IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 4, August 2026, pp. 3228~3239
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i4.pp3228-3239

Hybrid deep learning model for enhanced short-term gold price forecasting

Hoang Ha Nguyen[1], Minh Duc Nguyen[2], Cuong H. Nguyen-Dinh[3]
[1]Faculty of Information Technology, University of Sciences, Hue University, Hue, Vietnam
[2]Department of Economic Information Systems, University of Economics, Hue University, Hue, Vietnam
[3]Department of Data Science, University of Finance - Marketing, Ho Chi Minh, Vietnam

Article Info

Article history:

Received Jul 2, 2025
Revised Jul 13, 2026
Accepted Jul 21, 2026

Keywords:

1D convolutional neural network
Gold price prediction
Hybrid deep learning
Long short-term memory
Time series forecasting

ABSTRACT

Accurate short-term gold price forecasting is crucial for the financial decision-making. This paper introduces a short-term gold prediction network (STGP-Net), a novel hybrid deep learning model designed to enhance prediction accuracy by integrating one-dimensional convolutional neural network (1D-CNN) and long short-term memory (LSTM). STGP-Net leverages the 1D-CNN's ability to extract local temporal features and the LSTM capacity to model long-range dependencies within gold price time series. Various sliding window configurations are employed to generate input sequences for multi-step ahead prediction. Comprehensive experiments were conducted comparing STGP-Net against 1D-CNN + recurrent neural network (RNN) and 1D-CNN + bidirectional long short-term memory (BiLSTM) baseline models across three configurations using metrics like mean absolute error (MAE), root mean square error (RMSE), and determination (R^2). The results demonstrated that STGP-Net consistently provided better performance and robustness, proving more effective for short-term gold price forecasting than the alternative hybrid models tested.

Corresponding Author:

Cuong H. Nguyen-Dinh
Department of Data Science, University of Finance – Marketing
778 Nguyen Kiem St., Duc Nhuan Ward, Ho Chi Minh City, Vietnam
Email: cuongndh@ufm.edu.vn

1. INTRODUCTION

Recent advancements in deep learning have highlighted the significant advantages of hybrid models that combine convolutional neural network (CNN) architectures for spatial feature extraction with recurrent neural network (RNN) architectures for temporal sequence modeling; these synergistic approaches have enhanced performance across diverse research domains ranging from sentiment analysis [1] and object detection in video [2] to audio processing [3] and medical image analysis [4]. This trend has extended into the financial sector, where accurately predicting time series data is crucial [5]. Numerous studies have explored hybrid models combining elements from the CNN and RNN families for forecasting volatile financial markets, including cryptocurrency price prediction [6], stock market analysis [7], and commodity forecasting, such as gold price prediction [8]. Within the CNN family, one-dimensional convolutional neural networks (1D-CNNs) possess a distinct advantage for time series analysis, as their kernels operate along the temporal dimension, making them highly effective at extracting local patterns and relevant features directly from sequential data [9]. Recognizing the potential synergy between 1D-CNNs' feature extraction power and the sequence-handling capabilities of different RNN variants, this study explores and evaluates hybrid models combining 1D-CNN with members of the RNN family—specifically standard RNN, long short-term memory (LSTM), and bidirectional long short-term memory (BiLSTM)—to identify a robust and effective candidate model for the challenging task of short-term gold price forecasting.

Gold holds significant importance in global financial landscape, widely recognized as a safe-haven asset and a crucial tool for hedging against inflation, currency fluctuations, and geopolitical uncertainty. Consequently, the accurate prediction of its price movements, particularly in the short term, is of considerable value to investors, traders, portfolio managers, and policymakers for making informed decisions, managing risk, and capitalizing on market opportunities. Addressing this need, this paper introduces the short-term gold prediction network (STGP-Net), a novel hybrid deep learning model specifically designed for enhanced short-term gold price forecasting. The major contributions of this study are threefold: i) propose the novel STGP-Net architecture, integrating a 1D-CNN with LSTM to effectively capture both local temporal features and long-range dependencies in gold price data; ii) systematically apply different sliding window scales (configurations) to generate input data, enabling the model to learn from varying historical sequence lengths for multi-step ahead prediction; and iii) conduct comprehensive experiments comparing STGP-Net against two relevant baseline hybrid models (1D-CNN + RNN and 1D-CNN + BiLSTM), demonstrating the superior performance and robustness of proposed model through rigorous evaluation metrics.

