

رائد التتان

Hybrid LSTM%E2%80%93XGBoost Model for Gold Price Forecasting.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:338029593

Submission Date

Dec 17, 2025, 10:00 PM GMT+2

Download Date

Dec 17, 2025, 11:27 PM GMT+2

File Name

Hybrid LSTM–XGBoost Model for Gold Price Forecasting.docx

File Size

916.5 KB

23 Pages

6,074 Words

37,072 Characters

49% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups

-  **33 AI-generated only 49%**
Likely AI-generated text from a large-language model.
-  **0 AI-generated text that was AI-paraphrased 0%**
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.





Hybrid LSTM–XGBoost Model for Gold Price Forecasting

Prepared By

Name:	Raed Abdalnaser Tattan
ID No.:	243000058

Submitted in fulfillment of the requirements for the degree Master of Science in Artificial Intelligence

At

College of Informatics
Midocean University
UAE

Supervisor: **Dr. Wessam Ahmed**
College of Informatics, Midocean University

Acknowledgements

I would like to thank my supervisor Dr.Wesam Ahmed for her inspiration, patience, and encouragement, your insight and mastery aided me at every point of this project.

Abstract

Forecasting of the price of gold continues to be a difficult problem because of the high volatility and nonlinear dynamics of global financial markets. The abrupt regime changes, macroeconomic instability, and intricate interactions of financial components make accurate prediction difficult to achieve with conventional econometric approaches for predictive precision. In consequence, all these drawbacks have led to the use of machine learning and hybrid artificial intelligence models to approach such behavior more correctly; hence their ability to describe nonlinear behaviors and dynamics. This thesis has introduced an integrated hybrid forecasting framework using XGBoost with a residual-based Long Short-Term Memory (LSTM) model. In the methods proposed, XGBoost is utilized initially to model nonlinear relationships between designed price-dependent features and a robust prediction. Afterwards, an LSTM network then learns the remaining residual errors from the XGBoost model, capturing all the residual time dependencies, which are not completely captured by the baseline model. The proposed framework's performance is compared to standalone XGBoost and LSTM models on standard error metrics (RMSE and MAE) and directional accuracy. The experimental results show that the numerical forecasting errors show its lower behavior using XGBoost, while the residual-based hybrid model retains competitive performance and high directional accuracy. This implies that hybrid architecture can improve robustness for prediction of trends, provided the numerical error reductions are limited. Overall, we also look more into how a hybrid AI model might be used to forecast gold prices in volatile market conditions.

List of Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
XGBoost	Extreme Gradient Boosting
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
USD	United States Dollar

List of Figures

Figure No	Title
Figure 2.1	Gold Price Chart (2005–2025)
Figure 4.1	Research framework of the proposed hybrid model
Figure 4.2	Hybrid LSTM–XGBoost forecasting pipeline
Figure 5.1	Actual vs. Predicted Gold Prices using the XGBoost Model on the Test Set
Figure 5.2	Training and validation loss curves for the residual LSTM model
Figure 5.3	Actual vs. Predicted Gold Prices using the residual-based hybrid model on the test set

List of Tables

Table	Title
Table 3.2.1	Summary of Machine Learning-Based Studies for Gold Price Forecasting.
Table 3.3.1	Summary of LSTM-Based Studies for Gold Price Forecasting
Table 5.1	Comparative Performance of the Forecasting Models on the Test Set

Chapter 1: Introduction

Gold is among the most important resources in world financial markets [1]. Its price can be affected by macroeconomic and financial factors ranging from inflation prospects to currency exchange rates, interest rates, geopolitical tensions, and changes in global risk perception. As a result, accurate estimation of gold's price is one of the central goals of government policymakers as well as financial institutions — a fundamental task of investors. Nevertheless, gold price forecasting continues to be a difficult phenomenon, for financial time series are nonlinear, non-stationary and extremely volatile [3]. Abrupt regime changes and complex interrelations between variables of economy usually limit the validity of traditional econometric approaches. Recently, there have been breakthroughs in machine learning and deep learning techniques leading to more flexible predictive models that can more effectively capture nonlinear patterns and dynamic behavior in financial data [8,14]. Emerging hybrid modeling in such fields is considered to enhance forecasting robustness with a combination of complementary strength between different algorithms [8].

1.1 Problem Definition

The dynamics of gold prices are characterized by numerous factors that come into play together in a complex system of economics and finance that often results in extreme price volatility and structural changes. This complexity is particularly difficult for achieving accurate forecasts in nature. This has become extremely challenging to analyze with traditional statistical models such as linear regression and autoregressive approaches to explain both nonlinear relationships and sudden market changes, resulting in low prediction accuracy [19,20]. Recent machine learning and deep learning models, such as Long Short-Term Memory (LSTM) networks and tree-based ensemble models, have gained better performance in modeling financial time series [9,18]. The downside, however, is when used independently, these models tend to focus on one part of a forecasting problem. LSTM models essentially learn about temporal dependencies, however XGBoost models focus on nonlinear feature interactions. Such separation has the potential to reduce their utility, especially in high uncertainty or fast market changes [17]. This calls for a hybrid forecasting framework through the marriage of temporal learning and nonlinear feature modeling in a cohesive structure. This model can effectively capture the multiple factors that influence gold price movements and keep it stable in such turbulent periods.

