Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD24518

Improved Gold Price Prediction Based on the LSTM-ARIMA Hybrid Model

Yucheng Ye

Aberdeen Institute of Data Science and Artificial Intelligence, South China Normal University, Foshan, China
y.ye.22@abdn.ac.uk

Abstract. Gold price forecasting is a crucial task in financial markets due to gold's unique attributes as a commodity, precious metal, and currency. Traditional time series models such as ARIMA are effective in capturing linear trends and seasonal components but struggle with nonlinear dependencies present in gold price data. In contrast, deep learning models, especially Long Short-Term Memory (LSTM) networks, excel at modeling complex nonlinear relationships and long-term dependencies. However, standalone LSTM models may overlook certain linear patterns. This paper proposes a novel hybrid forecasting approach that integrates LSTM and ARIMA models to leverage the strengths of both methodologies. The LSTM model first learns nonlinear features from historical gold price data, and the ARIMA model further models the residuals to capture remaining linear trends. Empirical analysis is conducted using daily gold price data from August 19, 2013, to November 22, 2024. Experimental results demonstrate that the hybrid LSTM-ARIMA model significantly outperforms the standalone LSTM model across all major evaluation metrics, with remarkable reductions in forecasting errors and improvements in accuracy and robustness. The proposed hybrid model offers a more reliable and precise tool for gold price prediction, providing valuable quantitative support for investors and policymakers in the gold market.

Keywords: Artificial Intelligence, Deep Learning, Gold Price Forecasting, LSTM-ARIMA Hybrid Model, Time Series Analysis.

1. Introduction

Gold, as an important financial asset, possesses the triple attributes of commodity, precious metal, and currency [1], playing a crucial role within the international monetary system and global financial markets [2]. Gold prices not only reflect changes in the global economic and financial environment [3] but also serve as a "barometer" and "alarm signal" for global political and economic developments [4]. Fluctuations in gold prices exert profound impacts on global economic and financial stability; thus, accurately predicting and analyzing gold price trends has become a primary concern for investors, fund managers, and policymakers [5].

Traditionally, classical time series analysis methods, such as the Autoregressive Integrated Moving Average (ARIMA) model, have been widely employed for gold price forecasting. These methods perform well in capturing linear trends, seasonal patterns, and cyclical behaviors inherent in time series data [7]. However, the main limitation of the ARIMA model lies in its inability to effectively

DOI: 10.54254/2755-2721/2025.LD24518

handle nonlinear relationships [8]. In recent years, with the rapid advancements in data science and artificial intelligence technologies, machine learning techniques—particularly those based on deep learning—have increasingly been applied to gold price prediction research. Among these techniques, the Long Short-Term Memory (LSTM) network, a specialized variant of Recurrent Neural Networks (RNNs), has emerged as a widely adopted approach due to its powerful nonlinear modeling capabilities and superior ability to capture long-term dependencies [9]. Consequently, LSTM models have been extensively utilized for predicting prices in financial markets [10].

Although the LSTM model exhibits excellent performance in modeling complex nonlinear patterns, studies have demonstrated its relatively lower sensitivity to certain linear trends and seasonal patterns embedded within data [11]. Consequently, relying solely on the LSTM model may not adequately capture all pattern features present in gold price data. In recent years, the scholarly community has proposed the concept of hybrid models, combining traditional linear models with modern deep-learning models to improve forecasting accuracy and robustness [12]. For instance, Achkar compared forecasting results using data from Facebook and Google stocks as well as Bitcoin prices, demonstrating that hybrid models based on LSTM significantly outperform traditional backpropagation multi-layer perceptrons (BPA-MLP) models [13].

Motivated by the above considerations, this study proposes a hybrid modeling approach that integrates LSTM and ARIMA models. The goal of this approach is to leverage the strengths of the ARIMA model in capturing linear trends, seasonal patterns, and cyclical behaviors, along with the LSTM model's strong capability for modeling nonlinear relationships and long-term dependencies, thereby enhancing the accuracy of gold price forecasts [14]. Specifically, this study first employs the LSTM network to model the gold price time series data, learning complex nonlinear dependencies inherent within. Subsequently, the ARIMA model is utilized to further analyze the residuals between actual values and LSTM predictions, capturing potential linear trends and cyclical patterns in these residuals, thus effectively reducing forecasting errors.

