Performance Comparison of Gold Price Forecasting Models and Effectiveness of Hybrid Methods

Honghao Dan *

Shenzhen Audencia Financial Technology institute, Shenzhen University, Shenzhen, China

* Corresponding Author Email: 2023361010@email.szu.edu.cn

Abstract. Gold has long been regarded as an important risk-averse asset and an investment tool. Predicting its price trend has been a challenging topic for financial time series analysis. This study constructs six models, including artificial neural network (ANN), Long Short-Term memory (LSTM), bidirectional long Short-Term memory (Bi-LSTM) and their different hybrid strategies with generalized autoregressive conditional heteroscedastic (GARCH) model in gold price forecasting, and systematically compares their performance, based on the daily data between 2015/10/10 and 2025/10/10. Through multiple rounds of random seed experiments, the performance of each model is evaluated by the Root Mean Square error (RMSE), Mean Absolute error (MAE), Mean Absolute percentage error (MAPE), $R^2$and trend prediction accuracy. The results indicate that the ANN model is the most robust in comprehensive performance. The further analysis shows that simply introducing GARCH volatility in a linear way or using bi-directional structure (Bi-LSTM) may reduce the stability of the model. However, adopting the non-linear fusion method by combining the point prediction from the LSTM and the volatility features from GARCH (LSTM-ANN) shows potential in fitting volatility and trend. The study reveals the characteristics of different models in gold price prediction, provides a basis for model selection, and points out that the introduction of external macroeconomic and news factors is a key direction for improving forecasting performance in the future.

Keywords: Gold price prediction; Deep learning; Long Short-Term Memory; Generalized Autoregressive Conditional Heteroskedasticity; Hybrid model.

1. Introduction

Gold price fluctuates due to the effect of economic condition, policy and geopolitics, showing a high degree of non-linearity and non-stationarity, which makes the accurate prediction of gold price challenging and valuable.

In recent years, machine learning and deep learning techniques surpassed the traditional statistical methods and become more popular in financial time series prediction. From Autoregressive Integrated Moving Average Model (ARIMA) to Artificial Neural Network (ANN), Long Short-Term Memory (LSTM) and various hybrid models, there are an increasing number of models available for selection. Recently, Ranadhan et al. successfully verified the effectiveness of the hybrid method integrating LSTM-ANN network and GARCH model in gold price prediction, which provide a valuable idea for accurate prediction model [1]. Nevertheless, there is still a lack of systematic comparison of the comprehensive performance of these hybrid models and classic benchmark models (such as pure ANN, LSTM, etc), as well as the effectiveness of different GARCH integration strategies.

Inspired by this, this research aims to achieve the following goals through systematic experiments: First, to comprehensively evaluate the performance of hybrid models; Second, to explore the advantages and limitations of different models in dealing with time series characteristics; Third, to verify the effectiveness of different GARCH integration strategies.

This article is structured in following order: Chapter Two is a literature review; Chapter Three elaborates on the research methods; Chapter Four presents and analyzes the experimental results; Chapter Five summarizes the entire article and looks forward to the future.

2. Literature Review

2.1. The Adoption of Traditional Statistical Method in Gold Price Prediction

Early research on gold price prediction mainly depended on traditional statistical models. The ARIMA model adopted by BOX and JENKINS is a classic linear prediction method for time series, which makes the sequence stationary through differencing and models it using autoregressive and moving average terms[2]. However, ARIMA assumes a constant sequence variance and cannot effectively capture the common volatility clustering phenomenon in financial time series data. Therefore, the ARCH/GARCH model proposed by Engle and Bollerslev has been widely used to characterize and predict time-varing variances, becoming an essential tool in financial econometrics[3,4]. These models have a solid theoretical foundation but are insufficient in capturing complex nonlinear relationships.

2.2. Artificial Neural Network Model

Artificial Neural Network (ANNs), through mimicking the connection method of human brains, is equipped with strong ability to find complex patterns, which allow it to skip the strict assumption of the distribution of the input data. Early study, such as Mombeini and Yazdani-Chamzini's work, has proved that the ANN significantly performances better than ARIMA models when predicting the trend of the gold price [5]. As one of the basic models, ANNs are often used as benchmark for the models performance comparison.

2.3. The Development and the Adoption of Deep Learning Models

Deep learning models, especially the recurrent neural networks (RNNs), change how we analyze sequential data like gold price. However, the basic RNNs were faced with the problem of vanishing and exploding, which prevented them from extracting long-term dependencies. The Long Short-term memory proposed by Hochreiter and Schmidhuber has effectively solved the problem by introducing gating mechanisms (input gate, forget gate, and output gate) [6]. This advantage makes LSTM one of the optimal models for time series prediction. After that, more variants are proposed, such as Bidirectional LSTM (Bi-LSTM), which are capable of capture the information from both past and future for prediction, and the Gated Recurrent Unit, which simplifies the structure of LSTM with fewer parameters. Furthermore, the introduction of the attention mechanisms allows models to dynamically capture the more important part within the input series, which further enhances the model performance.

Recently, more studies have been conducted to deeply explore these models. Yurtsever compared LSTM, Bi-LSTM and GRU for gold price forecasting using economic indicators, like crude oil price and consumer price index, as features, which found that LSTM worked the best among the three [7].

2.4. Development in Hybrid Models

With the thought of combining the advantages of each model, in recent years, more and more studies have been designed to explore the efficient combination. Researchers typically blend models in two ways: either mixing different deep learning architectures or pairing deep learning with traditional statistical methods. This dual approach aims to capture both overall price trends and short-term volatility patterns.