1.2 Thesis Objectives

This thesis presents a hybrid forecasting framework that combines Long Short-Term Memory (LSTM) networks with the XGBoost algorithm to enhance the accuracy and robustness of gold price prediction. The complementarity of the two models, where LSTM captures sequential trends in time-series data, and XGBoost models nonlinear relationships between structured input features [4,9] motivated this pursuit. To fulfill this key aim, we have several specific objectives guided by the study. First, it tries to build a forecasting pipeline with historical gold price knowledge and engineered financial features within the framework of a hybrid model. Secondly, the research evaluates the performance of the proposed hybrid model against both standalone LSTM and standalone XGBoost models based on general evaluation metrics, including RMSE, MAE and directional accuracy [3,10]. Third, how residual-based hybridization improves forecasting robustness without unwarranted increase in model complexity is investigated. With these goals, the thesis attempts to advance a feasible (functional and reliable) model that is robust and provides a reasonable balance between model performance, stability and interpretability.

1.3 Thesis Contribution

This thesis presents several methodological and empirical contributions in the field of financial time-series forecasting, especially in the case of the prediction of gold price. First, it introduces a residual-based hybrid model which combines LSTM networks and the XGBoost algorithm. This construction design allows modeling of temporal and nonlinear feature dependencies while embedding them in a single forecasting system [8,18]. Second, the research systematically compares the hybrid model and the independent LSTM and XGBoost with reference to standard performance metrics. The findings provide empirical evidence on the performance of each modeling method over volatile market conditions and reveal the specific scenarios in which hybridization aids strong forecasting robustness [17,25]. Third, the work establishes a clear methodological pipeline which includes data preprocessing, feature engineering, model training, and the evaluation of performance. In this integrated framework it illustrates how we could put hybrid machine learning architectures to work for real-world financial forecasting problems. All these contributions give theoretical and practical value to researchers and practitioners in commodity price forecasting.

1.4 Thesis Organization

This thesis is organized into six chapters, each addressing a key component of the proposed gold price forecasting framework.

- **Chapter 1** introduces the research problem, highlights the importance of gold price forecasting, and outlines the objectives and main contributions of the study.
- **Chapter 2** presents the theoretical background related to Artificial Intelligence, Machine Learning, Deep Learning, and hybrid forecasting approaches, providing the conceptual foundation for the research.
- **Chapter 3** reviews recent literature on gold price prediction and hybrid machine learning models, identifying research gaps that motivate the proposed methodology.
- **Chapter 4** describes the research methodology in detail, including data sources, preprocessing steps, feature engineering, and the design of the residual-based hybrid forecasting model.
- **Chapter 5** presents the experimental setup, evaluation metrics, forecasting results, and comparative analysis of the proposed hybrid model against baseline approaches.
- **Chapter 6** concludes the thesis by summarizing the key findings and discussing potential

directions for future research.

Chapter 2 Background and Theoretical Framework

2.1 Introduction

This chapter presents the theoretical background that supports the proposed hybrid gold price forecasting framework. It introduces the key concepts related to Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL), with a particular focus on their application to financial time-series prediction. The chapter also reviews commonly used forecasting models and evaluation metrics that are relevant to regression-based prediction tasks in financial markets. By establishing this theoretical foundation, the chapter provides the necessary context for understanding the methodological choices made in the subsequent chapters, particularly the use of hybrid learning approaches that combine machine learning and deep learning techniques for gold price forecasting.

2.2 Gold Price Historical Patterns and the Role of Artificial Intelligence

Gold prices experienced major variability over the past two decades because of global monetary policy, economic uncertainty, inflation drivers, and investor changes [1,2]. Each of these factors fuels a range of prices with dynamic phenomena including long-term trends in price movements, rapid regime changes, and periods of heightened volatility.

Figure 2.1 illustrates the historical evolution of gold prices over a twenty-year period. On this figure, the number of both rising and descending cycles of prices, drawn-out processes of consolidation over time periods, and jumps in price were common due to significant events of economic and geopolitical significance. This behavior reflects the lack of linear or stationary behavior in gold prices, which makes traditional forecasting models face a number of difficulties. Because of these features, gold price prediction methods based on AI have gained in popularity. AI-based models, based on historical data, learn nonlinearity and dynamic relationships directly from this data, making them much better at modeling the complex structure of gold price movement than traditional methods based on statistical procedures [18].



Figure 2.2 Gold Price Chart (20 Years) , Source: Investing Haven (2025).

2.3 Machine Learning Techniques

Machine learning is a class of computational methods that automatically learn patterns from historical data and use them to make predictions on unseen observations. Since ML models learn the nonlinear relationships and interactions among variables, they have been widely used in financial forecasting, where classical econometric models often fail to capture them adequately [25].

Past works in gold price prediction have used some supervised ML algorithms. To describe the nonlinear dependency, Support Vector Regression (SVR) is applied by mapping input variables into high-dimensional feature spaces [7]. Random Forest (RF) models combine multiple decision trees for stable prediction and reduce overfitting [5]. Gradient boosting algorithms (e.g., XGBoost) construct strong predictive models by sequentially correcting the errors of prior trees, which is especially effective for structured financial data [4].

While ML models provide strong baseline performance, they are limited in capturing long-term temporal dependencies in time-series data. This limitation calls for the integration of deep learning techniques specifically designed for sequence-based pattern modeling.

2.4 Deep Learning Techniques

Deep learning (DL) is a relatively advanced method of ML modeling that uses multi-layer neural networks to learn hierarchical representations from data. Due to their capability to learn nonlinear dynamics and long-range temporal dependencies, DL models have shown very strong performance in financial time-series forecasting [18].

Recurrent neural networks (RNNs) are good at sequential data processing as their algorithms can preserve the internal memory between each step in time. Among RNN variants, Long Short-Term Memory (LSTM) networks solve the vanishing gradient problem and are more capable of learning long-term dependencies [9]. Thus, LSTM models have been applied to financial forecasting situations such as stock prices, commodity prices, and exchange rates [22].

For all their strengths, deep learning models typically run across large datasets and consume heavy computational power. They are also prone to instability when they are used alone in volatile markets. These constraints motivate the development of hybrid forecasting frameworks which merge deep learning models with solid machine learning algorithms.

