2020, Journal of Scientific Technology and Engineering Volume 25 No.1 April 2020).

Stocks have become a popular instrument in the Indonesian capital market. This is indicated by the growth of funds of 34% per year in Indonesia (Kurnia, 2005, p. 1). In terms of mutual fund customers, in September 2007 there were 249,670 customers (Capital Market Supervisory Agency, 2007, p. 32). Equity funds attracted more investors than other types of mutual funds with the highest percentage of investors, namely 3.80% or as many as 86,883 investors of the total number of mutual fund investors in the Indonesian capital market (Capital Market Supervisory Agency, 2007, p. 32).

The general characteristic of stock prices is that they have a degree of uncertainty. This uncertainty is in terms of stock price movements in the short term, or in the long term. This characteristic is not liked by investors because it poses a risk to their investment. This uncertainty is also unavoidable in investing. To deal with uncertainty in price movements, what investors can do is reduce this uncertainty. One of the tools / tools to reduce uncertainty is to predict stock prices (Van den Goorbergh, 1999, p. 8).

The stock price prediction process is divided into two groups. The first group is those who believe that there is a way to predict stock prices. The second group is those who believe that the market is efficient, and if new information is obtained, the market will absorb it and correct itself. This second group believes that stock prices are unpredictable. This is called the theory of the Efficient Market Hypothesis (EMH), as mentioned by (Gryc, 2006, p. 4). This group assumes that the stock market follows a random walk pattern, which means that the best predictions can be obtained about the future price of a shares are based on the current share price (Gryc, 2006, p. 4).

In this modern era, there are several methods for predicting stock prices. One method is machine learning, using a sample (example) of stock prices and trying to see patterns that are formed from these stock price movements using an algorithm that can study patterns. The algorithm produces an approximation function that relates the input variables to the output variables. The input variable here can be the historical stock price (Lawrence, 1997, p. 3)

Stock market predictions are important and very interesting because successful stock price predictions can promise attractive returns. To predict future stock prices, a model that can predict stock prices is needed using stock price data stored in a database or data warehouse several years ago. To create an architectural model to be able to see stock movement patterns, a method is needed, namely using machine learning, using a sample (example) of stock prices and trying to see patterns that are formed from these stock

Mulia Estuari and Dr. Lintang Yuniar Banowosari, "Comparative Study of LSTM Algorithm Performance (Long Short Term Memory and CNN (Convolutional Neural Network) in Making Machine Learning Model in Google Inc Share," *International Research Journal of Advanced Engineering and Science*, Volume 6, Issue 2, pp. 34-36, 2021.

price movements using an algorithm that can study patterns. The algorithm produces an approximation function that relates the input variables to the output variables. The input variable here can be its historical share price. With the existence of various forecasting methods and the development of forecasting methods with historical data that are quite rapid, there are many choices of methods that can be used. In predicting stock movement patterns as needed, it is necessary to compare one method with another in order to obtain forecast results with high accuracy.

II. Literature Review

A. Understanding Machine Learning

Machine learning is an application from the discipline of artificial intelligence (Artificial Intelligence) that uses statistical techniques to generate an automatic model from a set of data, with the aim of giving computers the ability to "learn". Machine learning or machine learning allows the computer to learn a number of data (learn from data) so that it can produce a model to carry out the input-output process without using explicit program code. The learning process uses special algorithms called machine learning algorithms. There are many machine learning algorithms with different efficiency and case specifications.

B. Long Short Term Memory Recurrent Neural Network

Long short term memory recurrent neural network or LSTM RNN is a neural network development capable of studying long-term dependencies. Introduced by Hochreiter and Schmidhuber in 1997, the RNN LSTM was further refined and popularized until now, the RNN LSTM can work very well on a wide variety of problems. There are several LSTM unit architectures. The general architecture consists of the forget gate, input gate, output gate. An example of the LSTM RNN architecture can be seen in figure 1.

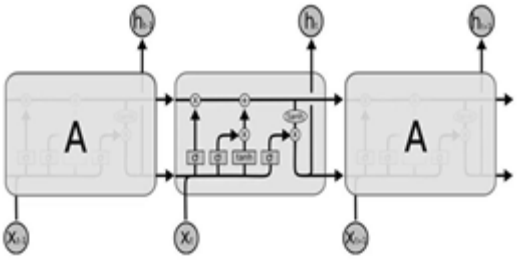

Fig. 1. LSTM architecture

C. Convolutional Neural Network

Convolutional Neural Network is a machine learning method from the development of Multi Layer Perceptron (MLP) which is designed to process two-dimensional data. CNN is included in the type of Deep Neural Network because of its deep network level and is widely implemented in image data. CNN has two methods; namely the classification using feedforward and the learning stage using backpropagation. The way CNN works is similar to MLP, but in CNN, each neuron is presented in two dimensions, unlike MLP, where each neuron is only one dimension.