Several studies show the promise of this approach. Santika's team found that a combined CNN-LSTM model beat either model alone for daily gold price prediction [8]. Risse used wavelet decomposition with support vector regression to analyze gold returns from both time and frequency perspectives, achieving better forecasts [9]. Bao and colleagues went further, building a system that integrates wavelet transforms, autoencoders, and LSTM to extract cleaner features from noisy financial data [10].

Beyond these, Singh and Varshney showed that ensemble methods like Random Forest can effectively identify complex patterns in gold prices [11]. Meanwhile, Garcia-Medina's work on cryptocurrency volatility demonstrated that combining LSTM with GARCH models is feasible [12]. Most relevant to our work, Ramadhan and colleagues successfully merged LSTM-ANN networks with GARCH for gold price forecasting - a finding that directly inspired our own hybrid model experiments [1].

The exploration in the hybrid model continues to expand these two paths. Noveldata sources are being adopt, as seen in the work of Kian Peer et al., who combined Google Trends and mied news sentiment using a CNN model [13]. Meanwhile, the latest research is evolving to integrate cutting-edge components. For instance, Yao et al. developed a Transformer-LSTM quantile regression model, demonstrating the advantage of Blending advanced neural networks for capturing complex financial dependencies [14]

While these studies confirm that smart model combinations can improve predictions, we still lack a clear comparison of which hybrid strategies work best. Our study aims to fill this gap by systematically testing different fusion approaches under the same conditions.

3. Methodology

3.1. Description on the Data

The gold price data resources of this study are investing.com (hettps: // cn. investing.com/ currencies/ xau-usd), covering a ten-year period from October 10, 2015, to October 10, 2025. There are 2,604 trading days in total within the period. The data set was divided into a training set (80%, 2033 samples) and a test set (20%. 509 samples) in time series order to maintain the natural time sequence. This approach preserves the real-world flow of market information and gives us a fair assessment of how each model wold perform in practice.

Feature variables: This study selected five key price features as model inputs:

- Open price of the day (opent)
- High price of the day (hight)
- Lowest price of the day (lowt)
- Closing price of the day (closet)
- Previous day's close (closet-1, lag feature)

To prepare the data for modelling, this study applied two key steps. First, we used Min-MAX normalization to scale the features between 0 and 1. This step is crucial for stabilizing the training process and helping the models converge. Next, the sequential data using a sliding window approach was structured. Here, each input sample packs the market data from the previous 60 days, creating a 60*5 matrix that the model uses to forecast the following day’s close. A 60-day window was set due to the initial tests indicated ti was long enough to capture meaningful medium-term trends in gold volatility and learn the patterns of short-term gold price fluctuations.

3.2. Model Structures

In this study, six deep learning prediction models were constructed and compared, with the parameters of each model being strictly controlled between 93,000 and 103,000, ensuring the fairness and comparability of the comparison experiments

3.2.1. Baseline Models

ANN: A feedforward neural network architecture is used as a baseline model to evaluate the basic nonlinear mapping capability. The detailed structure is shown in Table 1.

Table 1. Structure of ANN model.

Layer (type)	Output Shape
dense_9 (Dense)	(None, 200)
dense_10 (Dense)	(None, 120)
dense_11 (Dense)	(None, 60)
dense_12 (Dense)	(None, 30)
dense_13 (Dense)	(None, 1)

LSTM: A standard non-directional LSTM architecture specifically designed to capture the Long-term dependencies and temporal patterns in time series. The detailed structure is shown in Table 2.

Table 2. Structure of LSTM model.

Layer (type)	Output Shape
lstm (LSTM)	(None, 60, 88)
lstm_1 (LSTM)	(None, 88)
dense (Dense)	(None, 64)
dense_1 (Dense)	(None, 32)
dense_2 (Dense)	(None, 1)

Bi-LSTM: It leverages both past and future temporal information through a bi-directional architecture (using backward passing at training time), Theoretically capable of providing richer content. The detailed structure is shown in Table 3.

Table 3. Structure of Bi-LSTM model.

Layer (type)	Output Shape
bidirectional (Bidirectional)	(None, 60, 110)
bidirectional_1 (Bidirectional)	(None, 110)
dense (Dense)	(None, 28)
dense_1 (Dense)	(None, 6)
dense_2 (Dense)	(None, 1)

3.2.2. Volatility Enhanced Models

LSTM-GARCH / Bi-LSTM-GARCH: These two models both implement linear correction after the point prediction from the basic models. Detailed progress is described below:

LSTM and Bi-LSTM generate a benchmark price prediction.

GARCH (1,1) model is then fitted to the prediction residual sequence, and the conditional volatility is extracted.

Linear regression is used to establish the "forecast error-volatility" relationship and apply posterior correction to the benchmark forecast.

The mathematical formula is as follows:

$$\hat{y}_{\text{LSTM}} = f_{\text{LSTM}}(\mathbf{x}) \tag{1}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2}$$

$$where\ \epsilon = y_t - y_t \tag{3}$$

$$adjustment_t = \beta_0 + \beta_1 * \sigma_t \tag{4}$$

$$\hat{y}_{\text{final}} = \hat{y}_{\text{LSTM}} + \text{adjustment}_t \tag{5}$$

ANN-LSTM (Non-linear Fusion Model): It adopts a nonlinear deep fusion strategy, and its architecture is based on the LSTM-ANN network idea proposed by Ramadhan et al. [1]. This model aims to verify the advantages of feature-level nonlinear fusion over prediction-level linear correction. The detailed structure is shown in Table 4.

Table 4. Structure of LSTM-ANN model.