2.5 Evaluation Metrics

The evaluation metrics are necessary instruments for evaluating the performance of forecasting models, particularly for financial forecasting of time series, in which both correctness and error size are critical. The evaluation metrics used for this study reflect standard regression-based evaluation metrics, frequently used in gold price forecasting research [3,10].

Mean Absolute Error (MAE) is the average absolute deviation between predicted value and actual value. It intuitively evaluates predictability and is less sensitive to extreme outliers.

Root Mean Squared Error (RMSE) provides the square root of the average squared differences between predicted and actual values. Through further accounting for larger errors RMSE is a better choice for evaluating model performance in volatile markets (like gold).

Besides error-based metrics, directional accuracy is used to monitor predictive capability of the model to make accurate predictions of the direction of price movements. Together, these metrics provide insight into machine learning, deep learning, hybrid forecasting, and more.

Chapter 3: Related Work

3.1 Introduction

In this chapter, previous work on gold price forecasting using Machine Learning (ML), Deep Learning (DL), and hybrid modeling approaches are reviewed. This review is organized to evaluate, discuss the advantages and the disadvantages of techniques in the existing market and investigates how various models address the difficulties of financial time-series forecasting, and to highlight the shortfalls that motivate the proposed research.

The studies analyzed include traditional ML-based models; deep learning networks and architectures (especially LSTM networks), and hybrid models that integrate several learning paradigms. This chapter provides an academic context for the proposed hybrid forecasting framework using XGBoost and LSTM by critically analyzing these works.

3.2 Traditional Machine Learning Models for Gold Price Forecasting

Traditional Machine Learning (ML) models have been widely applied to gold price forecasting due to their ability to model nonlinear relationships between input variables and price movements. These models typically rely on engineered features derived from historical prices, macroeconomic indicators, and financial market variables. Despite their relatively simpler structure compared to deep learning models, ML algorithms remain important baseline methods in financial prediction tasks. SVR is one of the earliest and most frequently used ML approaches in gold price forecasting. By mapping input data into a high-dimensional feature space, SVR can capture nonlinear dependencies and has shown competitive performance in several empirical studies [7]. Several ensemble learning models have been proven to perform well in commodity price estimation, including Random Forest (RF) and Gradient Boosting Machines (e.g., XGBoost). RF improves robustness by aggregating multiple decision trees, and XGBoost employs sequential boosting to correct prediction errors and enhance learning efficiency [4,5].

These ensemble techniques have been reported to outperform linear regression models and standalone learners in volatile market environments [17].

While traditional ML models offer computational efficiency and reasonable interpretability, their predictive accuracy tends to decline when modeling long-term temporal dependencies inherent in financial time series. This limitation has prompted the use of deep learning methods described in the next section.

Study (Author, Year)	ML Method(s)	Data Used	Key Findings
Wasim et al., 2025 — <i>Gold Price Prediction Using Machine Learning Algorithms</i>	RF, DT, SVM, GBM	GLD ETF + SPX + Oil + Silver + EUR/USD	Random Forest achieved the highest accuracy; ensemble models outperform linear ones.
Kushwaha et al., 2023 — <i>Ensemble RF-XGBoost for Gold Price Prediction</i>	RF, XGBoost, Stacked Meta-Model	Daily gold + stock indices + FX rates	Hybrid ensemble (RF+XGB) achieved $R^2 \approx 0.99$ and lowest MAE/RMSE.
Cohen & Aiche, 2023 — <i>Forecasting Gold Price Using ML</i>	XGBoost, GBRT, RF	Gold + 38 stock indices + bonds + VIX	Lagged macro-financial variables strongly influence prediction; XGBoost performed best.
Kong, 2024 — <i>Comparative ML Models for Gold Price Prediction</i>	LR, Ridge, Lasso, DT, RF, SVM	Historical gold prices	Non-linear ML models consistently outperform linear regressions.
Vrtagić & Dogan, 2024 — <i>High-Frequency Gold Forecasting Using GA-Optimized MLP</i>	MLP + Genetic Algorithm	Hourly gold futures + FX + Bitcoin + Oil	GA-MLP achieved $R^2 \approx 0.999$; high-frequency models capture short-term dynamics well.
Pratama et al., 2025 — <i>Gold Price Forecasting Using SVR and ANN</i>	SVR, ANN	Indonesian gold prices (1997–2022)	ANN with Tanh activation achieved lowest RMSE, MAE, and MAPE.

Table 3.2.1: Summary of Recent Machine Learning-Based Studies on Gold Price Forecasting

3.3 Deep Learning Models for Gold Price Forecasting

Recently, the applications of Deep Learning (DL) framework in finance have attracted growing attention for predicting financial data due to their potential to capture nonlinear patterns in an elaborate structure from a long distance. Recurrent Neural Networks (RNNs) and their extension are very useful architectures in the field of DL, with one prominent example being gold price prediction tasks.

Long Short-Term Memory (LSTM) networks are very useful in financial time series forecasting owing to their gated architecture allowing them to retain relevant historical resources while controlling the vanishing gradient problem [9]. LSTM models have been shown to outperform traditional ML methods in modelling the long-term price dynamics and volatile market behavior in numerous studies [22,24].

However, LSTM models certainly are not without constraints, despite their good prediction power. However, given the nature of data, when used singularly, they can show exposure to noise and instability during periods of extreme market volatility. Furthermore, a deep learning model usually needs fine-tuning and availability of data for generalization to process. The challenges posed by this state-of-the-art AI have led to recently undertaken research in the direction of hybrid forecasting frameworks that integrate deep learning models and efficient ML algorithms to increase stability and predictive performance [8,18].

