2021 24th International Conference on Computer and Information Technology (ICCIT), 18-20 December, 2021.

Analysis and Prediction of Gold Price using CNN and Bi-GRU based Neural Network Model

Md. Masum Billah, and Sunanda Das
Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
masumpranto12@gmail.com, and sunanda@cse.kuet.ac.bd

***Abstract*—In recent decades, gold has been one of the most sought-after commodities for long-term and short-term investments, as investors perceive gold as a hedge against unanticipated market occurrences. Gold can be purchased, stored, and is rarely utilized as a payment method. However, it is pretty simple to convert gold into cash in almost any currency. As gold is essential for maintaining value, investment, and national economic stability, it is undoubtedly vital to forecast the price of gold accurately. In this paper, we proposed a hybrid method for forecasting the price of gold based on the combination of 1D Convolutional Neural Networks (CNN) and Bidirectional Gated Recurrent Unit (Bi-GRU). Though CNN-GRU, CNN-LSTM, CNN-RNN based hybrid networks or, the individual CNN, Bi-GRU, and GRU provide satisfactory results, our proposed hybrid approach is more reliable since it outperforms other networks and achieves the MAE, MAPE, MedAE, RMSE, MSE, MSLE of 11.88, 0.67%, 8.41, 15.59, 242.90, 76.08×10^{-6} respectively and R^2 Score of 93.56%.**

***Keywords*—Gold Price Forecasting, Deep Learning, CNN, Bi-GRU, Neural Network, Time Series Analysis.**

I. Introduction

Gold is glossy, metallic, and can be transformed effectively into bars, coins, or adornments [1]. It is the most obvious choice as a medium of exchange for goods and services. It does not rust, corrode, or rot, making it a long-term store of value, and humans are attracted to it physically and emotionally. It is not used up or consumed like other commodities such as oil or wheat. Investors reserve gold in their portfolios due to its unique inherent features. Societies and economies treasure gold, thereby maintaining its worth. The gold standard is a fixed monetary system under which government currency is settled and freely convertible into gold. A gold-based monetary system in which some or all nations fix their exchange rates on the basis of the relative gold parity values between individual currencies. Most proponents of commodity money select gold as a medium of trade because of its intrinsic characteristics. Gold is a safe asset to invest in, and it has always been significant throughout history.

Gold is appreciated not only for investment and jewelry, but it is also used in the production of a wide range of electrical and medical products [2], [3]. Central banks keep paper money and gold in reserve. Many countries in the world have reserves of gold. Demand for gold, central bank reserves, the value of the US dollar, and the desire to hold onto gold as a hedge against inflation and currency devaluation are all driving the gold price. A rising gold price is typically an indicator of underlying economic issues, as gold is viewed as a safe haven. Traders and individuals can invest in gold since it is a commodity that can help them weather financial turmoil. Gold is the metal we will fall into when other forms of money fail, which means gold will always be valuable in tough and good times.

CNN and Bi-GRU-based hybrid neural network model for predicting the opening price of gold in a 1-day interval has been developed in this paper. The proposed model is compared with CNN-GRU, CNN-LSTM, CNN-RNN, and several individual models employing R^2 Score and different errors such as MAE, MAPE, MedAE, RMSE, MSE, MSLE to evaluate its performance.

II. Related work

In the last decades, different approaches have been utilized for forecasting the gold price. Salis et al. [4] proposed a Long Short Term Memory (LSTM) model with four layers that predict the price of gold for the next day. They used 128 neurons in each layer and concluded that LSTM is able to produce precise results. To measure the model's performance, they used mean absolute error (MAE) and root mean squared error (RMSE). They also provided expert advice so that users could get guidance from the regular buyer of the stock.

Al-Dhuraibi et al. [5] analyzed whether the movement of the gold price would rise or fall in the near future. They applied different machine learning algorithms such as Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Linear Regression to predict the direction of gold price and witnessed that KNN produces a satisfactory performance.

Dubey [6] developed Support Vector Regression (SVR) and Adaptive Neural Fuzzy Inference System (ANFIS) for predicting the daily gold price. They observed that support vector regression outperformed the ANFIS model. At the same time, 'Adaptive Neural Fuzzy Inference System Grid Partitioning' (ANFIS-GP) performed slightly better than the 'Adaptive Neural Fuzzy Inference System Subtractive Clustering' (ANFIS-SC) model.