Layer Branch	Layer (type)	Output Shape
LSTM	lstm_input (Input)	(None, 60, 5)
	lstm (LSTM)	(None, 60, 80)
	lstm_1 (LSTM)	(None, 80)
	dense (Dense)	(None, 48)
	lstm_output (Dense)	(None, 1)
GARCH	garch_residuals	(None,)
	garch_volatility	(None,)
ANN	ann_input (Input)	(None, 7)
	dense_1 (Dense)	(None, 48)
	dense_2 (Dense)	(None, 24)
	ann_output (Dense)	(None, 1)
Fusion	concatenate (Concatenate)	(None, 2)
	dense_3 (Dense)	(None, 80)
	dense_4 (Dense)	(None, 80)
	dropout (Dropout)	(None, 80)
	dense_5 (Dense)	(None, 80)
	final_prediction (Dense)	(None, 1)

3.3. Experiment Setup

Adam optimizer, along with mean square error loss function and early stopping strategy was adopted. Ten different random seeds were used to train and test each model independently. All the final performance index will be presented in form of mean ± standard deviation.

4. Result and Analysis

4.1. Comprehensive Comparison of Models Performance

Table 5 shows the comprehensive performance of the six models of over 10 independent runs.

Table 5. Comparison of prediction performance of six deep learning models (ranked by RMSE).

Models	RMSE ($)	R^2	MAPE (%)	MAE ($)	Accuracy (%)	Parameters
ANN	157.32 ± 54.54	0.899 ± 0.063	4.07 ± 1.32	119.74 ± 41.01	32.24 ± 11.49	93,441
Bi-LSTM	236.71 ± 57.06	0.784 ± 0.106	5.43 ± 1.42	168.47 ± 43.70	30.24 ± 7.00	103,169
LSTM	244.21 ± 112.28	0.737 ± 0.258	5.95 ± 3.18	181.92 ± 94.49	29.35 ± 10.34	103,201
ANN-LSTM	255.86 ± 85.25	0.735 ± 0.144	7.96 ± 2.91	220.67 ± 82.44	14.56 ± 18.59	97,843
Bi-LSTM-GARCH	289.16 ± 115.68	0.647 ± 0.295	6.97 ± 3.47	213.39 ± 99.72	27.03 ± 14.66	103,169
LSTM-GARCH	321.15 ± 203.54	0.474 ± 0.666	8.50 ± 6.19	254.37 ± 179.45	20.14 ± 12.88	103,201

From table 5, the paper can conclude three key discoveries:

1)Significant differences in model stability:

ANN has the best stability and relatively small standard deviation. LSTM has large fluctuations; The GARCH enhanced model is extremely unstable, and the RMSE standard deviation of LSTM-GARCH reaches $203.54.

2)Negative effects of GARCH linear correction:

Compared with LSTM, the RMSE of LSTM-GARCH increases by 31.5%, and the R^2 decreases by 35.7%. Compared with Bi-LSTM, the RMSE of Bi-LSTM-GARCH increases by 22.1% and the R^2 decreases by 17.5%.

3) The relative success of ANN-LSTM

Compared with LSTM-GARCH, the RMSE of ANN-LSTM reduces by 20.3%, and the R^2 improves by 55.1%. Compared with Bi-LSTM-GARCH, the RMSE of ANN-LSTM reduces by 11.5% and the R^2 improves by 13.6%.

In conclusion, ANN delivers the best overall performance and robustness in this study, while the linear integration strategy for GARCH fails to deliver the anticipated benefits

4.2. Detailed Analysis for each Model

The metrics from table only tell part of the prediction performance. To relly understand how each model behaves, next section discusses their prediction visually. In this section, we examine prediction curves (Figs 1-6) that show how well each model follows actual gold price movements. The paper also analyzes residual plots that reveal systematic biases, and scatter plots that display error patterns. These visual tools help us see the practical strengths and weaknesses that pure numerical scores might miss.

4.2.1 ANN Model

The ANN model shows excellent nonlinear fitting ability and excellent stability, and the feedforward structure gives it efficient deployment characteristics. ANN ranks first in all four core indicators, and its RMSE is only \$157.32, 33.5% lower than the second place Bi-LSTM. The R-squared reaches 0.899 and the MAPE is 4.07%. In the multi-seed test, ANN also shows the best stability with the smallest standard deviation among all the models. However, the ability of ANN to capture short-term sharp fluctuations is relatively insufficient, and the response lags when the market has sudden large fluctuations.