2. Methodology

The construction of a hybrid LSTM-ARIMA model consists of four steps: LSTM prediction, residual series calculation, ARIMA modeling of residuals, and hybrid forecasting.

2.1. LSTM prediction

Mathematically, the basic mechanism of LSTM prediction can be expressed as [14]:

$$\hat{y}_t^{\mathrm{LSTM}} = f_{\mathrm{LSTM}}(y_{t-1}, y_{t-2}, \ldots, y_{t-p}, X_{t-1}, X_{t-2}, \ldots, X_{t-q}) \tag{1}$$

where y_t denotes the actual gold price at time t, and $\hat{y}_t^{\mathrm{LSTM}}$ represent the predicted gold price at time t, obtained from the LSTM model. The function $f_{\mathrm{LSTM}}(\cdot)$ denotes the trained LSTM network mapping function. p is the length of the historical price time window used for prediction. X_{t-i} represents the vector of exogenous variables (e.g., macroeconomic indicators, technical indicators, etc.) at time $t-i$. q is the length of the historical time window for the exogenous variables.

2.2. Residual series calculation

Calculate the residual series between the actual values and the LSTM-predicted values:

$$e_t = y_t - \hat{y}_t^{\mathrm{LSTM}} \tag{2}$$

DOI: 10.54254/2755-2721/2025.LD24518

where e_t denotes the residual at time t.

2.3. ARIMA modeling of residuals

The residual series e_t obtained from the LSTM predictions is modeled using an ARIMA(p′, d, q′) model as follows [15]:

$$\phi(B)(1-B)^d e_t = \theta(B)\varepsilon_t \tag{3}$$

Where B denotes the backshift operator, defined as $Be_t = e_{t-1}$. $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_{p'} B^{p'}$ is the autoregressive polynomial, $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_{q'} B^{q'}$ is the moving average polynomial, and $(1-B)^d$ is the differencing operator, where d represents the order of differencing. The parameter ε_t is a white noise process, and p' and q' represent the autoregressive and moving average orders of the ARIMA model, respectively.

The ARIMA-based prediction of the residual at time $t+h$ can be expressed as the conditional expectation:

$$\hat{e}_{t+h} = E[e_{t+h} \mid \{e_t, e_{t-1}, \ldots\}] \tag{4}$$

According to the computational process of the ARIMA model, this prediction can be further expanded as:

$$\hat{e}_{t+h} = \sum_{i=1}^{p'} \phi_i e_{t+h-i} - \sum_{j=1}^{q'} \theta_j \varepsilon_{t+h-j} \tag{5}$$

2.4. Hybrid forecasting

The final prediction of the hybrid model is derived by combining the LSTM prediction and the ARIMA residual prediction, represented mathematically as follows:

$$\hat{y}_{t+h}^{Hybrid} = \hat{y}_{t+h}^{LSTM} + \hat{e}_{t+h} \tag{6}$$

Where $\hat{y}_{t+h}^{Hybrid}$ denotes the final hybrid model prediction of the gold price at time $t+h$, $\hat{y}_{t+h}^{LSTM}$ denotes the LSTM model prediction of the gold price at time $t+h$, and $\hat{e}_{t+h}$ represents the ARIMA model prediction of the residual at time $t+h$.

2.5. Data source and model implementation

The proposed hybrid model is applied to gold price forecasting analysis using daily-frequency gold price time series data covering the period from August 19, 2013, to November 22, 2024. The dataset contains five key price indicators for each trading day: opening price, closing price, highest price, lowest price, and trading volume, providing comprehensive dynamic information on gold prices for the modeling process.

The constructed hybrid forecasting framework effectively integrates the advantages of deep learning and traditional econometric methods. Specifically, the Long Short-Term Memory (LSTM) network, with its specialized gating mechanisms, can effectively capture nonlinear characteristics and long-term dependencies inherent in gold price data. Meanwhile, the Autoregressive Integrated Moving Average (ARIMA) model demonstrates significant strengths in handling linear components, seasonal fluctuations, and autocorrelation structures within residual series. By systematically combining the

DOI: 10.54254/2755-2721/2025.LD24518

complementary capabilities of these two methodologies, the hybrid model effectively overcomes the prediction limitations of single-model approaches in complex market environments, thereby significantly enhancing both the overall accuracy and robustness of gold price time series forecasting.