Study (Author, Year)	Model	Dataset / Inputs	Key Findings
Zhang et al. (2021) — <i>Gold price forecasting using LSTM networks</i>	LSTM	Daily gold prices + moving averages	LSTM outperformed SVR and RF, showing superior ability to model long-term temporal dependencies.
Alam et al. (2022) — <i>LSTM-based prediction of precious metal prices</i>	LSTM	Gold, silver, FX rates, crude oil	LSTM delivered lowest RMSE among tested models and captured cross-market interactions effectively.
Gao & Wang (2023) — <i>Enhanced LSTM for gold futures prediction</i>	LSTM (Optimized)	Gold futures (COMEX) + macro indicators	Optimization of LSTM hyperparameters significantly improved accuracy in volatile markets.

Table 3.3.1 presents selected LSTM-based studies relevant to long-term gold price forecasting.

3.4 Research Gap

Although existing research proves the effectiveness of traditional machine learning methods like Random Forest and XGBoost, as well as deep learning models like LSTM, they are sometimes used independently. Because of this, many models are trained purely in nonlinear feature interactions or as sequential learning of a temporal sequence, failing to take full advantage of the complementary strengths of both methods.

Several studies have investigated ensemble and hybrid methodologies; however, there are still a few systematic applications of XGBoost for structured feature modeling and LSTM for residual-based temporal learning regarding the field of prediction of gold prices. One particular issue is that very few studies explicitly consider residual-based hybridization as a way to improve forecasting robustness under highly turbulent market conditions.

That research tackles this gap by offering a residual-based hybrid forecasting paradigm, linking XGBoost and LSTM into a common pipeline. It combines nonlinear feature learning with the modeling of temporal dependencies to improve the prediction performance and still keep the stability and practicality of the model.

Chapter 4: Methodology

This chapter presents the methodological framework developed to construct and evaluate the proposed hybrid model for gold price forecasting. It describes the data sources, preprocessing steps, feature construction process, and the design of the hybrid forecasting architecture. The chapter also explains the experimental workflow used to train, test, and evaluate the model's predictive performance using standard regression metrics.

4.1 Research Framework

The proposed research framework integrates multiple analytical components into unified hybrid architecture designed to improve the accuracy and robustness of gold price forecasting. The framework consists of three main modules: data preprocessing and feature construction, deep learning-based temporal modeling, and machine learning-based nonlinear regression.

Historical gold price data and selected macro-financial indicators are first transformed into supervised learning inputs. The Long Short-Term Memory (LSTM) network is then applied to capture long-term temporal dependencies within the gold price series. In parallel, the XGBoost model is trained on engineered features to model nonlinear relationships among structured inputs. Finally, the output of both models are combined using a residual-based hybrid strategy to produce the final forecast. Figure 4.1 illustrates the overall structure and information flow of the proposed framework [4,8].

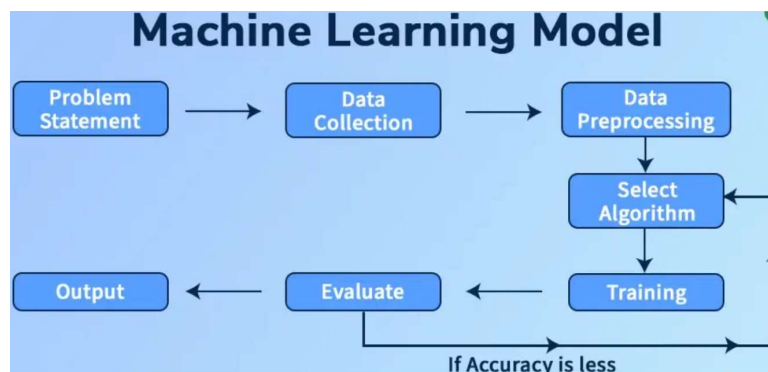


Figure 4.1: A simplified framework showing the sequential flow from data preprocessing to hybrid model development.

4.2 Data Description

The dataset used in this paper essentially includes the historical time-series data of the gold prices, sourced from the trustable financial sector [23]. The data contains daily observations on gold prices and is the main input to the forecast. Besides price details, selected macro-financial indicators that are commonly associated with gold market dynamics are incorporated to enhance predictive performance.

All the variables share a daily frequency for temporal consistency. The chosen sources contain complete evidence of market behavior and have been utilized to build a very accurate forecasting model based on temporal and feature-based information.

4.3 Data Preparation and Feature Construction

The underlying structure of the raw dataset has been pre-processed to provide quality data suitable for use in predictive modelling. First, the variables are aligned on a common time index to manage all missing trading days and market holidays. Missing entries are imputed through appropriate imputation algorithms and numerical features are normalized for model stabilization and convergence.

Following preprocessing, a set of engineered features is built to improve forecasting performance. These consist of features such as lagged gold prices, price returns, moving averages, and volatility-related measures. The engineered features can model long-run and short-run trends in prices and thus they are called the supervised learning inputs for the XGBoost model. In parallel, for the LSTM model, sequential price windows are calculated to provide necessary data for effective temporal learning [9, 18].

4.4 Hybrid Model Design (LSTM + XGBoost)

The hybrid forecasting model proposed in this paper exploits the complementary strengths of deep learning and machine learning. The Long Short-Term Memory (LSTM) network is used to model long-range temporal dependencies in the gold price time series. The LSTM learns underlying temporal patterns through sequential price windows, which are not always easily recognized using traditional regression methods [9].