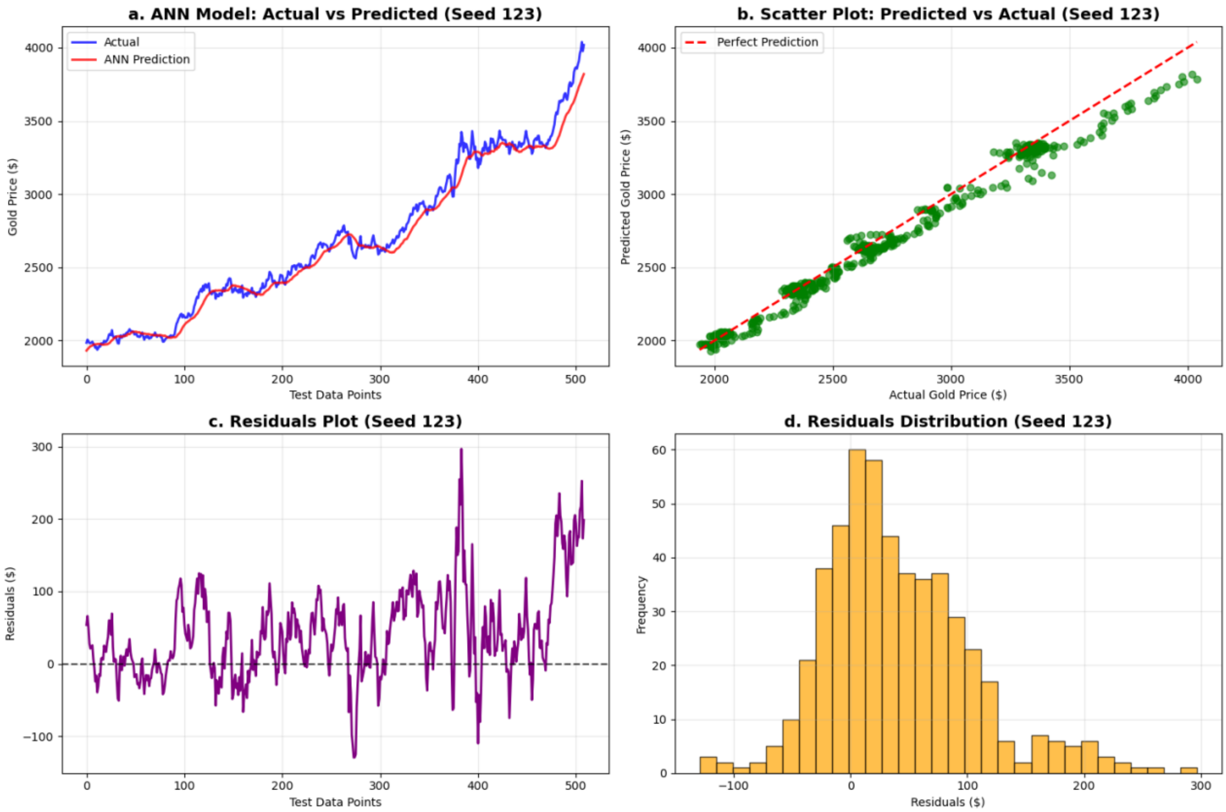

Fig 1. Visual Comparison of ANN Model Performances.

4.2.2. LSTM and Bi-LSTM

LSTM has moderate performance but poor stability. Bi-LSTM has good comprehensive performance and better stability than unidirectional LSTM. Both models show systematic forecast a bias later in the test set, indicating the limitations of purely technical models in responding to external shocks.