The rest of this paper is organized as follows. Section 2 elaborates on the state-of-the-art studies relevant to hybrid deep learning models for financial time series forecasting, providing context for this work. Section 3 introduces the proposed STGP-Net architecture in detail, elaborating on its constituent 1D-CNN and LSTM components and explaining the mechanism by which it processes data for prediction. Section 4 presents the experimental setup, including data preparation, configurations, baseline models, and evaluation metrics, followed by a presentation and discussion of the comparative results. Finally, section 5 concludes the paper by summarizing the key findings and suggesting potential directions for future research.

2. RELATED WORK

This section provides a review of the relevant literature concerning methodologies employed for forecasting gold prices, a task that has attracted significant research interest due to gold's economic importance. The evolution of predictive models applied to this challenge is examined, covering traditional statistical time series techniques, conventional machine learning algorithms, and more recent advancements in deep learning, including the development of hybrid architectures. This overview aims to situate the present study within the landscape of existing research and highlight the context for exploring novel hybrid deep learning approaches for enhanced short-term gold price prediction.

Historically, gold price forecasting was primarily addressed using traditional statistical time series models. For example, Makala and Li [10] compared the performance of autoregressive integrated moving average (ARIMA) and support vector machine (SVM) models for predicting daily gold prices using data spanning from 1979 to 2019. Their results indicated that the SVM model significantly outperformed the ARIMA model. Based on substantially lower root mean square error (RMSE) and mean absolute percentage error (MAPE) values, the authors concluded that SVM is a more accurate and suitable method for gold price prediction. Such findings suggested that even conventional machine learning models like SVM could offer improved predictive capabilities over traditional statistical time series methods like ARIMA for handling the complexities inherent in gold price data.

As a machine learning approach to address the limitations of gradient descent in neural network optimization, Doush *et al.* [11] proposed an archive-based Harris hawks optimizer (AHHO) to tune multi-layer perceptron (MLP) parameters for gold price forecasting. The researchers employed Pearson's correlation and a categorized correlation for feature selection and benchmarked the model against eight alternative machine learning and swarm intelligence algorithms. Their findings demonstrated that the AHHO-neural network model produced higher accuracy and lower error metrics than the comparative models, suggesting that the integration of an archive effectively improves performance in complex forecasting tasks.

To further the development of machine learning trends in forecasting financial price data, recent literature has shifted toward hybrid frameworks designed to capture the volatility and multi-scale dependencies inherent in these series. While Ghosh and Jana [12] focused on the extreme volatility of high-frequency intraday movements by utilizing maximal overlap discrete wavelet transformation and Bayesian structural models to maintain predictive resilience during global economic shocks, Song and Chen [13] addressed the broader temporal features of the shipping market by employing variational mode decomposition (VMD) with light gradient boosting machine (LightGBM) to extract multi-scale information from lower-frequency time series. Distinguishing itself from these decomposition-centric methods, the work of Bhambu *et al.* [14] introduced the recurrent ensemble deep random vector functional link (RedRVFL) network, which leverages stable, fixed-weight recurrent layers and deep representation learning to achieve superior accuracy across diverse financial datasets without requiring the specific signal-processing stages used in the aforementioned studies.

Cohen and Aiche [15] explored advanced machine learning techniques, specifically comparing random forest (RF), gradient boosted regression trees (GBRT), and extreme gradient boosting (XGBoost) for forecasting gold prices. The models utilized a comprehensive set of input features including lagged global stock indices, volatility index (VIX), commodity futures, and bond yields from several major economies. The research identified that certain lagged stock indices (ASX and S&P500), US/Japan bond yields, gas/silver prices, and the VIX were particularly influential predictors, confirming the impact of economic uncertainty. Ultimately, the study concluded that ensemble models like GBRT and XGBoost proved valuable for making informed gold investment decisions based on these diverse financial indicators. A comprehensive overview summarizing the application of various machine learning techniques to gold price prediction over the preceding decade was provided [16].