Manjula et al. [7] used three machine learning algorithms such as linear regression, random forest regression, and gradient boosting regression for predicting gold price. They demonstrated a comparative analysis among three models and concluded that random forest regression performs better in a longer period and, in a shorter period, gradient boosting regression is better.

He et al. [8] proposed the LSTM-CNN model with Attention Mechanism for a daily gold price forecast. They found that their proposed model outperforms other conventional models such as Support Vector Regression (SVR), Autoregression Integrated Moving Average (ARIMA), deep regression, CNN, LSTM, and also performs better than the CNN-LSTM network.

Bingol et al. [9] applied four algorithms such as linear regression, Support Vector Machine (SVM), Vector Autoregression model (VAR), and ARIMA in their prediction system. They audited that the ARIMA model outperformed the other regression model's performance by scoring the lowest RMSE.

Christina et al. [10] used type-2 neuro-fuzzy modeling for gold price prediction. Their findings are compared to the most widely used prediction method, ARIMA. They showed that type-2 neuro-fuzzy modeling has smaller errors, i.e., RMSE, MAE, and MAPE, than the ARIMA model for gold price prediction.

Sihananto et al. [11] proposed a hybrid approach between Evolution Strategies (ES) and Fuzzy Inference System (FIS) for forecasting gold price movement. The proposed method was compared with regular Mamdani FIS and evaluated using MAE. They concluded that the proposed ES-FIS had a lower error than regular FIS since the proposed method produced an error of 2.734 while regular Mamdani FIS provided an error of 4.036.

III. Methodology

A. Data Collection

Gold price data has been collected from Investing.com [12]. The data consists of six features, i.e., Date, Price, Open, High, Low, and Change. The 'Open' column has been taken and contains opening prices of gold in United States Dollars from 1st January 2016 to 30th June 2021. An actual time series plot for the open price of gold can be seen in Fig. 1a.

B. Data Preprocessing

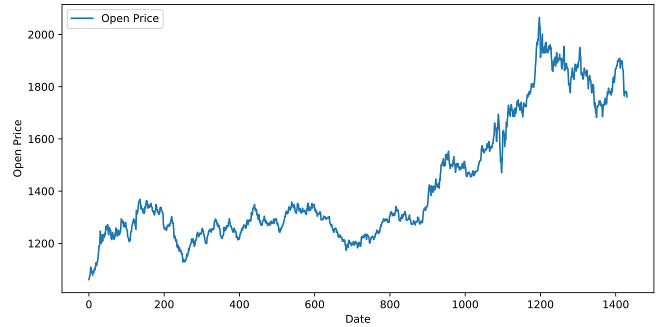
To maintain the open price between 0 and 1, the dataset was normalized using the min-max method. The min-max normalization formula for a range between 0 and 1 is applied as follows:

$$v' = \frac{v - min(v)}{max(v) - min(v)} \tag{1}$$

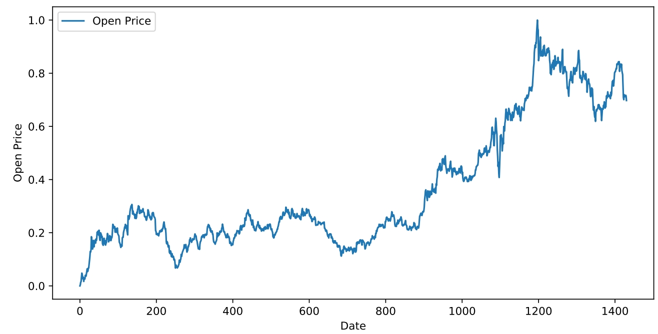
Where v is the actual value, and v' is the normalized value. After min-max normalization, the time series plot for the open price of gold can be seen in Fig 1b.

C. The Proposed Model

The proposed hybrid model consists of a sequentially working input layer, 1D-CNN layer, Bi-GRU, GRU, and Dense layer. The proposed model is illustrated in Fig. 2 and the summary of the model is given in Table I.