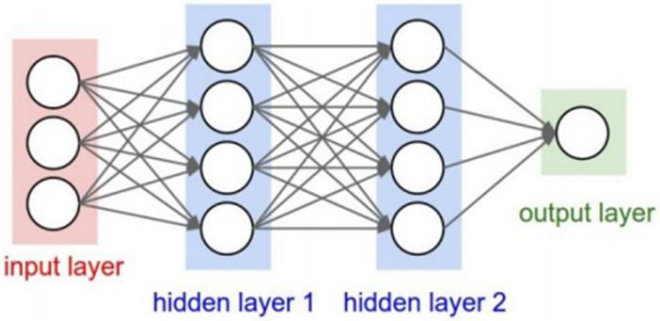

Fig. 2. MLP Architecture

An MLP as in figure 2. It has i layers (red and blue squares) with each layer containing ji neurons (white circles). MLP accepts one-dimensional data input and propagates the data on the network to produce output. Each link between neurons in two adjacent layers has a one-dimensional weight parameter that determines the quality of the mode. In each input data layer, a linear operation is carried out with the existing weight values, then the computation results are transformed using a non-linear operation called the activation function.

III. Research Method

A. Research Steps

The research stages are carried out so that the research can be completed scientifically. With the stages of research, it will make it easier for researchers to solve the problems at hand. The following are the steps taken can be seen below:

1. Installing Google Colab
2. Import csv and library data to Google Colab
3. Create the results of the train dataset and test dataset
4. Added MinMaxScaler
5. Make predictions
6. Doing Reshaping
7. Imports Keras or Tensor Flow libraries
8. Viewing Network Design
9. Run LSTM to the training set
10. Predicting Stock Prices
11. Prediction result visualization

IV. Research Result And Discussion

A. Comparison visualization of prediction results and real data using long short term memory

```
Epoch 1/10
38/38 [==============================] - 5s 125ms/step - loss: 0.0393 - accuracy: 0.0017
Epoch 2/10
38/38 [==============================] - 5s 126ms/step - loss: 0.0067 - accuracy: 0.0017
Epoch 3/10
38/38 [==============================] - 5s 128ms/step - loss: 0.0056 - accuracy: 0.0017
Epoch 4/10
38/38 [==============================] - 5s 128ms/step - loss: 0.0058 - accuracy: 0.0017
Epoch 5/10
38/38 [==============================] - 5s 127ms/step - loss: 0.0051 - accuracy: 0.0017
Epoch 6/10
38/38 [==============================] - 5s 127ms/step - loss: 0.0050 - accuracy: 0.0017
Epoch 7/10
38/38 [==============================] - 5s 128ms/step - loss: 0.0046 - accuracy: 0.0017
Epoch 8/10
38/38 [==============================] - 5s 127ms/step - loss: 0.0039 - accuracy: 0.0017
Epoch 9/10
38/38 [==============================] - 5s 132ms/step - loss: 0.0043 - accuracy: 0.0017
Epoch 10/10
38/38 [==============================] - 5s 135ms/step - loss: 0.0042 - accuracy: 0.0017
<tensorflow.python.keras.callbacks.History at 0x7f4a64b32828>
```

Fig. 3. Long short term memory accuracy results

Mulia Estuari and Dr. Lintang Yuniar Banowosari, "Comparative Study of LSTM Algorithm Performance (Long Short Term Memory and CNN (Convolutional Neural Network) in Making Machine Learning Model in Google Inc Share," *International Research Journal of Advanced Engineering and Science*, Volume 6, Issue 2, pp. 34-36, 2021.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

In figure 3 is the stage where the long short term memory begins to learn from the training set (X_train and y_train), here we only need to wait a few moments until the iteration display is complete. If the iteration is complete, we can find out the accuracy results of the Lomg short term memory algorithm.