3. Result and discussion

3.1. Evaluation metrics for forecasting performance

This study employs multiple evaluation metrics to comprehensively quantify and compare the forecasting performance of the LSTM model and the hybrid LSTM-ARIMA model. Table 1 provides a detailed comparison of the performance metrics for both models on the test dataset, along with the improvement rates achieved.

Table 1. Comparison of forecasting performance metrics between LSTM and hybrid LSTM-ARIMA models

Assessment of Indicators	MAE	MSE	RMSE	MAPE	SMAPE	R^2	DA	maximum error	accuracy
LSTM	129.139306	19084.251566	138.145762	0.053231	5.483550	0.596436	51.739130	233.216713	0.946769
LSTM+ARIMA	19.468537	608.690949	24.671663	0.008197	0.820077	0.987128	53.478261	89.784918	0.991803
Improvement rate (%)	84.924391	96.810507	82.140847	84.601779	85.044776	65.504564	3.361345	61.501508	4.756638

As observed from Table 1, the hybrid model significantly outperforms the standalone LSTM model across all evaluation metrics. Particularly noteworthy is the improvement in mean squared error (MSE), wherein the hybrid model achieves an impressive reduction of 96.81%. Similarly, mean absolute error (MAE) and Root Mean Squared Error (RMSE) declined by 84.92% and 82.14%, respectively. The mean absolute percentage error (MAPE) notably decreases from 5.32% in the LSTM model to 0.82% in the hybrid model, indicating a substantial increase in forecasting precision. Additionally, the coefficient of determination (R^2) improves markedly from 0.596 to 0.987, closely approaching the ideal value of 1, reflecting a significant enhancement in the hybrid model's fitting capability. Moreover, the maximum forecasting error is reduced by 61.50%, suggesting that the hybrid model demonstrates greater robustness in handling extreme cases.

3.2. Visualization analysis of forecasting accuracy

To intuitively evaluate the forecasting capabilities of the two models, Figure 1 provides a scatter plot comparison between predicted and actual values on the test dataset for both the LSTM and the hybrid LSTM-ARIMA models.