(a) Gold price before normalization



(b) Gold price after normalization

Fig. 1. Opening Price of Gold Before and After Normalization.

TABLE I. Summery of the Proposed Model.

Layers	Output Shape	Parameters
Conv1D	1×32	1,632
Bi-GRU	1×64	12,672
GRU	32	9,408
Dense	32	1,056
Dense	1	33
Total Parameters		**24,801**

1) Input Layer: The input layer is the first layer of the model, and it takes the normalized data and passes it on to the hidden layer.

2) Convolutional Neural Network (CNN): 1D-CNN has recently shown immense promises for different applications [13], [14]. The mathematical formula for 1D-CNN is as follows:

$$H(j) = \sum_{n=1}^{N_k} I(j+n) \times K(n) \tag{2}$$

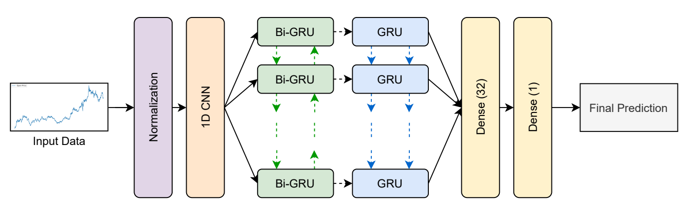


Fig. 2. Architecture of the Proposed Method.

TABLE II. MAE DEPENDENT ON OPTIMIZERS AND LOSS FUNCTIONS.

Loss Functions \ Optimizers	SGD	RMSprop	Adam	Adamax	Nadam	Adagrad
mean absolute error	17.89	15.73	11.88	12.11	13.20	21.57
mean absolute percentage error	84.26	13.62	12.38	12.20	13.03	13.87
mean squared error	24.10	17.76	12.19	12.21	13.24	49.48
mean squared logarithmic error	23.69	17.51	12.25	12.42	13.76	101.63
huber	23.93	16.59	12.42	12.58	13.09	97.79
log cosh	23.84	16.76	12.64	12.29	12.89	76.45

TABLE III. HYPER-PARAMETERS OF THE PROPOSED MODEL.

Parameters	Values
Window Size	10
CNN Filters	32
CNN Kernel Size	5
CNN Activation	ReLU
GRU Units	32
GRU Activation	Hyperbolic Tangent
Dense Units	32
Dense Activation	ReLU
Loss	Mean Absolute Error
Optimizer	Adam
Learning Rate	10^{-3}
Epochs	512
Batch Size	64

where $1 \leq j \leq N_i$. I and H are the N_i-dimensional input and output feature maps, respectively, and K is the N_k-dimensional convolutional kernel. In the proposed model, we utilized a 1D-CNN with a Rectified Linear Unit (ReLU) as an activation function to determine when the network's neurons should be activated. In CNN, each filter has its own weight and a predefined kernel size. For example, 32 convolutional filters with a kernel size of 5 were applied in the proposed CNN layer.

3) Gated Recurrent Unit (GRU): The GRU [15] is a recurrent neural network and has an advantage over LSTM as it is faster and requires less memory. Moreover, GRU addresses the vanishing gradient problem, and with the help of update gate and reset gate, it only allows to pass the appropriate information to produce better predictions.

4) Bidirectional Gated Recurrent Unit (Bi-GRU): The data processing of GRU is carried out in the order of the timestamp of the data. The output at the current moment is not only related to the input at the current moment but also related to the historical input. A Bidirectional GRU, also known as a Bi-GRU, is a two-GRU-based sequence processing model. One takes input in a forward direction, while the other takes it backward, thus help the model in producing more reliable decisions.

