

Gold Price Prediction in COVID-19 Era

Seng Hansun and Alethea Suryadibrata

Informatics Department, Universitas Multimedia Nusantara, Tangerang,
Indonesia, seng.hansun@lecturer.umn.ac.id

Date of Submission: 15th April 2019 Revised: 27th June 2021 Accepted: 18th July 2021

How to Cite: Hansun, S. and Suryadibrata, A. (2021). Gold Price Prediction in COVID-19 Era. *International Journal of Computational Intelligence in Control* 12(2), pp.1-4.

Abstract - As one of the most frequently traded commodities in the world, gold has been hugely impacted by the COVID-19 crisis. In this study, we try to apply a famous Deep Learning method for time series analysis, namely the Long Short-Term Memory (LSTM) networks, for future gold price prediction. However, rather than using a complex network architecture, we propose simple three layers LSTM networks that were trained on 4,219 training records and tested on 1,055 test records. We found that the Root Mean Square Error (RMSE) value for the prediction results is 39.94162, while the Mean Absolute Percentage Error (MAPE) value is 17.66144. Moreover, the R^2 score of the prediction results could reach 97.242%, which is considered high and comparable with other more complex networks' architectures available in the literature.

Index Terms—COVID-19, Deep Learning, Gold price, LSTM, Prediction.

INTRODUCTION

The 2019 novel coronavirus disease, known as COVID-19, has happened for almost two years now. Its first case was reported in late 2019 [1], [2] and later it was declared as a pandemic by the World Health Organization (WHO) within the first semester of its occurrence [3]. It has affected many aspects of our life, from political and socio-economic aspects to life-threatening medical situations that are faced by all survivors of this disease. As of 22 May 2021, more than 166 million people have been confirmed to get this disease, with more than 3.4 million deaths worldwide [4].

One example of this pandemic effect on the economic aspect is the gold price fluctuation. As one of the most frequently traded commodities in the world, the gold price movement has a significant impact on the world economy and the financial markets [5]. Throughout the COVID-19 pandemic era, the gold price volatility has attracted many researchers to learn its characteristics and nature associated with the disease. Syahri and Robiyanto[6], for example, they

had studied the correlation of gold, foreign exchange rate, and the Indonesia Composite Stock Price Index (CSPI/IHSG) during the COVID-19 pandemic period. They found that the gold price movement has significantly affected the stock price volatility, during this pandemic era. Moreover, as can be seen in other studies by Atriet *al.* [7] and Yousef and Shehadeh [8], the gold price seems to have a positive correlation with the COVID-19 global new infections. Therefore, unlike other assets such as metal and oil, gold can be considered as a truly safe haven asset, especially in this COVID-19 era[9], [10].

In this study, our purpose is to predict the future price of gold as a safe haven asset in this pandemic era. Several studies have been conducted related to the gold price prediction, some using the statistical and stochastic approach, others using complex machine learning methods. In Sami and Junejo[11] work, they applied the Artificial Neural Networks (ANN) and Logistic Regression (LR) in predicting future gold rates by using 22 market variables. Livieriset *al.* [12] and Vidal and Kristjanpoller[13] also had tried to predict the gold price movement by using a popular Deep Learning method known as the Long Short-Term Memory (LSTM) networks. However, both studies were conducted before the COVID-19 crisis occurred with complex LSTM architectures. We propose to use a relatively simpler and smaller architecture of LSTM networks, yet could give an accurate prediction result of the gold price.

In the following section, we will explain the main Deep Learning method, namely the LSTM networks applied in this study. Some prediction error criteria as the performance evaluation metrics will also be given in the same section. Then in the next section, the Deep Learning model development and other pre-processing steps will be briefly described. The prediction results then will be discussed later, followed by some concluding remarks in the last section of this paper.

RESEARCH METHOD

The fundamental concept of the Long Short-Term Memory (LSTM) networks applied in this study will be first explained in this section. Two prediction error criteria will be briefly explained next.

I. LSTM Networks

Hochreiter and Schmidhuber had introduced LSTM networks to handle the limitation found in the conventional Recurrent Neural Networks (RNN) method. It can learn both short and long terms interdependency of states in the networks through its gate mechanism. There are three different gates in an LSTM cell, namely the Input, the Output, and the Forget gates [14]. Figure 1 illustrates an LSTM cell with its gate mechanism [15].