DOI: 10.54254/2755-2721/2025.LD24518

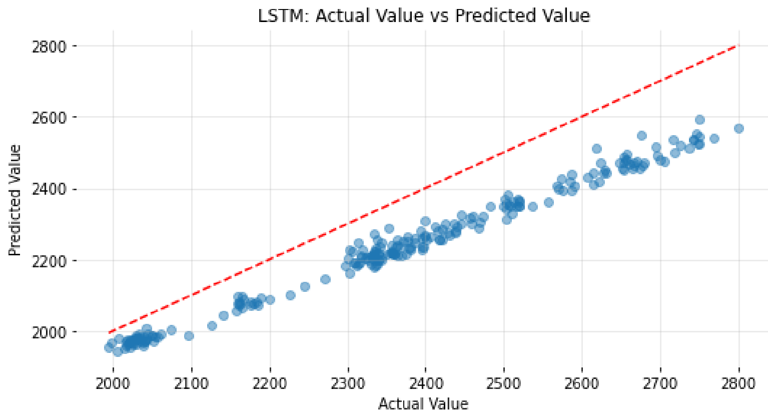

Figure 1. Scatter plot of predicted vs. actual values: LSTM model

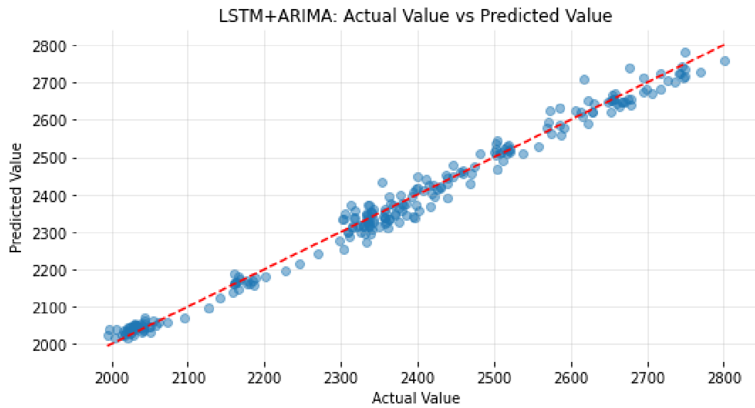

Figure 2. Scatter plot of predicted vs. actual values: hybrid LSTM-ARIMA model

As illustrated in Figure 1, the predictions of the standalone LSTM model exhibit a scattered distribution, markedly deviating from the ideal 45-degree reference line, and consistently underestimate actual prices within higher price ranges. In contrast, predictions from the hybrid LSTM-ARIMA model (as shown in Figure 2) closely cluster around the 45-degree line, indicating strong consistency between predicted and actual values. This visualization result confirms that the forecasting accuracy of the hybrid model is significantly superior to that of the individual LSTM model. Further analysis of the gold price prediction curves, as shown in Figure 3, reinforces this implication.

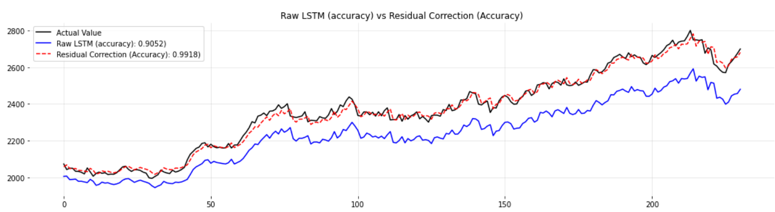

Figure 3. Time series comparison of actual vs. predicted gold prices

Figure 3 demonstrates that the standalone LSTM model (represented by the blue line) systematically underestimates gold prices throughout the entire testing period, with discrepancies between predicted and actual prices (black line) progressively increasing over time. Conversely, the predictions generated by the hybrid LSTM-ARIMA model (red dashed line) closely align with actual

DOI: 10.54254/2755-2721/2025.LD24518

gold prices, resulting in a remarkable improvement in forecasting accuracy from 90.52% to 99.18%. This demonstrates the superior predictive performance of the hybrid model.

3.3. Analysis of error distribution characteristics

To conduct an in-depth analysis of the error characteristics of both models, Figure 4 presents a comparative boxplot visualization of the absolute errors.

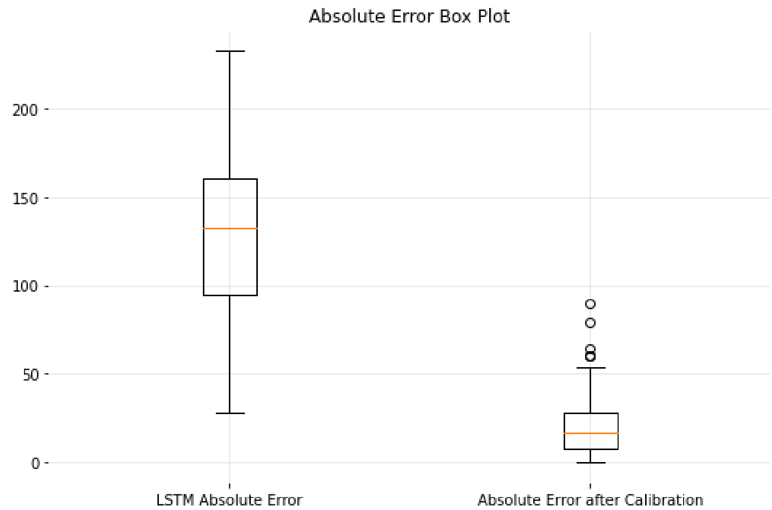

Figure 4. Boxplot comparison of absolute errors between LSTM and hybrid LSTM-ARIMA models

The boxplot clearly illustrates that the median absolute error for the LSTM model is approximately $135, with an interquartile range spanning from $90 to $160. In contrast, the hybrid model exhibits a significantly lower median absolute error of around $20, with the interquartile range notably compressed to between $10 and $30. These results indicate that the hybrid model not only substantially reduces the central tendency of forecasting errors but also significantly decreases error dispersion, demonstrating more precise and stable predictions.

Further insights into the difference in model performance are revealed by examining how forecasting errors evolve (as shown in Figure 5).