D. Performance Evalution

We applied six types of errors, such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Median Absolute Error (MedAE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Squared Logarithmic Error (MSLE) and R^2 Score for comparing the performance of our proposed hybrid model with others model. The following are the formulas for calculating the errors where v is an actual value, $\hat{v}$ is the predicted value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|v_i - \hat{v}_i| \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|v_i - \hat{v}_i|}{|v_i|} \tag{4}$$

$$MedAE = median(|v_1 - \hat{v}_1|, ..., |v_n - \hat{v}_n|) \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i - \hat{v}_i)^2} \tag{6}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(v_i - \hat{v}_i)^2 \tag{7}$$

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}(log(v_i + 1) - log(\hat{v}_i + 1))^2 \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(v_i - \hat{v}_i)^2}{\sum_{i=1}^{n}(v_i - \bar{v})^2} \tag{9}$$

IV. RESULT ANALYSIS

The whole dataset is divided into three sets, namely train, validation, and test set. The optimal window size of the proposed hybrid model is 10. In the 1D CNN layer, 32 filters, kernel size of 5, and ReLU activation have been used. Both GRU and Dense layer applied 32 units for building the proposed hybrid model. For determining the best optimizer and loss function of the proposed model, we varied different optimizers such as SGD, RMSprop, Adam, Adamax, Nadam, Adagrad, and loss functions such as mean absolute error, mean absolute percentage error, mean squared error, mean squared logarithmic error, huber, log cosh. The MAE for each combination of optimizers and loss function is inserted as matrix format in Table II. This experiment produces the lowest MAE of 11.88, when Adam optimizer with learning rate 10^{-3} and mean absolute error loss function were used. For fitting the model, 512 epochs and 64 batch size has been employed. The loss vs. epochs curve in Fig. 3 shows that the model is not overfitting. Hyper-parameters of the proposed method can be noticed in Table III.

TABLE IV. Comparative Analysis of Different Models.

Models	MAE ↓	MAPE(%) ↓	MedAE ↓	RMSE ↓	MSE ↓	MSLE(10^{-6}) ↓	R^2 Score(%) ↑
RNN	14.21	0.80	12.73	18.41	339.06	105.80	91.01
LSTM	14.05	0.79	11.80	18.29	334.36	104.21	91.14
2 Layer NN	13.87	0.78	12.23	18.13	328.69	102.38	91.29
GRU	13.93	0.78	12.14	18.24	332.75	104.01	91.18
Bi-GRU	13.83	0.78	11.37	18.02	324.86	101.32	91.39
CNN	13.63	0.77	11.43	17.84	318.18	99.37	91.56
CNN-RNN	13.19	0.74	10.32	17.40	302.75	94.95	91.97
CNN-LSTM	13.16	0.74	10.57	17.42	303.33	95.30	91.96
CNN-GRU	12.95	0.73	10.56	17.15	293.98	91.82	92.21
Proposed Model	**11.88**	**0.67**	**8.41**	**15.59**	**242.90**	**76.08**	**93.56**

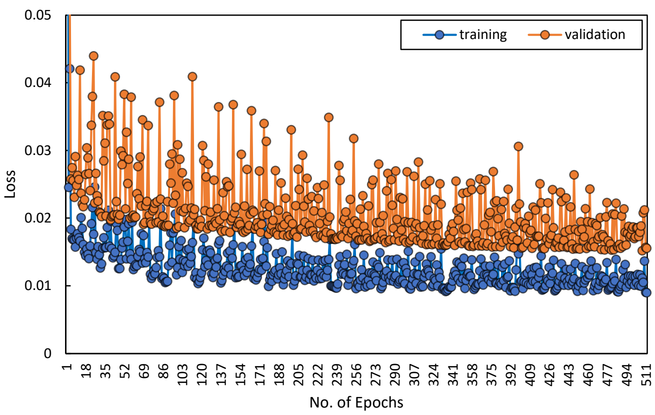


Fig. 3. Loss vs Epochs Curve of the Proposed Model.

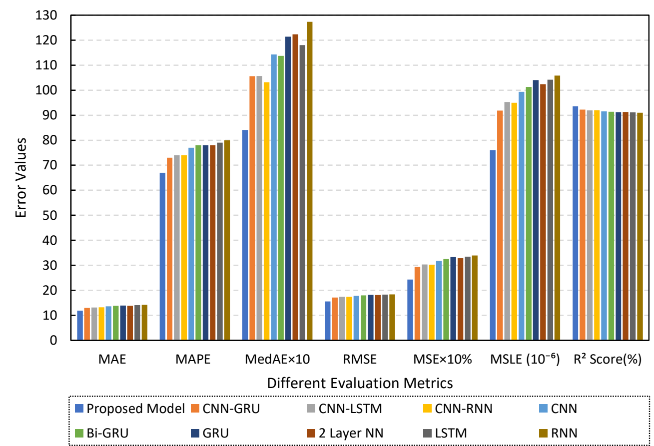


Fig. 4. Visual Representation of the Comparative Analysis of Different Models with Respect to Different Evaluation Metrics.