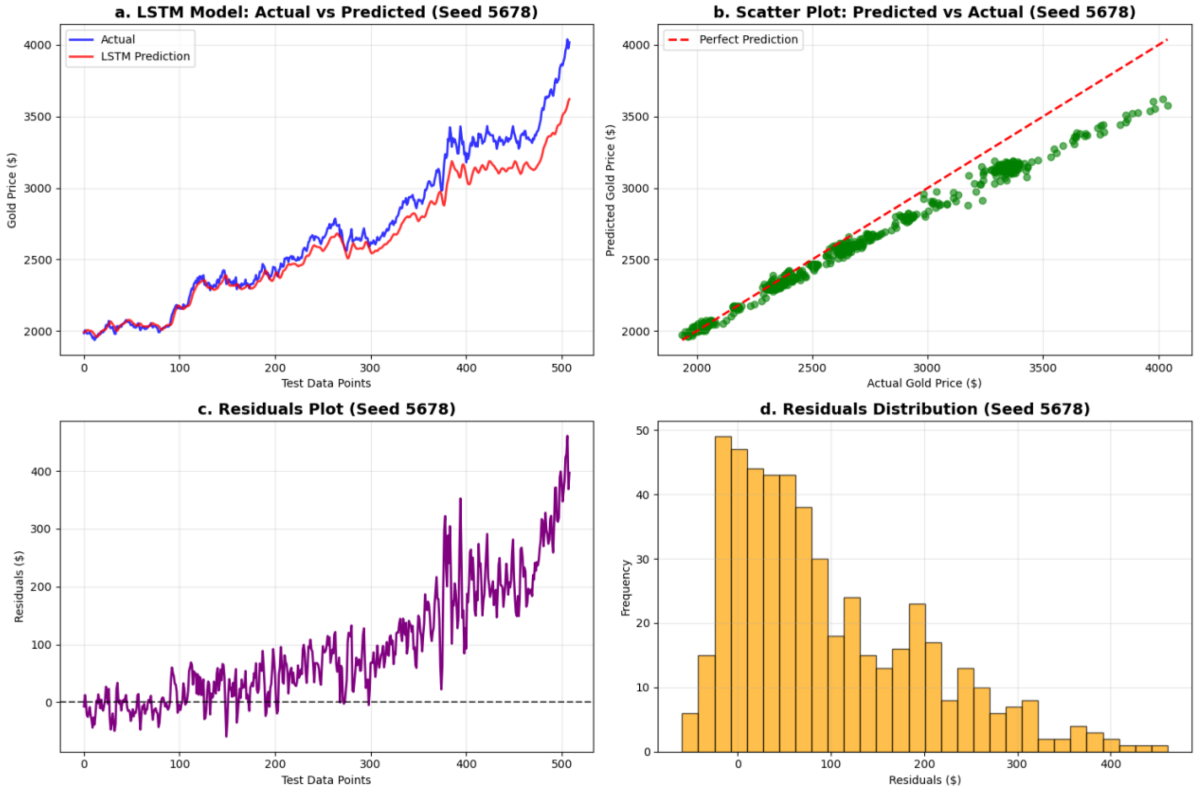

Fig 2. Visual Comparison of LSTM Model Performances

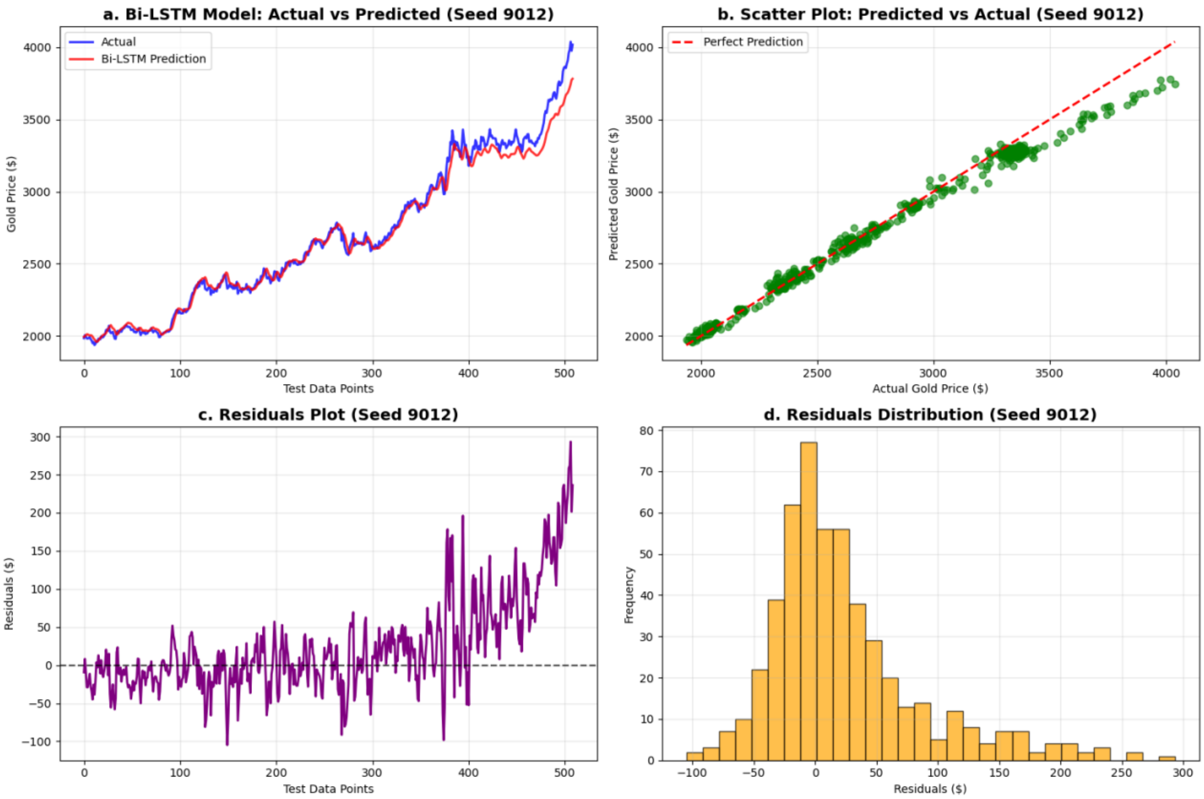

Fig 3. Visual Comparison of Bi-LSTM Model Performances.

4.2.3. GARCH Enhanced Models

The performance of the two GARCH enhanced models deteriorated significantly and their stability decreased substantially. The reason for this failure is that the simple linear regression correction assumption fails to capture the complex relationship between forecast errors and volatility. The volatility extracted by GARCH model itself may contain noise, and the linear correction instead amplified this noise, further increasing the systematic bias mentioned above.

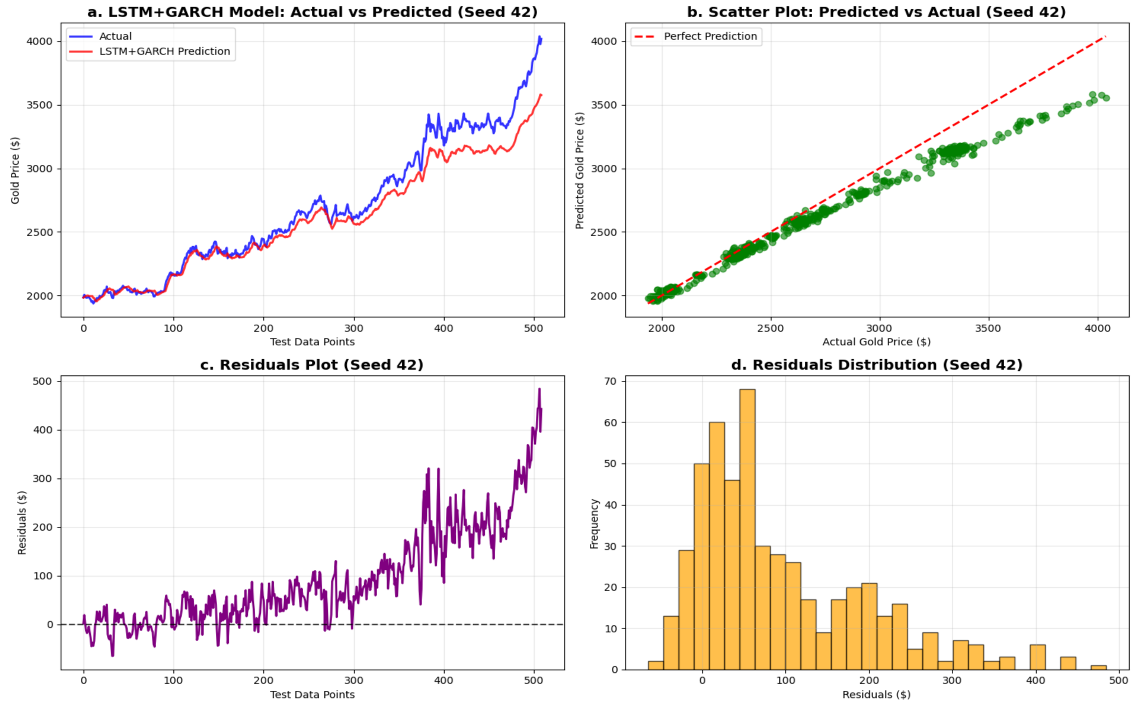

Fig 4. Visual Comparison of LSTM-GARCH Model Performances.