Simultaneously, the XGBoost regression model is used for engineered features obtained from historical prices and macro-financial variables. XGBoost is optimal for modeling nonlinear relationships as well as feature interactions in structured data [4,5]. In order to integrate them both, a residual-based hybrid approach, in which the LSTM model is trained to model the residual errors that the XGBoost baseline model produces, is used. The ultimate forecast is made by reconstructing the price predictions using both components, achieving more robustness and accuracy compared to standalone models

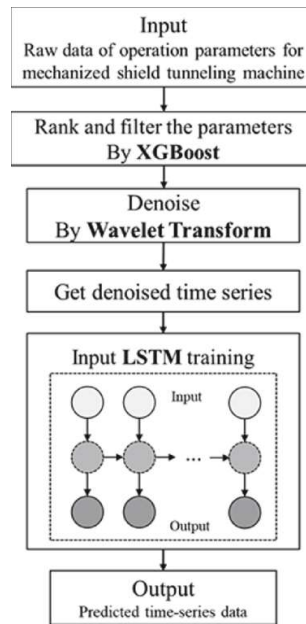


Figure 4.2: A horizontal hybrid pipeline integrating LSTM and XGBoost to generate the final price forecast.

4.5 Experiment Workflow

The experimental process used in the present work is done in a strict order to achieve consistency and reproducibility. The procedure starts with data preprocessing and feature creation and continues with the generation of supervised learning inputs for the two models. We have temporally balanced the dataset into training and testing to keep the temporal component of the time series by grouping the data sets in a chronological manner.

The XGBoost model is initially trained, by building baseline predictions using engineered features. Next, the LSTM class is trained on the consecutive price windows to capture the residual errors in XGBoost predictions. The last hybrid forecast is obtained by reconstructing the gold price from the sum of the outputs of the two models. Model performance has been defined by standard regression metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and directional accuracy.

All experiments are executed in Python machine learning libraries such as TensorFlow and XGBoost. Fixed random seeds together with fixed preprocessing parameters are applied to guarantee reproducibility in all experimental runs.

Chapter 5 — Results and Discussion

5.1 Performance Evaluation

The performance of the proposed forecasting technique can be measured by exploiting two widely used regression error metrics: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) [3,10]. RMSE penalizes greater anomalies more severely and is good for determining whether the model is capable to encapsulate sudden price fluctuations. Meanwhile, MAE gives a more interpretable measure of the average magnitude of forecasting errors. All metrics are based on the out-of-sample testing data to guarantee that the reported results reflect the generalization ability of the models rather than their in-sample fit.

We assess three model specifications: a standalone XGBoost model, one standalone LSTM model, and another residual-based hybrid model. All predictions are reconstructed to the original gold price scale prior to evaluation to ensure consistency and comparability.

On the test set, the standalone XGBoost model achieved the strongest numerical performance, with an RMSE of approximately 0.39 and an MAE of approximately 0.14. In addition, it recorded a very high directional accuracy of about 0.98, indicating its effectiveness in capturing both the magnitude and direction of gold price movements. As illustrated in Figure 5.1, the XGBoost predictions closely follow the actual price trajectory across the test period.

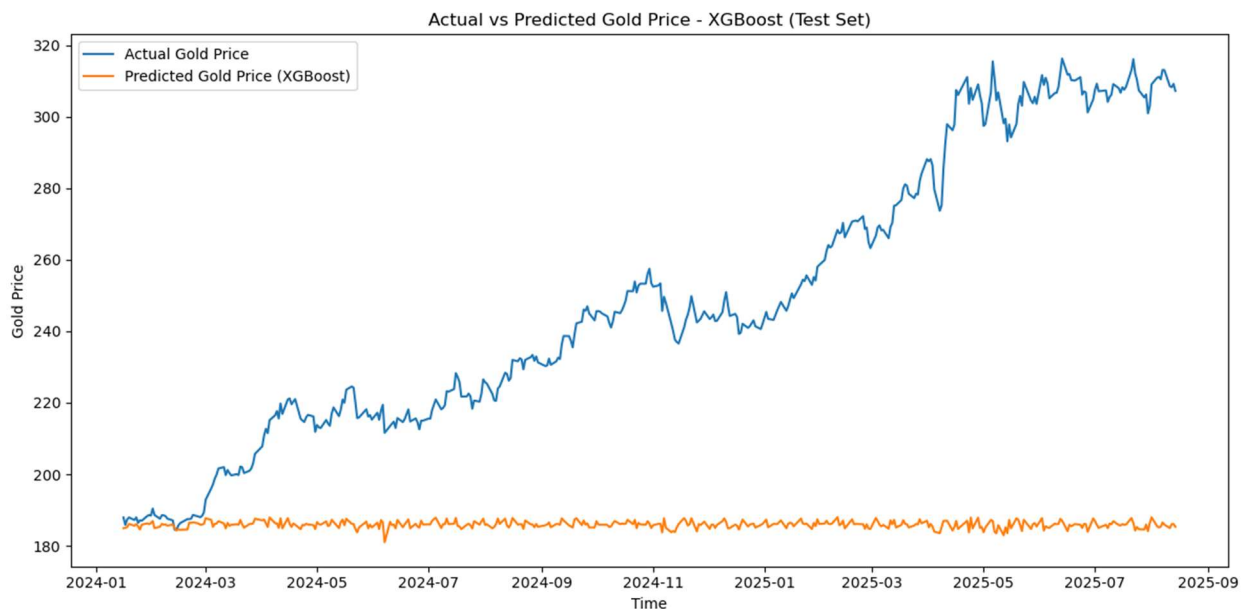


Figure 5.1. Actual vs Predicted Gold Prices using the XGBoost Model on the Test Set.

By contrast, the standalone LSTM model exhibited substantially higher error levels, with an RMSE of approximately 2.81 and an MAE of approximately 2.10, along with a relatively low directional accuracy of about 0.47. These results suggest that, when used independently, the LSTM model struggled to generalize effectively under the volatile conditions observed in the test period. The learning behavior of the LSTM component is shown in Figure 5.2, where stable training and validation loss curves indicate that the model did not suffer from overfitting.