A comparative analysis has been demonstrated in Table IV between different models with respect to R^2 score and several errors such as MAE, MAPE, MedAE, RMSE, MSE, and MSLE. Downward arrow has been assigned after each error in the Table IV which indicates that the lowest the error value, the better the model, and the R^2 score is assigned an upward arrow which suggests that the model with the highest R^2 score performs better than other models with lower R^2 score. From the Table IV, it is conceivable that the proposed model produces the lowest error in every category and achieves MAE, MAPE, MedAE, RMSE, MSE, MSLE of 11.88, 0.67%, 8.41, 15.59, 242.90, 76.08×10^{-6} respectively as well as produces the highest R^2 score of 93.56% and thereby successfully exceeds other models such as CNN-LSTM, CNN-GRU, CNN-RNN, CNN, etc. A visual representation of the comparative analysis can be found in Fig. 4.

The efficiency of the proposed model concerning various prediction lengths is described in Fig 5. The visual illustration ensures that our proposed model produces the lowest mean absolute error in every time step.

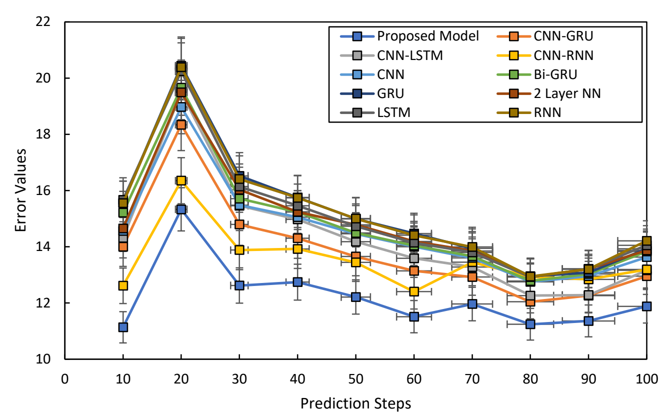
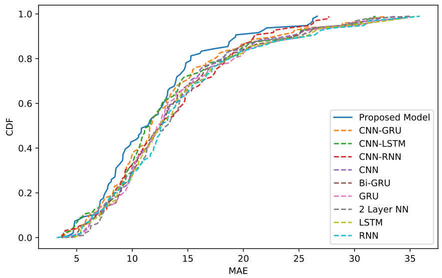
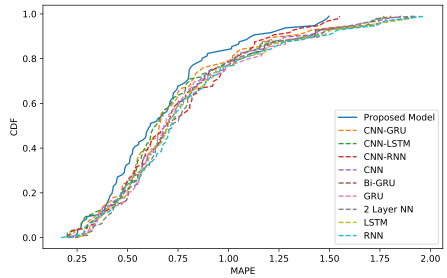


Fig. 5. Prediction Length Effect Comparison of Different Models

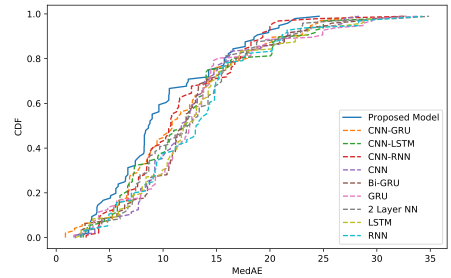
For statistical analysis of the proposed hybrid model, we have plotted the Cumulative distribution function (CDF) with respect to the different errors i.e., MAE, MAPE, MedAE, RMSE, MSE, and MSLE for all models as inspected in Fig. 6. The proposed model has a greater area under the curve for each error than other models, which indicates the robustness of the proposed model. The different time-step prediction curves of the proposed hybrid model are displayed in Fig. 7 where the blue line and orange line represent original and predicted gold price, respectively. It represents how accurately the proposed model can predict the open price of gold for the next 7 days, 30 days, 60 days, and 100 days.