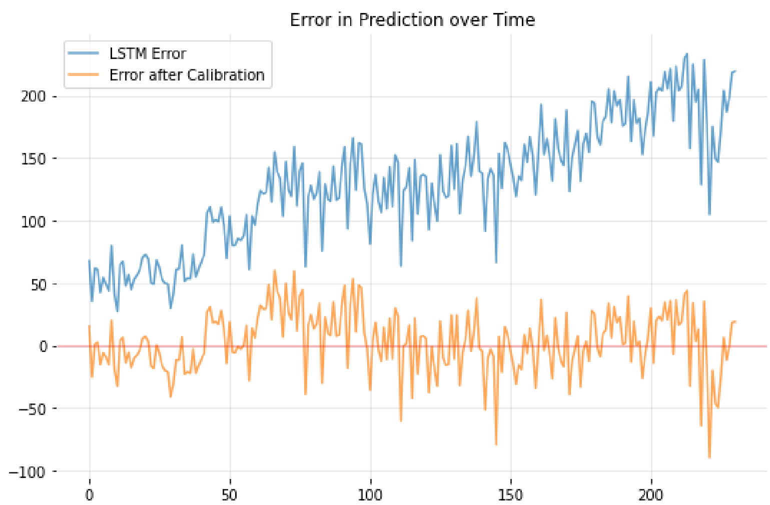

Figure 5. Temporal evolution of forecasting errors

DOI: 10.54254/2755-2721/2025.LD24518

Figure 5 demonstrates that the prediction errors of the LSTM model (blue line) display a clear upward trend over time, with maximum errors exceeding $200 and large fluctuations. In contrast, prediction errors from the hybrid model (orange line) remain consistently low throughout the test period, primarily within ±$50 without noticeable trend bias. This indicates that the hybrid model effectively mitigates the systematic error observed in the standalone LSTM model.

3.4. Cumulative error analysis

In practical financial forecasting applications, cumulative error is a crucial metric for evaluating a model's long-term predictive performance. Figure 6 presents a comparative analysis of cumulative absolute errors for the two models throughout the testing period

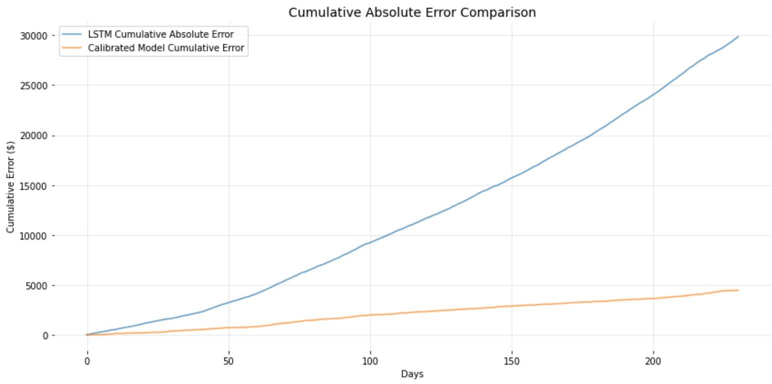

Figure 6. Comparison of cumulative absolute errors

As shown in Figure 6, the cumulative error of the LSTM model (blue line) exhibits a clear linear increasing trend, with the cumulative error approaching approximately $30,000 by the end of the testing period. In contrast, the cumulative error growth rate for the hybrid model (orange line) is significantly lower, with a cumulative error of only around $4,000 by the end of the testing period—equivalent to merely 13.3% of the error observed in the standalone LSTM model. This outcome indicates that the hybrid model possesses substantial advantages in long-term forecasting, effectively mitigating cumulative error effects and providing a more reliable reference for predicting long-term trends in gold prices.

3.5. Analysis of residual correction effectiveness

The core mechanism of the hybrid model is utilizing the ARIMA model to capture linear patterns within the residuals of LSTM predictions. Figure 7 illustrates the comparison between actual LSTM prediction residuals and the residuals predicted by the ARIMA model.

