COMPARATIVE ANALYSIS OF GOLD PRICE PREDICTIONS USING DIFFERENT TIME SERIES MODELS

N. Srilekha Ph.D Research Scholar Department of Statistics, Annamalai University Chidambaram, Tamilnadu India.

Dr.P.Arumugam Associate Professor Department of Statistics, Annamalai University Chidambaram, Tamilnadu India.

Dr.K. Rajyalakshmi Assistant Professor Department of Engineering Mathematics, College of Engineering, Koneru Lakshmaiah Education Foundation, Greenfields, Vaddeswaram, Guntur-522302, Andhra Pradesh.

Abstract:

Gold is considered as more powerful foil among all foils are mined from the Earth. Due to the importance of Gold in financial markets researchers are interested to predict the future value. In this current study, mainly focused on estimating the future value of gold prices using different time series models namely Auto regressive integrated moving average (ARIMA), Gated recurrent model (GRU), Long short-term memory (LSTM), Bi-directional long short-term memory (Bi-LSTM). Accuracy metrics for each model are also presented. Based on the analysis GRU model shows the less error (%) with values of mean absolute error 79.1788, Root means square error 82.8688. Also, predicted future values with more accuracy represented through time series plots.

Key words: Gold prices, ARIMA, GRU, LSTM, Bi-LSTM, Performance metrics.

Historically, humans treated Gold as more powerful foil among all foils are mined from the Earth. Investment in correct time and right place is crucial to grow financially. Gold is an isolated product in the commodity market and considered as one of the stock contract products. The mining companies, investors, financial institutions and related firms are required an accurate prediction model for investigating the fluctuations in the gold price for taking the correct decisions. But the gold future price data belongs to the time series forecast as well as it is complex to predict the price because of its noisy, chaotic and non-stationary features of data. Various elements are responsible for increment of prices in the gold. Particularly, Low interest/conventional rate, Savings plan with low loan fee and no liability on investors. The chosen auxiliary data shows the positive graphical portrayal based on the analysis of gold prices. Over the years many people choosing for preservation of wealth is in the form of physical gold. The special characteristics of gold lead to considered as a tangible asset. Many authors focused on the importance of gold in the present and past scenario. Undefined time frame trend analysis has carried out through occasion and no-occasional information by Akaike in the year (1974). He introduced the concept of non-seasonal model to determine the best order. Brock well and Davies (1996) extended the study of time frame data using different techniques. In the year (2015), Aarti and Saina estimate the forecasting prices of gold using ARIMA model. Guha and Bandyopadhyay (2016) made an attempt to study the estimation of gold prices using ARIMA model. By adopting the Box and Jenkin's methodology Fuller proposed Augmented Fuller test to check the stationarity. Using ARIMA model, non-stationarity can be eliminated by integrating series. Sandhyarani et al. (2017) predicts the life cycle of batteries using ARIMA. Surendra et al. (2021) observed the patterns of gold in India for different time periods by using ARIMA model. Ioannis et *al*. (2020) implemented a Convolutional Neural Network-Long Short-term Memory (CNN-LSTM) method for gold price time-series forecasting. Srilekha et al. (2021) examined the price of gold and observed that it has soared to a record due to Corona Virus pandemic in 2020. Navya Prathyusha et al. (2022) examined the performance measures on for predicting diabetes milletus using machine learning approach. In this study, mainly focused on estimating the future prices of gold and error (%) by using original values. Augmented Duckey Fuller test (ADF) used to remove the non-stationarity that exists in the data. In the present study, different time series models compared and mention the performance metrics of each model with actual and predicted gold prices of time series of data.

Methodology

ARIMA:

Non seasonal ARIMA model of Box and Jenkin (1976) is represented as ARIMA(p,d,q)
Here p, d and q are non-negative integers
P: number of autoregressive observations
d: number of non-seasonal differences -to make stationary
q: order of moving average model
Seasonal ARIMA model is represented as SARIMA(p,dq)(P,D,Q)$_s$
where p =non seasonal AR order, d is non seasonal differencing, q is non seasonal moving average order, P is seasonal AR order, D is seasonal differencing, Q is seasonal moving order and S is time span of repeating seasonal pattern.
For a stationary time series, the ARIMA model can be represented as follows

$$y_t = \mu + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} - \gamma_1 \epsilon_{t-1} - \gamma_2 \epsilon_{t-2} - \cdots - \gamma_q \epsilon_{t-q} + \epsilon_t \quad (1)$$

where μ is constant, φ : Autoregressive parameter, $\gamma_:$ moving average parameter and ϵ_t: error term at time 't'.
To cross check the fixed time series, using Augmented Duckey-Fuller test. Dismissing invalid speculation recommends of period series fixed. Finally, Lung-Box test help us to indicate that the data is independently distributed or exhibiting the serial correlation. The minimum values of Akaike information criteria (AIC) and Bayesian information criteria (BIC) are providing the best and appropriate model which would be the reality of the situation. The predicted values should compare with the table values and determine the model accuracy and perfect fit.