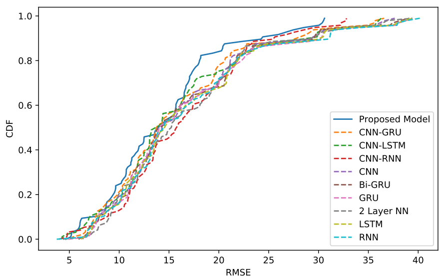
(a) CDF vs MAE



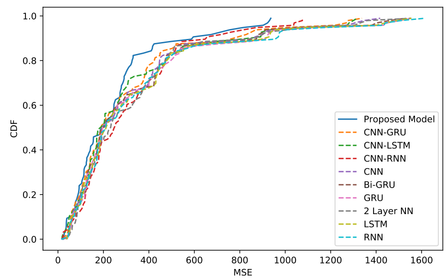
(b) CDF vs MAPE



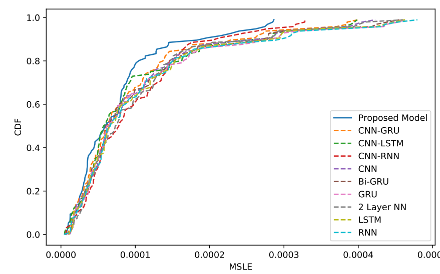
(c) CDF vs MedAE



(d) CDF vs RMSE

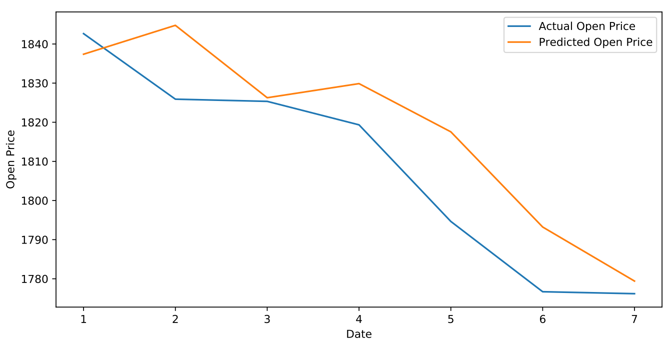


(e) CDF vs MSE

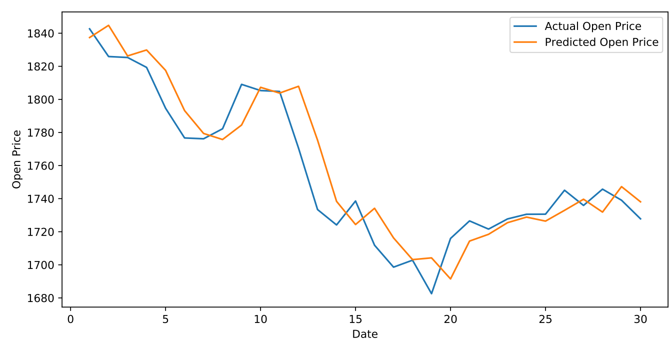


(f) CDF vs MSLE

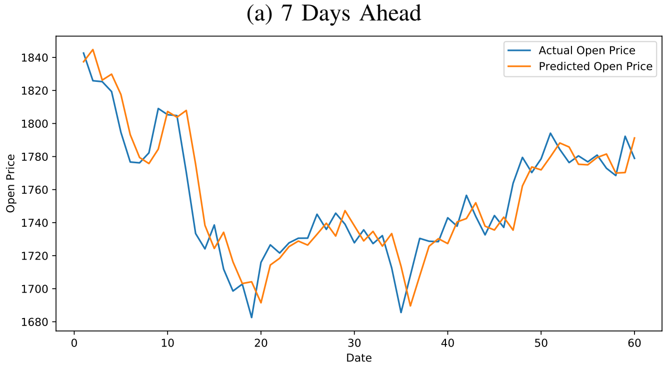
Fig. 6. Analysis of CDF Plots for Different Errors.