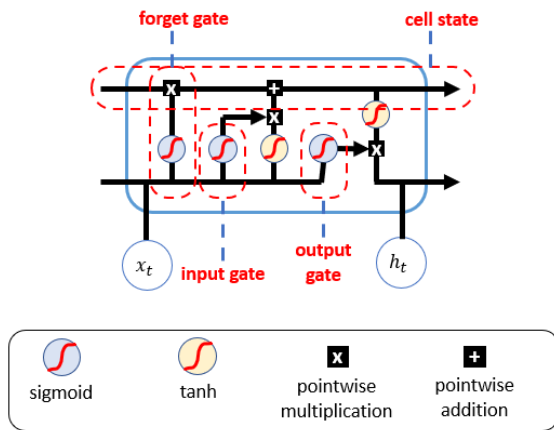


FIGURE 1
AN LSTM CELL DIAGRAM

The gate mechanism is used to control how much information can be passed throughout the LSTM networks. There are six formal equations related to this gating mechanism in an LSTM cell as given below.

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (1)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \quad (3)$$

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

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

where f_t is the forget gate at the current cell, i_t is the input gate at the current cell, o_t is the output gate value at the current cell, C_t is the current cell state, \tilde{C}_t is the candidate (next) cell state, $W_f, W_i, W_c, W_o, U_f, U_i, U_c, U_o$ are the

weights of the networks, b_f, b_i, b_c, b_o is the bias variable values, h_t is the current hidden state value, h_{t-1} is the prior hidden state value, and x_t is the new input value.

II. RMSE and MAPE

Two popular prediction error criteria, namely the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE), are used as the performance evaluation metrics in this study. If RMSE gives error degrees in a unit value, MAPE gives the error degrees in a percentage value. Both of them can be formulized as [14], [16], [17]

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \quad (7)$$

$$MAPE = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \right) \cdot 100\% \quad (8)$$

where n is the total number of data, Y_t is the actual value, and F_t is the forecasted value.

PRE-PROCESSING AND MODEL DEVELOPMENT

In this study, the recorded daily gold price was taken from Yahoo! Finance [18] on 15 May 2021. The maximum data available in the data source was collected, resulting in a total of 5,274 records. Although there are several attributes available in the dataset, only the Closing data attribute will be considered in this study for the gold price prediction. To handle any missing values in the dataset, we conducted a simple data imputation technique. A total of 112 missing records was found and then replaced with their last known records. Next, in the data splitting phase, we divided the dataset into training and test sets with an 80:20 ratio, resulting in a total of 4,219 records for the training set and a total of 1,055 records for the test set.

In the data normalization phase, we processed both training and test sets with a simple feature scaling method, which was then reshaped into a 3D array that can be accepted by the LSTM model in Keras. Keras is a Deep Learning package for Python which runs on top of the TensorFlow platform. Rather than using a complex network architecture, in this study, we propose simple three layers LSTM networks. The network consists of an LSTM layer with 100 neurons, a Dropout layer that drops 20% of the processed information in the networks to prevent overfitting, and a Dense layer with one neuron. Moreover, in compiling the model, we used the Mean Square Error (MSE) and Adam optimizer. The model will be trained on the training set for 20 epochs with a batch size of 32 each. Figure 2 shows the snipped code for the model development process in the Python programming language.

```
# BUILD the LSTM networks
# We use 1 LSTM block with 100 nodes

# Initialising the RNN-LSTM
model = Sequential()

# Adding the LSTM Layer and some Dropout regularisation
model.add(LSTM(100, return_sequences = True, input_shape=(x_train.shape[1], 1)))
model.add(Dropout(0.2))

# Adding the output layer
# For Full connection layer we use dense
# As the output is 1D so we use unit=1
model.add(Dense(1))

# Compiling the RNN-LSTM
model.compile(loss='mean_squared_error', optimizer='adam')

# Fitting the RNN to the Training set
history = model.fit(x_train, y_train, epochs=20, batch_size=32, validation_data=(x_test, y_test),
                    callbacks=[EarlyStopping(monitor='val_loss', patience=10)], verbose=1, shuffle=False)

# Show the model summary
model.summary()
```

FIGURE 2
MODEL DEVELOPMENT WITH LSTM NETWORKS

The learned model will be used to predict the Closing gold price on both the training and test sets. However, we need to invert the predicted results into the original scaling of the data first. Then, the predicted results can be found and plotted as described in the next section.

RESULTS AND ANALYSIS

After successfully built the model using the proposed three layers LSTM networks, we predicted the gold price results

on both the training and test sets. The actual and predicted closing prices are plotted in blue and red lines shown in Figure 3. As can be perceived from Figure 3, the proposed LSTM networks could predict the gold price movement very well. Moreover, the loss function results recorded in the model development on the training set are shown in Figure 4. It can be seen that the model has converged well and remained stable after 12 epochs.