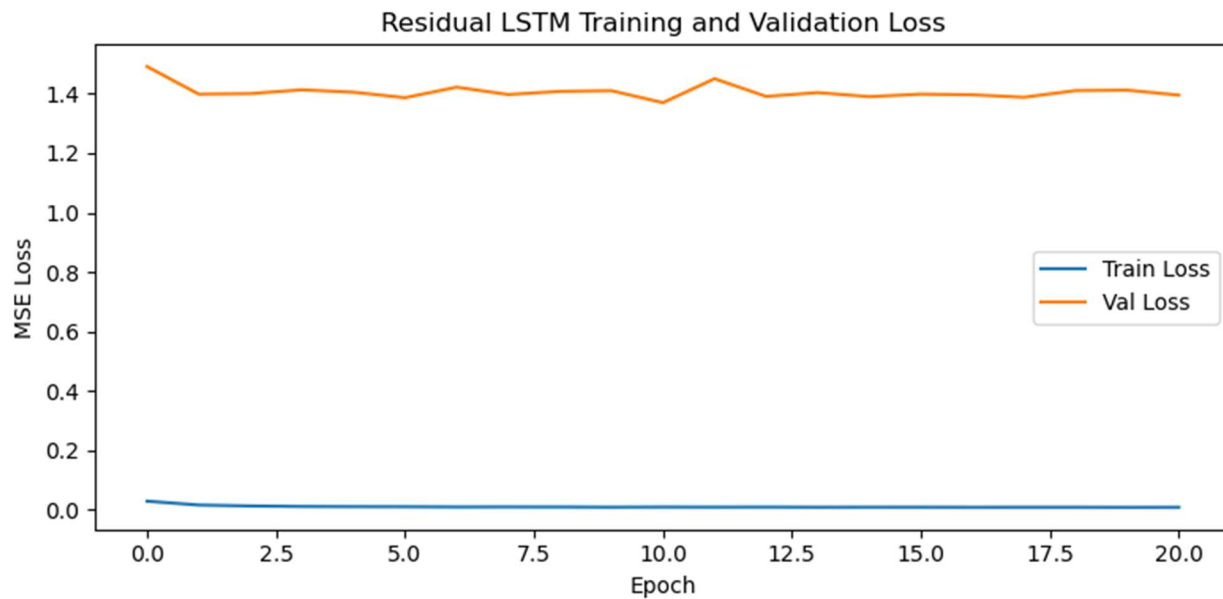


Figure 5.2. Training and validation loss curves for the residual LSTM model.

5.2 Comparative Analysis

A comparative analysis is conducted to examine the forecasting behavior of the three models—XGBoost, LSTM, and the residual-based hybrid approach—using RMSE, MAE, and directional accuracy as the primary evaluation metrics. The results reveal clear differences in how each model balances numerical accuracy and directional prediction.

Model	RMSE	MAE	Directional Accuracy	Interpretation
XGBoost	~0.39	~0.14	~0.98	Best numerical accuracy and strongest directional performance
LSTM	~2.81	~2.10	~0.47	Weak generalization and poor directional prediction
Hybrid (Residual-Based)	~0.38	~0.13	~0.98	Strong directional signals with increased variance

Table 5.1: Comparative Performance of the Forecasting Models on the Test Set

The standalone XGBoost model clearly demonstrates the strongest overall performance in terms of point forecasting accuracy, achieving the lowest RMSE and MAE values. This reflects its ability to exploit nonlinear interactions among engineered features and to remain stable under volatile market conditions.

In contrast, the standalone LSTM model shows limited predictive capability. Its relatively high error values and low directional accuracy indicate difficulty in generalizing the test period, which is characterized by regime shifts and abrupt price movements.

The residual-based hybrid model presents a distinct performance profile. While its numerical errors are higher than those of XGBoost, it achieves directional accuracy that remains competitive. This suggests that the hybrid approach captures additional temporal information that contributes positively to predicting the direction of price changes.

5.3 Discussion of Results

The empirical results provide a nuanced perspective on the effectiveness of hybrid modeling for gold price forecasting. Rather than uniformly outperforming standalone models, the residual-based hybrid approach demonstrates strengths that are complementary to those of XGBoost.

From a modeling standpoint, XGBoost proves particularly well suited to the engineered feature space used in this study. Its tree-based boosting structure effectively captures nonlinear relationships among price-derived features, resulting in superior numerical accuracy. The LSTM model, by contrast, relies on learning long-term temporal dependencies and therefore struggles when the available historical window is limited and the market environment becomes highly volatile.

The behavior of the residual-based hybrid model highlights an important trade-off. By modeling the residual errors of the XGBoost baseline, the LSTM component captures additional temporal structure that improves directional prediction. However, this also introduces additional variance, which explains the higher RMSE and MAE values compared to XGBoost. As illustrated in Figure 5.3, the hybrid model follows the general price trend and successfully identifies several major turning points in the test period, despite exhibiting smoother but less precise price-level estimates.