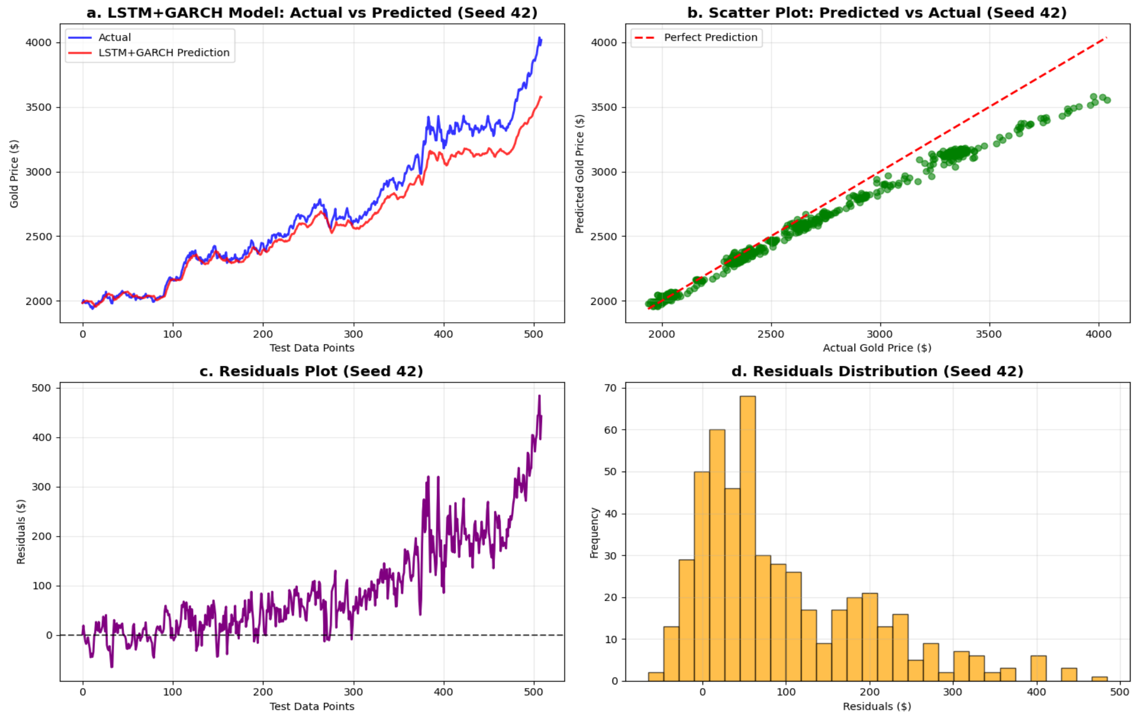

Fig 5. Visual Comparison of Bi-LSTM-GARCH Model Performances.

4.2.4. ANN-LSTM

ANN-LSTM performs better than the linear GARCH correction model, which proves the advantage of feature-level nonlinear fusion and to some extent reduces the systematic bias that appears in the late test. However, even with nonlinear fusion, the performance was still inferior to that of the pure ANN model, possibly because the fusion architecture needed further optimized, the contribution of GARCH features was limited, and the risk of overfitting increased due to the increasing complexity of the model.