FIGURE 3
GOLD PRICE PREDICTION RESULTS

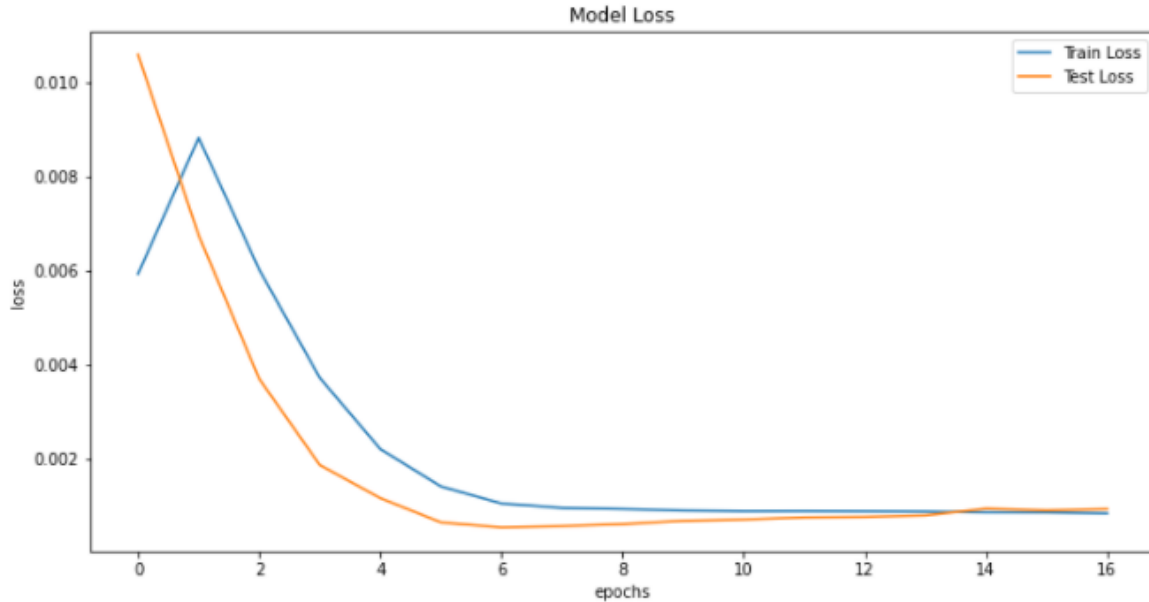


FIGURE 4
GOLD PRICE LOSS FUNCTION PLOTS

Furthermore, we also evaluated the prediction error results by using RMSE and MAPE criteria. The RMSE value on the test set is 39.94162, while the MAPE value is 17.66144. If we are using the typical interpretation of MAPE scores as stated by Moreno et al. [19], then the prediction results can be considered as a good forecasting result. Lastly, we also calculated the R-squared (R^2) score for the predicted results. R^2 is a commonly used criterion to measure a model's accuracy for a regression task in the Machine Learning domain. The R^2 score for the gold price prediction using the proposed LSTM networks could achieve 97.242% which is considered high and comparable with other more complex network architectures.

CONCLUSION

Gold, as one of the most frequently traded commodities in the world, has been hugely impacted by the COVID-19 pandemic. We tried to apply a famous Deep Learning method, i.e., the Long Short-Term Memory (LSTM) networks, to build a model that can predict the gold price movement in the future. By using simple three layers LSTM networks, we found a Deep Learning model that could achieve a prediction accuracy as high as 97.242%. However, there are also some limitations found. In this study, we do not focus on the comparison of several prediction methods for gold price prediction. Other Deep Learning methods, such as Artificial Neural Networks, Support Vector Regressors, and Bi-directional LSTM, can be applied shortly. Moreover, we also do not try to assess the gold's property as a safe haven, especially in this pandemic era, which can be answered in the future study.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and corrections. The authors would also like to extend our gratitude to Universitas Multimedia Nusantara for the support given during this study.