More recently, the field has witnessed a significant shift towards deep learning methodologies, leveraging complex neural network architectures to potentially capture more intricate temporal patterns in gold price data. For example, Chen and Huang [17] compared CNN and LSTM models, optimized using Bayesian optimization, to forecast S&P 500 movements based on various inputs including financial indicators as well as gold and crude oil price and volatility indices. The models were trained on data from 2010-2018 and tested on 2019 data. Experimental results, evaluated based on classification accuracy and back-tested return on investment (ROI), showed that the proposed CNN model with eight features (CNN8) achieved a higher ROI (13.23%) than the equivalent LSTM model (11.06%). The authors also noted that including gold and oil-related features proved particularly beneficial for predictions related to certain industries like semiconductors and petroleum.

Beyond comparing established models like CNN and LSTM, research also explored novel standalone deep learning architectures designed to capture complex market dynamics directly. For instance, Huang *et al.* [18] proposed a novel multilevel graph attention network (ML-GAT) to address limitations in existing graph-based stock prediction methods, particularly regarding the incorporation of diverse information and stock relationships derived from knowledge graphs like Wikidata. Applying ML-GAT to predict trends for hundreds of stocks in the S&P 500 and CSI 300 indices, the researchers found it significantly outperformed five popular baseline approaches. The model achieved substantial improvements in F1-score, accuracy, average daily return, and sharpe ratio, demonstrating the effectiveness of the sophisticated single-model architecture.

Similarly, Wang *et al.* [19] proposed a novel forecast model named special gated recurrent unit (SGRU)-attention mechanism (AM), based on a SGRU combined with an AM. This model was applied to forecast China's gold futures prices using Shanghai Futures Exchange data from 2008 to 2021. Comparative experimental results demonstrated that the proposed SGRU-AM model achieved superior performance in both forecast accuracy and efficiency compared to the baseline methods evaluated.

Recognizing the distinct advantages of different architectures, a significant recent trend focuses on developing hybrid deep learning models, particularly combining the feature extraction power of CNNs with the sequential modeling strengths of RNN variants for gold price forecasting. For instance, Livieris *et al.* [8] explicitly proposed a new hybrid deep learning model for predicting gold price and movement, aiming to capitalize on the complementary strengths of different architectures. The model was designed to exploit CNN for extracting relevant features from time-series data and LSTM layers for identifying temporal dependencies. Preliminary experimental analysis suggested that combining LSTM with convolutional layers offered a significant boost to forecasting performance compared to state-of-the-art baseline models.

Highlighting gold's role as a hedge investment, in another effort, Billah and Das [20] proposed a hybrid method combining 1D-CNN with bidirectional gated recurrent units (Bi-GRU) for forecasting gold prices. The authors asserted that while various hybrid and standalone models provided satisfactory results, their proposed 1D-CNN-BiGRU approach proved more reliable. Their experiments showed it outperformed other tested networks (including CNN-gated recurrent unit (GRU), CNN-LSTM, CNN-RNN, and standalone variants) based on multiple evaluation metrics like mean absolute error (MAE), RMSE, and achieved determination (R^2) score of over 93%.

Further innovations in data representation include in the work [21], where gold price data was transformed into images using the gramian angular field (GAF) technique before being processed by a CNN-LSTM network. Beyond structural hybrids, research has also integrated external factors to refine predictions. For instance, Varela *et al.* [22] incorporated FinBERT-based sentiment analysis of financial news with high-frequency technical indicators into an LSTM framework, which improved real-time trading decisions for gold/USD exchange rates.