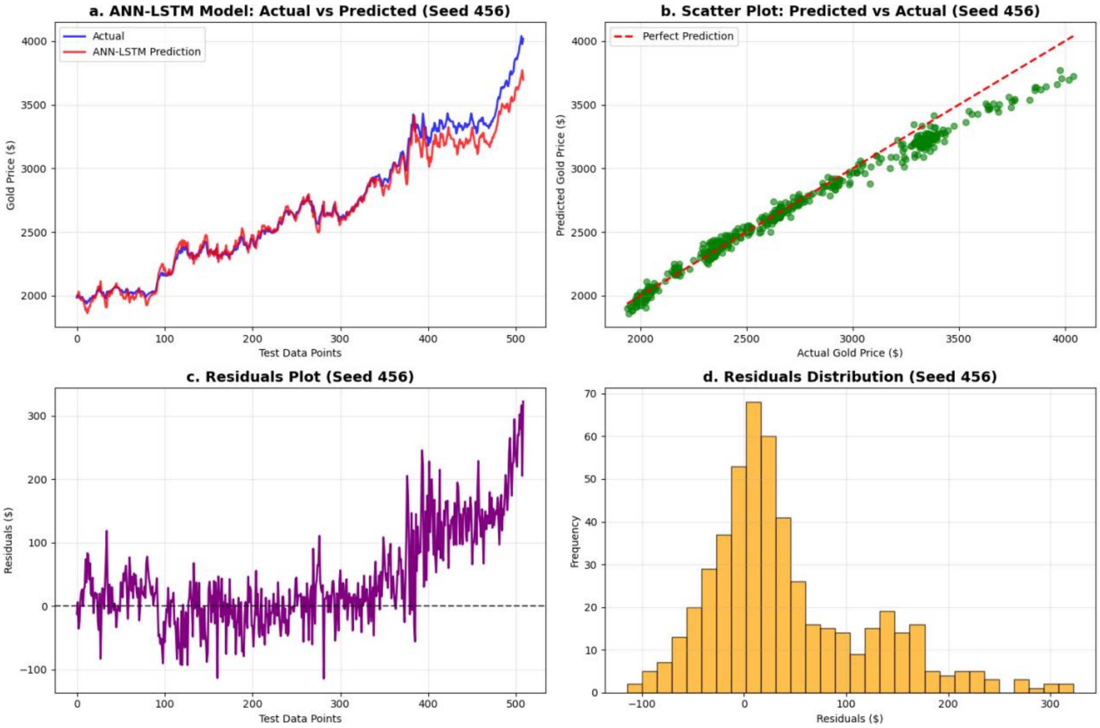

Fig 6. Visual Comparison of ANN-LSTM Model Performances.

5. Discussion

5.1. Underlying Reasons for the Performance Difference among the Models

The most surprising finding in this study is that, even with the simplest architecture, the ANN model surpassed other models with more complex models across all dimension. This unexpected outcome can be explained by the feature representation sufficiency hypothesis: after being flattened through a 60-day time window, we believe that the five features used in this study already implicitly contained sufficient temporal information, which was totally capable of providing enough data to learn primary price change patterns through nonlinear mapping.

From the perspective of matching model capacity with task complexity, with only the technical features, the learnable patterns of gold prices may be relatively limited. In this case, there is probability that the parameters set is sufficient for ANN to fit these patterns, while it is excessive for LSTM. Excessive model capacity may cause the model to learn more noise from the training data, leading to overfitting, especially when there is a lack of valid information. This finding aligns with the principle of Occam's razor: when prediction accuracy is similar, simpler models generally exhibit better generalization capabilities.

5.2. The Implications and Relative Advantages of the Nonlinear GARCH Fusion Strategy

Due to the limited feature set, the marginal benefit of the valid information within the GARCH volatility extracted from the time series price data was reduced. However, an important implication can be drawn from the relative success of ANN-LSTM model: the method of feature fusion determines whether the model "amplifies noise" or "extracts signals"

The improvement of ANN-LSTM does not prove that the GARCH volatility is a strong feature but indicates that a robust fusion architecture with nonlinear learning ability is able to extract residual and weak useful information from the feature with noise to some extent, while suppressing their negative impact. This provides a methodological reference for future integration of weakly correlated features in information constrained environments.

The effect of GARCH on LSTM and Bi-LSTM after linear adjustment can be referred to Fig 7.

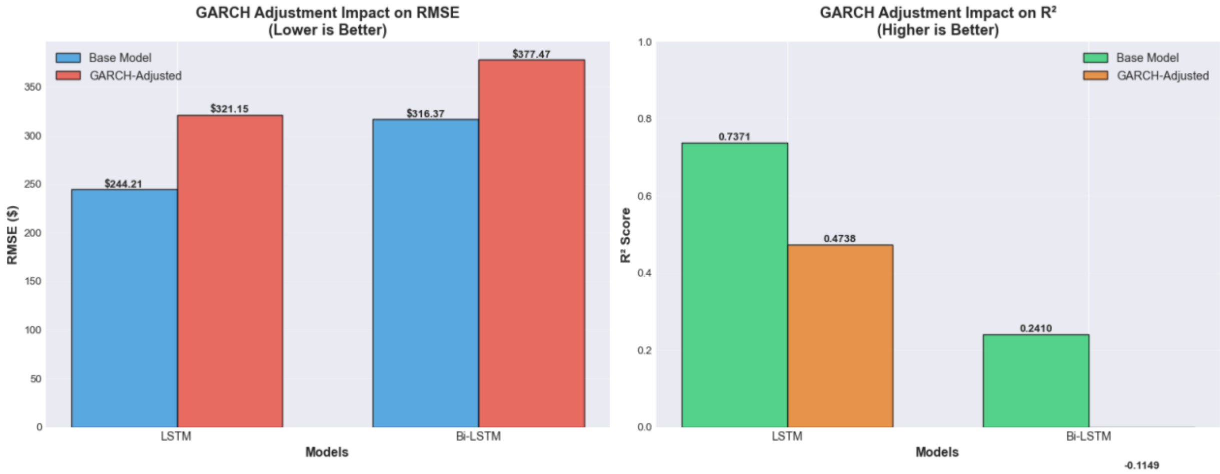

Fig 7. Comparison between GARCH enhance models and Base models.

6. Challenge and Prospect

6.1. Limitations

Despite efforts to ensure rigor, this study has the following limitations:

Feature Limitations: This study only contains five technical features and lacks external factors that have a significant impact on the gold price, such as macroeconomic indicators, geopolitical factors, market sentiment, and related asset prices.

Time range limitation: The test set mainly covers the period from 2024 to 2025, which coincided with many extreme market fluctuations (e.g. US election, escalation of geopolitical conflicts). The performance of the models in extreme market environments needs to be verified over a longer period, and their adaptability to different market cycles (bull/bear/turbulence) was not evaluated.

Model limitation: This study only compares traditional ANN, LSTM, Bi-LSTM and their simple combinations, and only two GARCH integration methods are adopted. The cutting-edge architectures such as Transformer series, time-specific architectures and neural architecture search are not involved in this study.

6.2. Future Direction

Based on the findings and limitations of this study, future research can focus on several key directions. Firstly, feature engineering could be enhanced by integrating multi-source data, including macroeconomic indicators, geopolitical factors, market sentiment, and related asset prices, as demonstrated in prior multi-variable approaches [7,10]. The incorporation of alternative data sources, such as web search trends and news sentiment, has also been shown to provide valuable predictive signals for gold prices [13].