GRU :

A Gated recurrent model (GRU) for accurate gold price prediction with less error values have been used in this research. When compared to LSTM, GRU consumes less memory and is quicker. For time series applications, GRU have a number of advantages over conventional Recurrent Neural Network (RNN). GRU uses an update and reset gate to deal with the RNN's vanishing gradient problem.

Update gate: The update gate works in helping the network to manage how much of previous data needs to be passed further which is significant because the network can choose to recognize all of the data from the past. The formula for updating gate is represented in Eq. (7)

$$Z_t = \sigma(W_z \odot x_t + U_z \odot h_{t-1} + b_z) \quad (2)$$

Where, Z_t is update gate activation vector, x_t is input vector, σ is sigmoid function, W, U are weights, b is the bias vector, $\odot$ is the symbol for Hadamard product, and h_{t-1} is output of previous vector.

Reset gate: The reset gate in the network determines how much of the prior information to forget. The formula for reset gate is represented in Eq. (8)

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (3)$$

Where, r_t is a reset gate activation vector.

$$\widetilde{h_t} = \tan h(W_h x_t + U_h(r_t * h_{t-1}) + b_h) \quad (4)$$

Where, $\widetilde{h_t}$ is a candidate vector

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \widetilde{h_t} \quad (5)$$

Initially, the update gate is computed using x_t, h_{t-1}, and σ. In the same way, the reset gate is calculated using its own weights and bias vectors.

LSTM:

LSTM is extended version of Recurrent Neural Networks (RNN) that used to easy remembering of past data in the memory. The issue of vanishing gradient occurred in RNN is solved by LSTM. Zhongqi Miao et *al* (2022) suggested an optimal portfolio approach related on real time

prediction of gold and bitcoin prices. Simple Moving Average (SMA) was utilized to predict the initial data and Long Short-Term Memory (LSTM) was utilized to predict the price the recent long-term data as well as upgraded actual time price data was projected in the aspect of data prediction.
The LSTM comprises of five main parts which are detailed below:

- Input gate (i_t): It alters the cell state with the goal of including a new data related to the current time step. Moreover, the included additional data may affects the stock price movement.
- Output gate (o_t): This used to decide what should be the next hidden state and provides the final related data for predicting the stock price.
- Hidden state (h_t): It is the multiplication of the output gate vector with cell state vector.
- Cell state (c_t): It has the data which exist in the memory after the earlier time step.
- Forget gate (f_t): It used to alter the cell state and used to avoid the values with less importance from the earlier time steps. This elimination is used to forget the unwanted data which doesn't has any effect on future price prediction.

The operations of LSTM are expressed in the following equations (6) to (11):

$$i_t = \sigma(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}) \quad (6)$$
$$f_t = \sigma(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}) \quad (7)$$
$$o_t = \sigma(W^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)}) \quad (8)$$
$$c_t = i_t \odot \tilde{c}_t + f_t \odot c_{t-1} \quad (9)$$
$$\tilde{c}_t = \tanh\,(W^{(c)}.[h_{t-1}, x_t] + b^{(c)}) \quad (10)$$
$$h_t = o_t \odot \tanh\,(c_t) \quad (11)$$

Where, input vector is x_t; previous cell state and hidden state are denoted as c_{t-1} and h_{t-1} respectively; U and W are hidden-to-hidden and input-to-hidden weight matrices; logistic sigmoid function is denoted as σ and element-wise multiplication is denoted as $\odot$.

Description of Data
To complete this work, the analyst has considered the information from Kaggle website. Train the ARIMA, GRU and LSTM models by using 90% training data and also test the model by remaining 10% testing data. The final output of each model is presented
.