A parallel development in the literature focuses on “divide-and-conquer” strategies, where complex price signals are decomposed into simpler components before modeling. Liang *et al.* [23] introduced the

ILCC model, which combined improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) decomposition with a joint LSTM-CNN-convolutional block attention module (CBAM) architecture. Their results indicated that the integration of feature enhancement via the CBAM significantly increased prediction precision.

The effectiveness of these decomposition techniques was further validated by Song *et al.* [24], who utilized a VMD-ensemble empirical mode decomposition (EEMD) technique across various futures contracts (including gold and zinc). Their empirical evidence showed that multiple price decompositions significantly reduced error metrics, such as MAE and RMSE, compared to single-decomposition models. Finally, to address limitations of point-data models, Yang *et al.* [25] developed interval-valued forecasting system for non-ferrous metals. This system utilized a two-stage ensemble learning approach with temporal fusion transformers and AMs, providing more robust insights into price ranges and market risks than traditional methods.

In essence, the reviewed literature shows a clear evolution from traditional statistical methods towards machine learning and increasingly sophisticated deep learning models for gold price forecasting, with hybrid architectures combining CNN and RNN elements showing particular promise. Despite this progress, a focused comparison evaluating 1D-CNN paired specifically with standard RNN, LSTM, and BiLSTM for short-term gold price prediction under various input window configurations remains a gap in the current research. This study directly addresses this gap by proposing and rigorously evaluating the STGP-Net (1D-CNN-LSTM) model against these specific hybrid counterparts across different experimental settings.

3. THE SHORT-TERM GOLD PREDICTION NETWORK MODEL ARCHITECTURE

This section details the methodology employed for short-term gold price forecasting. It begins by introducing the overall architecture of the STGP-Net model, which integrates 1D-CNN and LSTM networks. Subsequently, the sliding window technique used to prepare the time series data into appropriate input sequences for this model is described. Following this, the individual components of the hybrid model (1D-CNN and LSTM) are elaborated upon. Finally, the algorithm that utilizes the proposed model for prediction is presented.

Figure 1 illustrates the proposed STGP-Net architecture designed for short-term gold price prediction. Input sequences, created from historical gold price data using a sliding window technique, are initially fed into the 1D-CNN component. This module comprises an input layer followed by three blocks, each containing a convolutional layer and a pooling layer, to extract hierarchical temporal features and patterns within the input window. The resulting feature maps from the CNN module serve as input to LSTM layer, which is tasked with modeling longer-range sequential dependencies present in the extracted features.