Secondly, regarding model design, more advanced methods for volatility modeling and ensemble techniques should be explored. Rather than treating GARCH volatility as a simple linear input, future work could develop advanced fusion mechanisms, for example, by using attention mechanisms or custom neural network layers to selectively incorporate volatility information, building upon existing feature fusion and hybrid modeling ideas [8,12].

Thirdly, exploring cutting-edge architecture remains a promising frontier. This includes investigating models like Transformers to capture long-range dependencies, as well as designing advanced hybrid architectures that combine different neural network families (e.g., Transformers with RNNs) or integrate signal processing techniques such as wavelet decomposition [9]. The demonstrated success of architectures like Transformer-LSTM in related financial forecasting tasks underscores the potential of such complicated integrations [14].

Finally, to ensure practical robustness, future studies should emphasize cross-market and cross-cycle validation. Following approaches that test models across diverse market conditions, it is crucial to specifically evaluate performance during distinct phases such as bull markets, bear markets, and high-volatility periods [10]. This would give us much better insight into real-world reliability.

7. Conclusion

This study rigorously tested six deep learning models for gold price forecasting, including basic architecture and several hybrid combinations. To ensure fair comparison, all models were set similar in size (93,000-103,000 parameters) and ran each one 10 times with different random seeds to measure stability.

First, with sufficient feature engineering, simple model frameworks can surpass complex recurrent neural networks. Under adequate temporal feature engineering, the structurally simple Feedforward Neural Network (ANN) comprehensively outperformed all more complex LSTM and their variants in both prediction accuracy (RMSE: $157.32) and stability. This indicates that when historical price information is sufficiently encoded through sliding windows, the nonlinear capability of ANN is adequate to capture the primary price patterns, whereas excessive model parameters might lead to overfitting without yielding performance gains.

Second, nonlinear fusion is an effective path for extracting weak signals from noise. Simple linear correction using GARCH volatility severely degraded model performance (LSTM-GARCH's RMSE increased by 31.5%), proving that the volatility extracted using a limited features set contains significant noise. However, the relative success of the model using ANN for nonlinear fusion (ANN-LSTM) demonstrates that a fusion architecture with nonlinear learning capability can, to some extent, "purify" and softly utilize this information, performing significantly better than the crude linear correction method. This finding points toward a future direction for improvement: when perfect external features are unavailable, designing more sophisticated fusion mechanisms with strong noise resistance (such as attention mechanisms) is a feasible path to enhancing model performance. While this does not change our core conclusion that the feature set was insufficient, it does reveal a path for future development: by designing smarter models that can extract value from noisy or weak features, rather than being hampered by them.

Third, it is important to test model stability. Multi-seed testing in this study revealed crucial differences in model reliability. ANN proved remarkably stable, winning 8 out of 10 rounds with minimal performance variation. LSTM, however, showed wildly different results depending on the random seed - its error margins fluctuated by nearly half its average value. This means that running a model just once can be deeply misleading, and thorough testing is essential for trustworthy comparisons.

Our findings clearly show that complex models aren't always better. For anyone building a gold price prediction system, we recommend starting with a simple ANN model, which proved both accurate and stable, rather than defaulting to a more complex LSTM. The real performance gains likely lie not in model complexity, but in incorporating better external data and smarter fusion techniques, like attention mechanisms. In the unpredictable world of finance, a simple, reliable model is often the most powerful choice.

References

[1] Ramadhan T, Wahanani H E, Aditiawan F P. Gold Price Prediction Using Hybrid Deep Learning by Integrating LSTM-ANN Network with GARCH Model. Literasi Nusantara, 2025, 5(3): 315-328.

[2] Box G E P, Jenkins G M. Time Series Analysis: Forecasting and Control. Holden-Day, 1976.

[3] Engle R F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 1982, 50(4): 987-1008.

[4] Bollerslev T. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 1986, 31(3): 307-327.

[5] Mombeini H, Yazdani-Chamzini A. Modeling Gold Price via Artificial Neural Network. Journal of Economics, Business and Management, 2015, 3(7): 699-703.

[6] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation, 1997, 9(8): 1735-1780.

[7] Yurtsever M. Gold Price Forecasting Using LSTM, Bi-LSTM and GRU. European Journal of Science and Technology, (31), 341-347. 2021.

[8] Santika I W K G, Saadah S, & Yunanto P E. Gold price prediction using convolutional neural network-long short-term memory (CNN-LSTM). Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 6(3), 231-238. 2021.

[9] Risse M. Combining wavelet decomposition with machine learning to forecast gold returns. International Journal of Forecasting, 35(2), 601-615. 2019.

[10] Bao W, Yue J, & Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLOS ONE, 12(7), e0180944. 2017.

[11] Singh N, & Varshney J. Artificial Intelligence-Driven Model for Gold Price Prediction. International Journal of Scientific Research in Engineering and Management, 8(5). 2024.

[12] García-Medina A, Aguayo-Moreno E. LSTM-GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. Computational Economics, 2024, 63(1): 1511–1542.

[13] Kian Poor S., Fattahi S., Hajian M. Gold Price Forecasting: A Novel Approach Based on Text Mining and Big-Data-Driven Model. Studies in Business and Economics. 2024, 19(3): 157-171.

[14] Yao Yinhong, Wang Xiaoxu, Chen Wei, Chen Zhensong. Extreme Risk Spillover among Global Stock Markets Based on Transformer-LSTM Quantile Regression. Chinese Journal of Management Science. 2025, 33(8): 1-13.