Result analysis
Over the years, the prices of gold are increased day by day. The descriptive statistics of the data along with analysis using Python programming from 2013 to 2022 prices of gold are given as follows. Actual and predicted values of ARIMA, GRU, LSTM models are presented in the following figures (1), (2), (3) and (4).

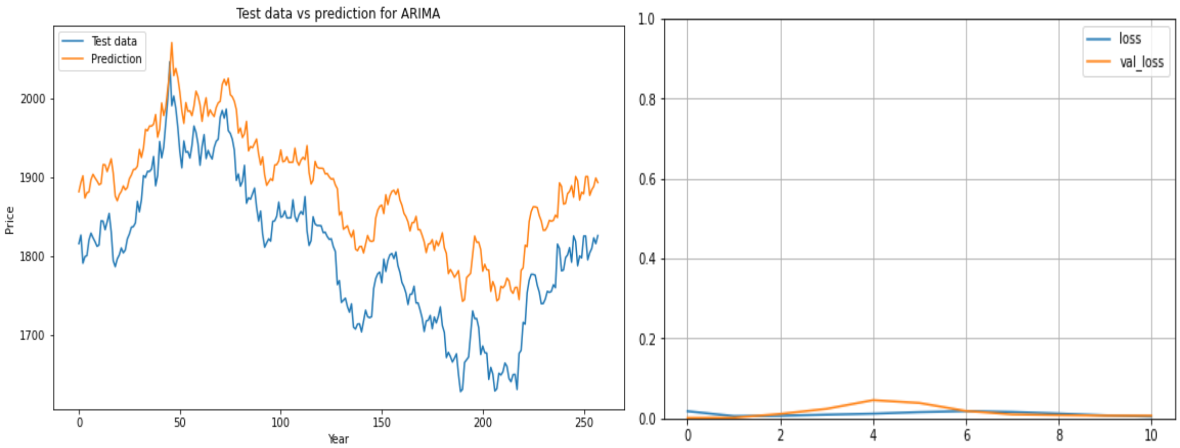

Figure (1): ARIMA model

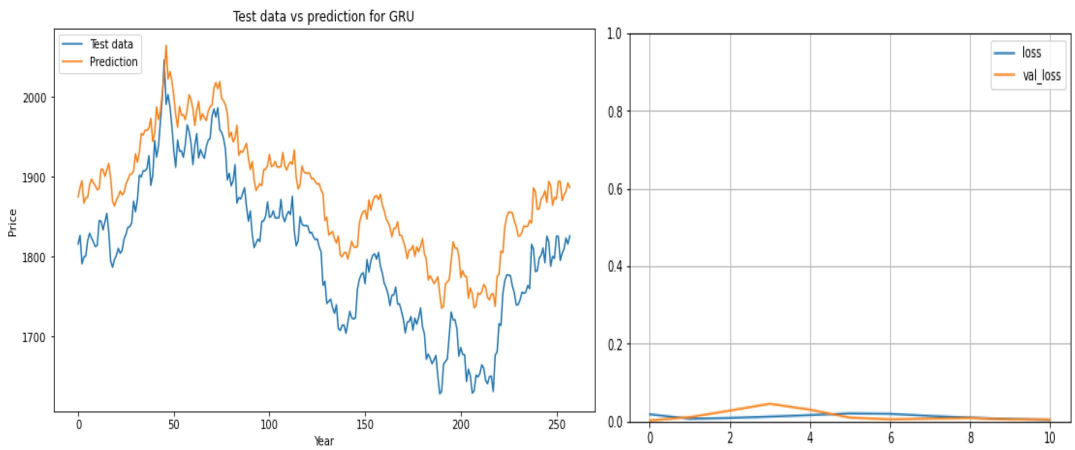

Figure (2): GRU Model

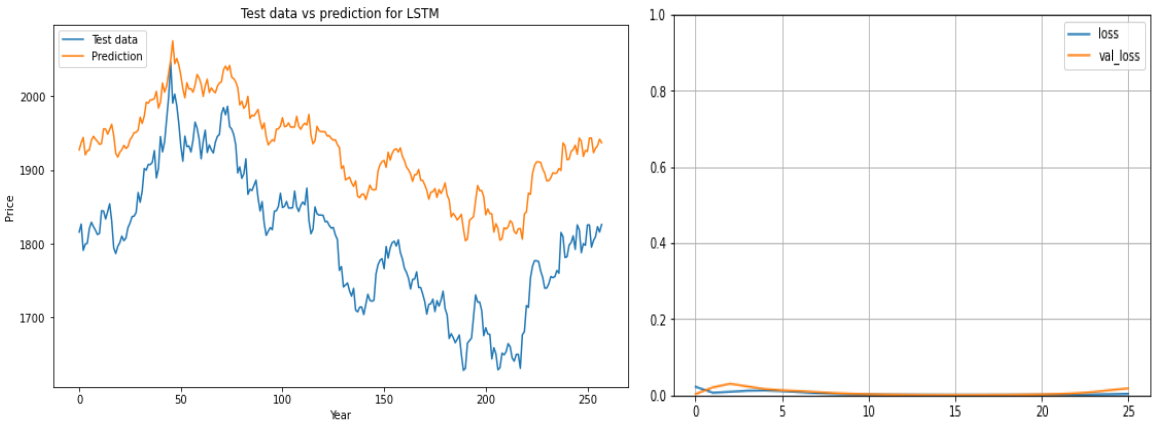

Figure (3): LSTM Model

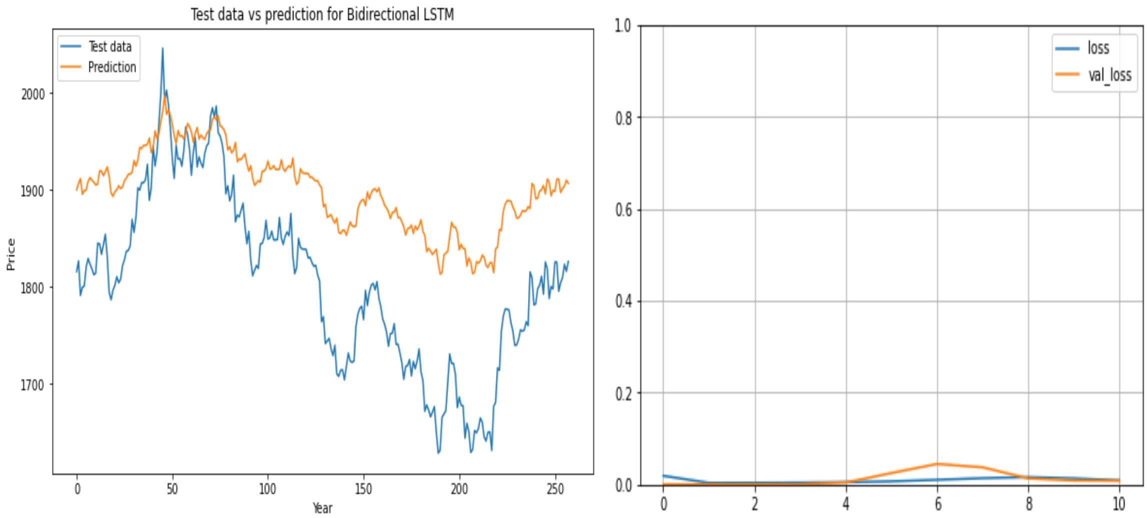

Figure (4): Bi-LSTM Model

ARIMA: 9/9 [==============================] - 1s 4ms/step
GRU: 9/9 [==============================] - 1s 3ms/step
LSTM: 9/9 [==============================] - 1s 3ms/step
Bi-LSTM 9/9 [==============================] - 2s 4ms/step

ARIMA:
Mean Absolute Error: 79.1788
Root Mean Square Error: 82.8688

GRU:
Mean Absolute Error: 72.3810
Root Mean Square Error: 76.3130

LSTM:
Mean Absolute Error: 122.7204
Root Mean Square Error: 127.3613

Bidirectional LSTM:
Mean Absolute Error: 96.2403
Root Mean Square Error: 107.8040

Conclusion

In the present paper, different time series models are suggested to forecast the gold prices of actual and estimated values and also calculated error (%) for each model also depicted the figures (1), (2), (3) and (4). Fitted models shows Loss and value loss functions for each model. Based on the performance metrics GRU model shows more accuracy with MAE of 72.3810 and Root mean square error of 76.3130 and it also takes less time to compute the data.