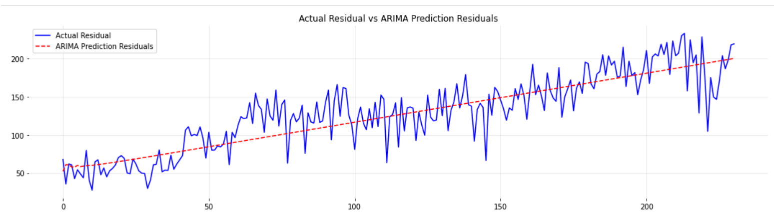

Figure 7. Comparison of actual residuals and ARIMA-predicted residuals

DOI: 10.54254/2755-2721/2025.LD24518

In Figure 7, the blue line represents the actual prediction residuals from the LSTM model, while the red dashed line indicates the residuals predicted by the ARIMA model. It is evident that the ARIMA model successfully captures the primary trends within the residual series, validating the presence of predictable linear patterns in the LSTM residuals. This explains why modeling the residuals using ARIMA can significantly enhance overall forecasting accuracy.

The effectiveness of residual correction is further verified through the analysis of the 20-point moving average errors (Figure 8).

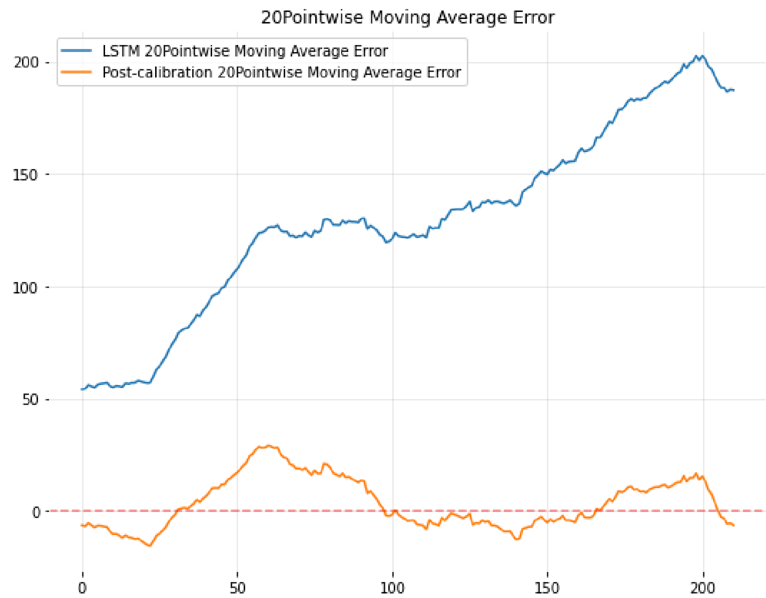

Figure 8. Comparison of 20-point moving average errors

Figure 8 clearly demonstrates that the moving average error of the standalone LSTM model (blue line) exhibits an evident upward trend over time. In contrast, the corrected errors from the hybrid model (orange line) remain consistently stable, closely oscillating around zero without any noticeable trend. This indicates that the ARIMA component successfully eliminates the systematic biases inherent in the LSTM predictions.

The analytical results consistently indicate that the LSTM-ARIMA hybrid model effectively combines the strengths of deep learning and classical time series analysis methods. It successfully mitigates the limitations associated with the standalone LSTM model in gold price forecasting, achieving significantly improved prediction accuracy. Consequently, the hybrid model provides a more reliable quantitative basis for investment decision-making in the gold market.

4. Conclusion

This study proposed a hybrid forecasting model integrating the Long Short-Term Memory (LSTM) neural network and the Autoregressive Integrated Moving Average (ARIMA) model to improve gold price prediction accuracy. Leveraging the complementary strengths of these two methodologies, the hybrid model was designed to effectively capture both nonlinear dependencies and linear characteristics inherent in gold price time series data.

Empirical analysis based on daily gold price data from August 19, 2013, to November 22, 2024, demonstrated that the hybrid LSTM-ARIMA model significantly outperforms the standalone LSTM model across multiple evaluation metrics. Specifically, the hybrid model achieved substantial

DOI: 10.54254/2755-2721/2025.LD24518

improvements in mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE), with reductions of 84.92%, 96.81%, 82.14%, 84.60%, and 85.04%, respectively. Moreover, the determination coefficient (R^2) increased notably from 0.596 to 0.987, indicating a superior explanatory power and predictive capability of the hybrid approach.

Visualization analyses further validated the effectiveness of the hybrid model. Scatter plots illustrated that predictions from the hybrid model closely aligned with actual values, markedly reducing deviations observed in standalone LSTM predictions. Time series comparisons revealed that the hybrid model consistently tracked real gold price fluctuations accurately, effectively correcting the systematic underestimations inherent in the standalone LSTM model. Additionally, analyses of error distribution characteristics indicated that the hybrid model significantly decreased error magnitude and variability, demonstrating stable and precise forecasting performance over time.