REFERENCES

- [1] Acter, T., Uddin, N., Das, J., Akhter, A., Choudhury, T.R., and Kim, S., 2020, "Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: A global health emergency," *Sci. Total Environ.*, 730, pp. 138996, doi: 10.1016/j.scitotenv.2020.138996.
- [2] El-Aziz, T.M. and Stockand, J.D., 2020, "Recent progress and challenges in drug development against COVID-19 coronavirus (SARS-CoV-2) - an update on the status," *Infect. Genet. Evol.*, 83, pp. 104327, doi: 10.1016/j.meegid.2020.104327.
- [3] Spinelli, A. and Pellino, G., 2020, "COVID-19 pandemic: perspectives on an unfolding crisis," *Br. J. Surg.*, 107(7), pp. 785–787, doi: 10.1002/bjs.11627.
- [4] Johns Hopkins University & Medicine, 2021, "Coronavirus Resource Center". <https://coronavirus.jhu.edu/> (accessed May 22, 2021).
- [5] Gharib, C., Mefteh-Wali, S., and Ben Jabeur, S., 2021, "The Bubble Contagion Effect of COVID-19 Outbreak: Evidence from Crude Oil and Gold Markets," *Financ. Res. Lett.*, 38, pp. 101703, doi: 10.1016/j.frl.2020.101703.
- [6] Syahri, A. and Robiyanto, R., 2020, "The Correlation of Gold, Exchange Rate, and Stock Market on Covid-19 Pandemic Period," *J. Keuang. dan Perbank.*, 24(3), pp. 350–362, doi: 10.26905/jkdp.v24i3.4621.
- [7] Atri, H., Kouki, S., and Gallali, M., 2021, "The Impact of COVID-19 News, Panic and Media Coverage on the Oil and Gold Prices: An ARDL Approach," *Resour. Policy*, 72, pp. 102061, doi: 10.1016/j.resourpol.2021.102061.
- [8] Yousef, I. and Shehadeh, E., 2020, "The Impact of COVID-19 on Gold Price Volatility," *Int. J. Econ. Bus. Adm.*, VIII(4), pp. 353–364, doi: 10.35808/ijeba/592.

- [9] Salisu, A.A., Raheem, I.D., and Vo, X.V., 2021, "Assessing the Safe Haven Property of the Gold Market During COVID-19 Pandemic," *Int. Rev. Financ. Anal.*, 74, pp. 101666, doi: 10.1016/j.irfa.2021.101666.
- [10] Dutta, A., Das, D., Jana, R.K., and Vo, X.V., 2020, "COVID-19 and Oil Market Crash: Revisiting the Safe Haven Property of Gold and Bitcoin," *Resour. Policy*, 69, pp. 101816, doi: 10.1016/j.resourpol.2020.101816.
- [11] Sami, I. and Junejo, K.N., 2017, "Predicting Future Gold Rates using Machine Learning Approach," *Int. J. Adv. Comput. Sci. Appl.*, 8(12), pp. 92–99, doi: 10.14569/IJACSA.2017.081213.
- [12] Livieris, I.E., Pintelas, E., and Pintelas, P., 2020, "A CNN–LSTM Model for Gold Price Time-Series Forecasting," *Neural Comput. Appl.*, 32(23), pp. 17351–17360, doi: 10.1007/s00521-020-04867-x.
- [13] Vidal, A. and Kristjanpoller, W., 2020, "Gold Volatility Prediction using a CNN-LSTM Approach," *Expert Syst. Appl.*, 157, pp. 113481, doi: 10.1016/j.eswa.2020.113481.
- [14] Shahid, F., Zameer, A., and Muneeb, M., 2020, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM," *Chaos, Solitons & Fractals*, 140, pp. 110212, doi: 10.1016/j.chaos.2020.110212.
- [15] Phi, M., 2018, "Illustrated Guide to LSTM's and GRU's: A step by step explanation," *towards data science*. <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21> (accessed Apr. 01, 2021).
- [16] Hansun, S., Charles, V., Gherman, T., Subanar, and Indrati, C.R., 2021, "A Tuned Holt-Winters White-Box Model for COVID-19 Prediction," *Int. J. Manag. Decis. Mak.*, 20(3), pp. 241–262, doi: 10.1504/IJMDM.2021.10034422.
- [17] Hansun, S., Charles, V., Indrati, C.R., and Subanar, 2019, "Revisiting the Holt-Winters' Additive Method for Better Forecasting," *Int. J. Enterp. Inf. Syst.*, 15(2), pp. 43–57, doi: 10.4018/IJEIS.2019040103.
- [18] yahoo! finance, 2021, "Gold". <https://finance.yahoo.com/quote/GC%3DF?p=GC%3DF> (accessed May 15, 2021).
- [19] Moreno, J.J.M., Pol, A.P., Abad, A.S., and Blasco, B.C., 2013, "Using the R-MAPE Index as a Resistant Measure of Forecast Accuracy," *Psicothema*, 25(4), pp. 500–506, [Online]. Available: <http://www.psicothema.com/pdf/4144.pdf>.