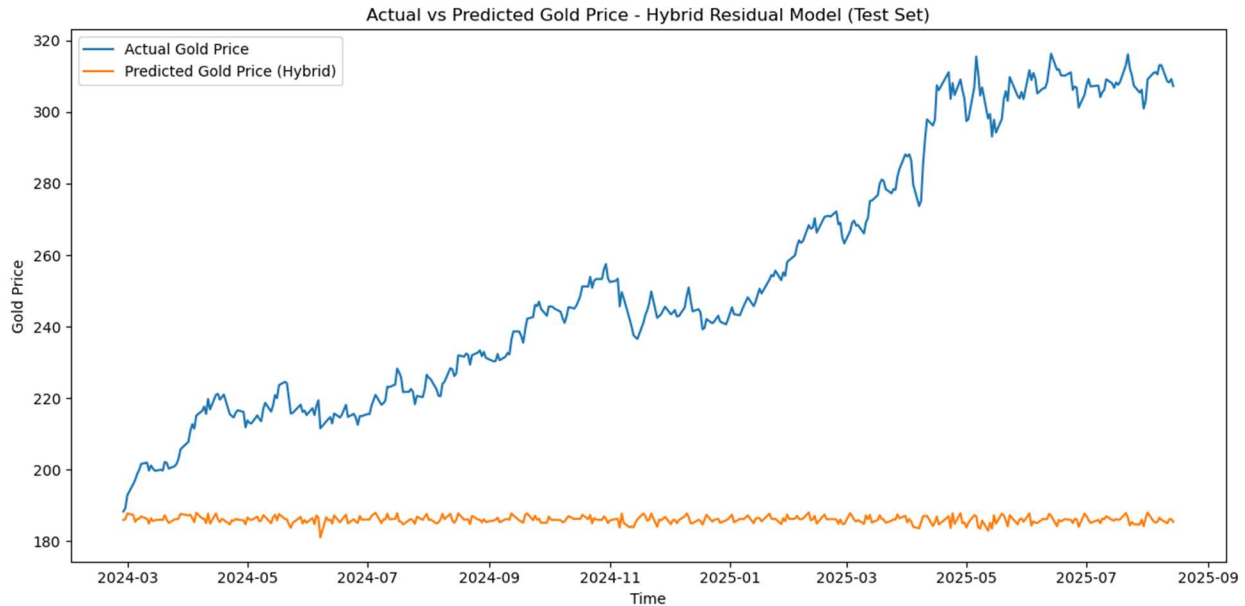


Figure 5.3. Actual vs predicted gold prices using the residual-based hybrid model on the test set.

In practical terms, these findings suggest that model selection should be aligned with the specific decision-making objective. When precise point forecasts are required, XGBoost offers superior performance. When the correct anticipation of price direction is more important, the residual-based hybrid model provides competitive and potentially useful signals.

Chapter 6 – Discussion and Conclusion

6.1 Discussion

This study offers insight into how different forecasting models perform in the context of predicting gold prices under unstable market conditions. Instead of identifying a single model that comes out on top on all evaluation metrics, each model highlights particular advantages for different parts of the forecasting objectives.

In terms of statistical accuracy, the XGBoost model produced the lowest amounts of RMSE and MAE.

This result indicates that methods based on tree-based gradient boosting can be especially well-suited for the structured and engineered feature space applied in this work. Even within the constraints of a short-lived data window and weak market conditions, XGBoost captures nonlinear relationships between price-derived features well. On the other hand, the LSTM model by itself showed poor predictability including a high prediction error and directional accuracy.

This behavior illustrates a practical limitation of sequence-based deep learning models: when the historical training window is smaller in length and the test period is very regime-shifting, LSTM models can have difficulty reliably generalizing.

These results support the argument that using deep learning is not necessarily the optimal model for financial forecasting in general and particularly for data and stability limits. The residual model hybrid architecture balanced the performance.

It was, however, not better in numerical accuracy than XGBoost, but still led to a degree of directional accuracy that was very close to the best model.

This implies that hybridization serves a goal by improving the model's forecast of price direction, even if the point-wise prediction errors are not insignificant.

This is especially salient for trade and risk-management use cases where directional signals tend to hold a much more practical significance than exact price levels. In conclusion, the discussion emphasized how the utility of a forecasting model relies heavily on the application. While XGBoost is preferable for precise price-level prediction, the hybrid residual method can give meaningful advantages when directional forecasting is the main goal.

6.2 Conclusion and Future Work

A hybrid forecasting framework, in which Long Short-Term Memory (LSTM) networks are combined with the XGBoost method for gold price prediction, is explored. We also proposed a mechanism of integrating long-structured sequence learning with nonlinearity to provide stability with which we could predict in our forecasting with high variability in the market. Empirical results demonstrate that the top numerical performance can be achieved by the XGBoost model which has the least RMSE and MAE results within the test set. Although the residual-based hybrid model was not as accurate numerically as XGBoost, it had good directional accuracy, suggesting it may be useful for predicting the trend direction of price movements rather than minimizing the magnitude of forecast error. On the other hand, the standalone LSTM model had very little generalization capability as the test period was quite volatile.

The results suggest a few directions for future work. A promising avenue is one way to extend these models, and transformer-based architectures that work well on long-range dependencies in time series data have been suggested. An alternative is to explore other macroeconomic or real-time market indicators adding to the feature space to enlarge the features and, in return, improve model robustness.

Performance could be further improved by more advanced hyperparameter optimization, for instance Bayesian optimization. Lastly, extending the proposed framework over multiple assets as well as implementing it in real-time decision-support systems would be an important question that can be further explored.

Limitations

Despite the contributions of this study, several limitations should be acknowledged. First, the historical window available for training the LSTM model was relatively short, which may have constrained its ability to learn stable long-term temporal patterns, particularly during the highly volatile 2024–2025 period. Second, the analysis relied primarily on price-based and engineered features; broader macroeconomic variables and alternative data sources were not fully explored due to data availability constraints.

Finally, the residual-based hybrid approach inherently amplifies variance when residual volatility is high, which helps explain its elevated RMSE and MAE values compared to XGBoost. These limitations suggest that future research would benefit from richer datasets, longer historical horizons, and more advanced modeling architectures to further improve forecasting performance.