BIBLIOGRAPHY:

1. Aarthi M Sharma and Saina Baby, Gold price forecasting in India using ARIMA modelling, GE-International Journal of Management Research 2015, 3(10), 14-33.
2. Akaike H, A New look of the statistical model identification, IEEE transaction on automatic control, 19 (6), 716-723, 1974.
3. Angelidis.T, Benos.A, and Degiannakis.S The use of GARCH models in VaR estimation. Statistical Methodology, 1(1-2):105–128, 2004.
4. Banhi Guha and Gautam Bandyopadhyay, Gold Price Forecasting Using ARIMA Model, Journal of Advanced Management Science, 4(2), 118-121, 2016.
5. Brockwell P. J. and Davis R. A, Introduction to Time Series and Forecasting. Springer, New York. Sections 3.3 and 8.3, 1996.
6. Box GEP and Jenkin G.M, Time series of analysis. Forecasting and Control, Sam Franscico, Holden Day, California, USA, 1976.
7. Barndorff-Nielsen.O.E and Shephard.N Realized Power Variation and Stochastic Volatility Models. Bernoulli, 9(2):243–265, 2003.
8. Barndorff-Nielsen O. E and Shephard.N Power and Bipower Variation with Stochastic Volatility and Jumps. Journal of Financial Econometrics, 2(1):1–37, 2004.
9. Barndorff-Nielsen O. E. and N. Shephard. Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. Journal of Financial Econometrics, 4(1):1–30, 2006.
10. Black F. and Scholes. M. The Pricing of Options and Corporate Liabilities. The Journal of Political Economy, 81(3):637, 1973.
11. Bollerslev.T Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3):307–327, 1986.
12. Bollerslev,T R. Y. Chou, and K. F. Kroner. ARCH modeling in finance. Journal of Econometrics, 52(1), 1992.
13. Bollerslev.T, R. F. Engle, and D. B. Nelson. Chapter 49 Arch models, volume 4, pages 2959–3038. 1994. Dhruvi Sarvaiya and Disha Ramchandani. Time Series Analysis and Forecasting of Gold Price using ARIMA and LSTM Model, International Journal for Research in Applied Science & Engineering Technology, 2022, 10(9), 168-173.
14. *Fuller W. A, Introduction to Statistical Time Series. New York: John Wiley and Sons, 1976, ISBN 0-471-28715-6.*
15. https://www.python.org/

16. Engle. R.F Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4):987–1007, 1982.
17. Engle.R.F GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. The Journal of Economic Perspectives, 15(4):157–168, 2001.
18. Engle R. F. and Patton. A. J. , What good is a volatility model? Quantitative Finance, 1(2):237–245, 2001.
19. Fama. E. F The Behavior of Stock Market Prices. Journal of Business, 38(1), 1965.
20. Fern´andez.C and Steel.M.F.J On Bayesian Modeling of Fat Tails and Skewness. Journal of the American Statistical Association, 93(441):359–371, 1998.
21. Livieris, Ioannis E., Emmanuel Pintelas, and Panagiotis Pintelas. "A CNN–LSTM model for gold price time-series forecasting." Neural computing and applications 32 (2020): 17351-17360.
22. Li, Y., Wang, S., Wei, Y. and Zhu, Q. A new hybrid VMD-ICSS-BiGRU approach for gold futures price forecasting and algorithmic trading. *IEEE Transactions on Computational Social Systems*, 2021, *8*(6), pp.1357-1368.
23. Navya Prathyusha M, Rajyalakshmi K, Apparao B V and Charankumar G. Impact of sleep on usage of the smart phone at the bed time-A Case study, Mathematics and Statistics, 9(1), 31-35, 2021.
24. Sandhyarani N, Veeraiah D, Shanmugam M, Raviraju B, Amudhavel J, An enhanced ARIMA model for predicting life cycle of the batteries for remote Wi-Fi enabled device, Journal of Advanced Research in Dynamical and Control Systems, 2017, 9(12),1183-1197.
25. Srilekha, Nallamothu., Rajyalakshmi, K. and Arumugam, P., 2023. Forecasting Gold Prices in India Using an ARIMA Model, AIP Conference Proceedings, 2707, 040012.
26. Surendra J, Rajyalakshmi K, Apparao B V and Charankumar G. Forecast and Trend analysis of gold prices in India using ARIMA, Journal of Mathematical and Computational Science, 11(2), 1166-1175, 2020.
27. Zhongqi, M and Wenxuan Huang. "An optimal portfolio method based on real time prediction of gold and bitcoin prices." Systems Science & Control Engineering 2022, 10, 1: 653-661.
28. Engle. R.F Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4):987–1007, 1982.