The following in figure 4 is the process of plotting the prediction results. When executed, it looks like this:

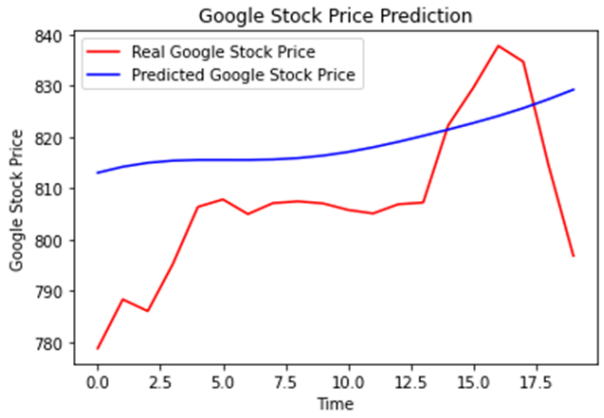

Fig. 4. Long short term memory prediction results

B. *Visualization of comparison of prediction results and actual data using a convolutional neural network*

```
Epoch 1/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0746 - accuracy: 8.3472e-04
Epoch 2/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 8.3472e-04
Epoch 3/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 0.0000e+00
Epoch 4/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 8.3472e-04
Epoch 5/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 8.3472e-04
Epoch 6/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 0.0017
Epoch 7/10
38/38 [==============================] - 2s 44ms/step - loss: 0.0744 - accuracy: 0.0000e+00
Epoch 8/10
38/38 [==============================] - 2s 43ms/step - loss: 0.0744 - accuracy: 8.3472e-04
Epoch 9/10
38/38 [==============================] - 2s 43ms/step - loss: 0.0744 - accuracy: 8.3472e-04
Epoch 10/10
38/38 [==============================] - 2s 43ms/step - loss: 0.0744 - accuracy: 8.3472e-04
<tensorflow.python.keras.callbacks.History at 0x7fd205a5bfd0>
```

Fig. 5. Convolutional neural network accuracy results

In figure 5 are the stages where the convolutional neural network begins to learn from the training set (X_train and y_train), here we only need to wait a while until the iteration display is complete. If the iteration is complete, we can find out the accuracy results of the convolutional neural network algorithm.

The following in figure 6 is the process of plotting the prediction results. When executed, it looks like this :

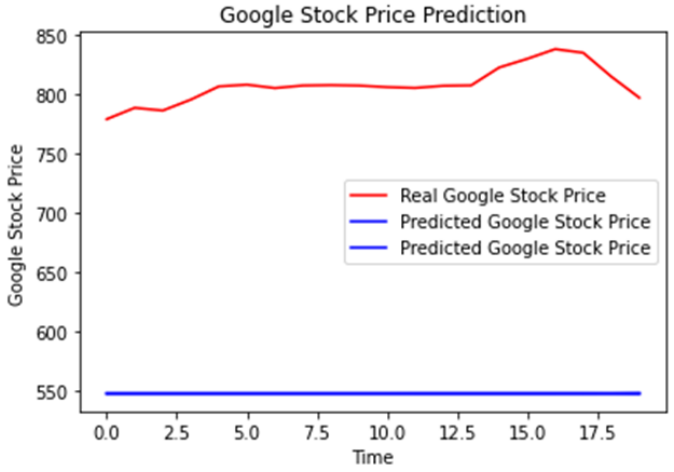

Fig. 6.

V. CONCLUSION

A. *Hypothesis Test Results*

Based on the research results, it can be concluded that the long short term memory algorithm is better for predicting stock prices than the convolutional neural network algorithm model. Because based on the graphical pattern and data from the LSTM (Long short term memory) algorithm the accuracy is 0.0017, while for the data from the CNN (Convolutional neural network) algorithm research the accuracy is 8.3472e-04. Because the accuracy of the LSTM algorithm is close to 1, it can be said that the LSTM algorithm is better than CNN. And seen from the visual image, the LSTM graphic pattern almost matches the original prediction data. Although there is still a gap (difference) between the actual price and the predicted price, visually it is good enough.

B. *Suggestion*

This model can still be improved and done by adding an LSTM layer, modifying existing algorithms, or using other activation functions, and much more.

REFERENCES

[1] https://journal.uii.ac.id
[2] https://colab.research.google.com
[3] https://indoml.com
[4] https://www.kajianpustaka.com
[5] https://subrutin.com
[6] https://www.datacamp.com
[7] https://www.kaggle.com

Mulia Estuari and Dr. Lintang Yuniar Banowosari, "Comparative Study of LSTM Algorithm Performance (Long Short Term Memory and CNN (Convolutional Neural Network) in Making Machine Learning Model in Google Inc Share," *International Research Journal of Advanced Engineering and Science*, Volume 6, Issue 2, pp. 34-36, 2021.