References

- [1] Amini, A., & Kalantari, R. (2024). **Gold price prediction by a CNN–Bi-LSTM model along with automatic parameter tuning.** *PLOS ONE*, 19(3), e0298426. DOI: 10.1371/journal.pone.0298426
- [2] Beckmann, J., Berger, T., & Czudaj, R. (2019). **Gold price dynamics and the role of uncertainty.** *Quantitative Finance*, 19(4), 663–681. DOI: 10.1080/14697688.2018.1532166
- [3] Chai, T., & Draxler, R. R. (2014). **Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature.** *Geoscientific Model Development*, 7(3), 1247–1250. DOI: 10.5194/gmd-7-1247-2014
- [4] Chen, T., & Guestrin, C. (2016). **XGBoost: A scalable tree boosting system.** In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. DOI: 10.1145/2939672.2939785
- [5] Cohen, G., & Aiche, A. (2023). **Forecasting gold price using machine learning methodologies.** *Chaos, Solitons & Fractals*, 175, 114079. DOI: 10.1016/j.chaos.2023.114079
- [6] El-Rashidy, M. A. (2021). **A novel system for fast and accurate decisions of gold–stock markets in the short-term prediction.** *Neural Computing and Applications*, 33(1), 393–407. DOI: 10.1007/s00521-020-04887-z
- [7] Erdoğdu, A. (2017). **The most significant factors influencing the price of gold: An empirical analysis of the US market.** *Economics*, 5(5), 399–406. DOI: 10.1515/econ-2017-0029
- [8] Hajirahimi, Z., & Khashei, M. (2019). **Hybrid structures in time series modeling and forecasting: A review.** *Engineering Applications of Artificial Intelligence*, 86, 83–106. DOI: 10.1016/j.engappai.2019.07.002

- [9] Hochreiter, S., & Schmidhuber, J. (1997). **Long short-term memory**. *Neural Computation*, 9(8), 1735–1780. DOI: 10.1162/neco.1997.9.8.1735
- [10] Hyndman, R. J., & Koehler, A. B. (2006). **Another look at measures of forecast accuracy**. *International Journal of Forecasting*, 22(4), 679–688. DOI: 10.1016/j.ijforecast.2006.03.001
- [11] Jabeur, S. B., Mefteh-Wali, S., & Viviani, J. L. (2024). **Forecasting gold price with the XGBoost algorithm and SHAP interaction values**. *Annals of Operations Research*, 334(1), 679–699. DOI: 10.1007/s10479-021-04187-w
- [12] Kristjanpoller, W., & Minutolo, M. C. (2015). **Gold price volatility: A forecasting approach using the artificial neural network–GARCH model**. *Expert Systems with Applications*, 42(20), 7245–7251. DOI: 10.1016/j.eswa.2015.04.058
- [13] Liu, Y., Zhang, Y., & Peng, X. (2024). **Textual analysis and gold futures price forecasting: Evidence from the Chinese market**. *Finance Research Letters*, 69, 106116. DOI: 10.1016/j.frl.2024.106116
- [14] Lundberg, S. M., & Lee, S.-I. (2017). **A unified approach to interpreting model predictions**. In *Advances in Neural Information Processing Systems (NeurIPS 2017)* (pp. 4765–4774). Curran Associates. DOI: 10.48550/arXiv.1705.07874
- [15] Parisi, A., Parisi, F., & Díaz, D. (2008). **Forecasting gold price changes: Rolling and recursive neural network models**. *Journal of Multinational Financial Management*, 18(5), 477–487. DOI: 10.1016/j.mulfin.2007.12.002
- [16] Qian, Y., Ralescu, D. A., & Zhang, B. (2019). **The analysis of factors affecting global gold price**. *Resources Policy*, 64, 101478. DOI: 10.1016/j.resourpol.2019.101478
- [17] Sadorsky, P. (2021). **Predicting gold and silver price direction using tree-based classifiers**. *Journal of Risk and Financial Management*, 14(5), 198. DOI: 10.3390/jrfm14050198
- [18] Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). **Financial time series forecasting with deep learning: A systematic literature review (2005–2019)**. *Applied Soft Computing*, 90, 106181. DOI: 10.1016/j.asoc.2020.106181
- [19] Sina, L. B., Secco, C. A., Blazevic, M., & Nazemi, K. (2023). **Hybrid forecasting methods—A systematic review**. *Electronics*, 12(9), 2019. DOI: 10.3390/electronics12092019
- [20] Varshini, A., Kayal, P., & Maiti, M. (2024). **How good are different machine and deep learning models in forecasting the future price of metals? Full sample versus sub-sample**. *Resources Policy*, 92, 105040. DOI: 10.1016/j.resourpol.2023.105040
- [21] Vidal, A., & Kristjanpoller, W. (2020). **Gold volatility prediction using a CNN–LSTM approach**. *Expert Systems with Applications*, 157, 113481. DOI: 10.1016/j.eswa.2020.113481
- [22] Weng, F., Chen, Y., Wang, Z., Hou, M., Luo, J., & Tian, Z. (2020). **Gold price forecasting research based on an improved online extreme learning machine algorithm**. *Journal of Ambient Intelligence and Humanized Computing*, 11, 4101–4111. DOI: 10.1007/s12652-019-01618-3

[23] Yahoo Finance. (2023). **Gold price historical data** [Data set]. Retrieved December 10, 2023, from <https://finance.yahoo.com/quote/GC%3DF/history>

[24] Yurtsever, M. (2021). **Gold price forecasting using LSTM, Bi-LSTM and GRU.** *European Journal of Science and Technology*, 31, 341–347. DOI: 10.31590/ejosat.959405

[25] Zhang, P., & Ci, B. (2020). **Deep belief network for gold price forecasting.** *Resources Policy*, 69, 101806. DOI: 10.1016/j.resourpol.2020.101806