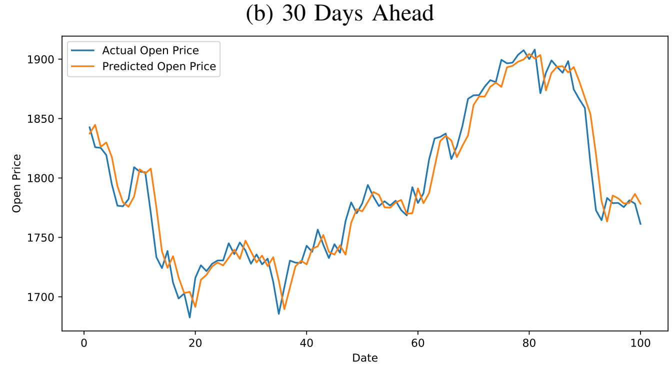
(a) 7 Days Ahead



(b) 30 Days Ahead



(c) 60 Days Ahead



(d) 100 Days Ahead

Fig. 7. Multi-step Ahead Gold Price Prediction Using the Proposed Hybrid Model.

V. Conclusion

Gold is widely used as a monetary standard and as an alternative to paper currency. This paper aimed to develop a method to aid gold investors in deciding the best time in the future to buy or sell gold. This paper combines 1D CNN, Bi-GRU, and GRU networks with 2 Layer NN to form a hybrid model. In addition, other predictive hybrid deep learning networks such as CNN-GRU, CNN-LSTM, CNN-RNN, and individual networks such as CNN, GRU, LSTM, etc., were also developed for comparison purposes. The comprehensive result analysis confirms the proposed hybrid model's reliability compared to other models. Thus, the proposed model is able to predict gold prices accurately, allowing gold to be purchased and sold at profitable conditions while reducing investment risk.

References

[1] Investing in gold. Available: https://www.investopedia.com/ask/answers/06/investingingold.asp. [Accessed Aug. 25, 2021].

[2] S. P. Pricker, "Medical uses of gold compounds: past, present and future," *Gold bulletin*, vol. 29, no. 2, pp. 53–60, 1996.

[3] P. Goodman, "Current and future uses of gold in electronics," *Gold bulletin*, vol. 35, no. 1, pp. 21–26, 2002.

[4] V. Salis, A. Kumari, and A. Singh, *Prediction of Gold Stock Market Using Hybrid Approach*, 01 2019, pp. 803–812.

[5] W. A. Al-Dhuraibi and J. Ali, "Using classification techniques to predict gold price movement," in *2018 4th International Conference on Computer and Technology Applications (ICCTA)*, 2018, pp. 127–130.

[6] A. D. Dubey, "Gold price prediction using support vector regression and anfis models," in *2016 International Conference on Computer Communication and Informatics (ICCCI)*, 2016, pp. 1–6.

[7] K. A. Manjula and P. Karthikeyan, "Gold price prediction using ensemble based machine learning techniques," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1360–1364.

[8] Z. He, J. Zhou, H.-N. Dai, and H. Wang, "Gold price forecast based on lstm-cnn model," in *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 2019, pp. 1046–1053.

[9] S. Bingol, S. Sadik, R. M. Abu Bakar, H. Mouhcine, J. Albaf, S. Chaabene, and A. Auta, "Gold price prediction in times of financial and geopolitical uncertainty: A machine learning approach," 09 2020.

[10] C. Christina and R. F. Umbara, "Gold price prediction using type-2 neuro-fuzzy modeling and arima," in *2015 3rd International Conference on Information and Communication Technology (ICoICT)*, 2015, pp. 272–277.

[11] A. N. Sihananto and F. A. Bachtiar, "Gold price movement forecasting using hybrid es-fis," in *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, 2017, pp. 321–326.

[12] Xau usd historical data. Available: https://www.investing.com/currencies/xau-usd-historical-data. [Accessed Jul. 1, 2021].

[13] H. Cho and S. M. Yoon, "Divide and conquer-based 1d cnn human activity recognition using test data sharpening," *Sensors*, vol. 18, no. 4, p. 1055, 2018.

[14] M. Azizjon, A. Jumabek, and W. Kim, "1d cnn based network intrusion detection with normalization on imbalanced data," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, 2020, pp. 218–224.

[15] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.