The cumulative error analysis highlighted the hybrid model's significant advantage in long-term forecasting scenarios, effectively mitigating cumulative prediction errors. Residual analysis confirmed that the ARIMA component successfully captured and corrected linear trends within LSTM residuals, substantially enhancing overall model performance.

In conclusion, the proposed hybrid LSTM-ARIMA model effectively integrates deep learning's ability to model nonlinear complexities with traditional econometric methods' strength in linear analysis, offering a robust and accurate forecasting tool for gold price prediction. This methodological innovation significantly improves forecasting accuracy and reliability, providing investors, fund managers, and policymakers with a more dependable quantitative basis for strategic decision-making in gold market investments. Future research may explore incorporating additional variables and alternative deep learning architectures to further refine forecasting precision and adaptability.

References

[1] Sharma, Dipasha. (2016). Nexus between financial inclusion and economic growth: Evidence from the emerging Indian economy. Journal of Financial Economic Policy. 8. 10.1108/JFEP-01-2015-0004.
[2] Baur, D.G. and Lucey, B.M. (2010), Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. Financial Review, 45: 217-229. https: //doi.org/10.1111/j.1540-6288.2010.00244.x
[3] Bams, Dennis and Blanchard, Gildas and Honarvar, Iman and Lehnert, Thorsten, Does Oil and Gold Price Uncertainty Matter for the Stock Market? (July 5, 2017). Journal of Empirical Finance, Vol. 44, 2017, Available at SSRN: https: //ssrn.com/abstract=2692987 or http: //dx.doi.org/10.2139/ssrn.2692987.
[4] Shafiee, S., & Topal, E. (2010). An overview of global gold market and gold price forecasting. Resources Policy, 35, 178-189.
[5] Saumendra, Das & Nayak, Janmenjoy & Rao, B. & Kanithi, Vakula & Routray, Ashanta. (2022). Gold Price Forecasting Using Machine Learning Techniques: Review of a Decade. 10.1007/978-981-16-2543-5_58.
[6] Sadorsky, Perry. (2021). Predicting Gold and Silver Price Direction Using Tree-Based Classifiers. Journal of Risk and Financial Management. 14. 198. 10.3390/jrfm14050198.
[7] Tunnicliffe Wilson, Granville. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. Journal of Time Series Analysis. 37. n/a-n/a. 10.1111/jtsa.12194.
[8] Zhang, Peter. (2003). Zhang, G.P.: Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Neurocomputing 50, 159-175. Neurocomputing. 50. 159-175. 10.1016/S0925-2312(01)00702-0.
[9] D. M. Q. Nelson, A. C. M. Pereira and R. A. de Oliveira, "Stock market's price movement prediction with LSTM neural networks, " 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 2017, pp. 1419-1426, doi: 10.1109/IJCNN.2017.7966019.
[10] Maknickas A, Maknickiene N. Support System for Trading in Exchange Market by Distributional Forecasting Model. Informatica. 2019; 30(1): 73-90. doi: 10.3233/INF-2019-1201.
[11] S. Siami-Namini, N. Tavakoli and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series, " 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA,

DOI: 10.54254/2755-2721/2025.LD24518

2018, pp. 1394-1401, doi: 10.1109/ICMLA.2018.00227.
[12] Zhang, Peter & Qi, Min. (2005). Neural network forecasting for seasonal and trend time series. European Journal of Operational Research. 160. 501-514. 10.1016/j.ejor.2003.08.037.
[13] Achkar, R., Elias-Sleiman, F., Ezzidine, H., & Haidar, N. (2018). Comparison of BPA-MLP and LSTM-RNN for Stocks Prediction. 2018 6th International Symposium on Computational and Business Intelligence (ISCBI), 48-51.
[14] Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. Neural Comput. 2000 Oct; 12(10): 2451-71. doi: 10.1162/089976600300015015. PMID: 11032042.
[15] Hyndman, R. J., & Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software, 27(3), 1–22. https: //doi.org/10.18637/jss.v027.i03.
[16] Lai, G., Chang, W., Yang, Y., & Liu, H. (2017). Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks. The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.