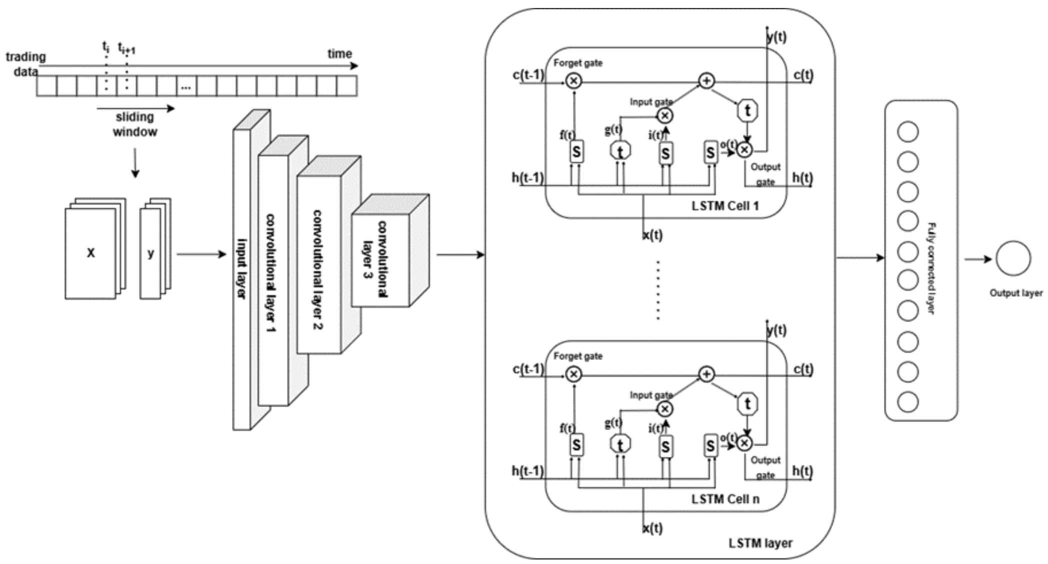

Figure 1. STGP-Net architecture

Finally, the output sequence from the LSTM layer is processed by a fully connected (dense) output layer. This type of layer, where each neuron receives input from all neurons in the previous layer, effectively maps the learned high-level sequential features to the final predicted gold price(s). This architecture aims to enhance prediction performance by combining the 1D-CNN's proficiency in local pattern extraction with the LSTM's capability for modeling temporal dynamics.

3.1. Time series segmentation using sliding windows

Given the historical transactional dataset of gold as $X = \{x_t | \ x_t \in R^n; t = 0, \pm 1, \pm 2, \cdots\}$. This time series dataset ($X \in R^{m \times n}$) maintains the chronological order of observations. Each data point is denoted as $x_t \in R^n$ represents the prices of a gold transaction at a specific time *t*. To prepare this sequential data for input into our predictive model, a sliding window method is applied. This technique extracts fixed-size subsequences (windows) by moving across the original dataset. The mathematical definition of the sliding window method is presented in (1).

$$\varphi(x_t, w): R^{m \times n} \rightarrow R^{k \times w \times n}, \forall x_t \in X \quad (1)$$

Where $X \in R^{m \times n}$, *w* is the window size, and $k \ll m$.

In short, the function $\varphi(x_t, w)$ maps the original dataset from a tensor with rank 2 ($R^{m \times n}$) to a tensor with rank 3 ($R^{k \times w \times n}$). The tensor resulting from the sliding window operation is well-suited for subsequent processing. It aids model training and feature engineering, assists in capturing short-term temporal patterns, and can contribute to reduced computation. The specific implementation of the sliding window process on the time series data is outlined in Algorithm 1.

Algorithm 1. Sliding window
```
function slideWindow(data, window_size, forecast_size)
        data_size ← |data|
        X ← [ ]
        y ← [ ]
        for i from window_size to data_size - window_size - forecast_size:
                X ← X ∪ data[i − window_size − forecast_size : i − forecast_size, : ]
                y ← y ∪ data[i − forecast_size : i, : ]
        return  X, y
end function
```

3.2. One-dimensional convolutional neural network component for feature extraction

1D-CNN is a specialized CNN architecture tailored for sequential data, like time series or natural language, focusing on identifying patterns along the temporal dimension. A typical 1D-CNN comprises several key layers: convolutional layers are primarily responsible for extracting relevant local features from the input sequences. Activation functions applied within these layers introduce non-linearity, enabling the network to model complex patterns. Pooling layers are commonly included to reduce the dimensionality of the feature maps, which can improve computational efficiency and robustness. In the context of this research, the 1D-CNN module is integrated with an LSTM module.

Formally, a 1D convolutional operation calculates an output feature $y(t)$ at timestep *t* by applying a kernel *w* of size *k* to a corresponding segment of the input sequence *x*, as shown in (2).

$$y(t) = (x * w)(t) = \sum_{i=0}^{k-1} x(t + i) \cdot w(i) \quad (2)$$

Where $*$ denotes the convolution operation, $x(t + i)$ is the $(t + i)th$ input element, and $w(i)$ is the $i - th$ element of the kernel. The rectified linear unit (ReLU) serves as the activation function within this hybrid model; its formula is presented in (3).

$$ReLU(z) = \max(0, z) \quad (3)$$

The max pooling operation, utilized within the hybrid model, is formally defined by (4).

$$y(t) = \max(x(t), x(t + 1), \cdots, x(t + p - 1)) \quad (4)$$

The 1D-CNN module is trained by adjusting its filter weights via backpropagation. A loss function first quantifies the error between the predicted and actual outputs. Backpropagation then calculates the gradients of this loss relative to the network's weights. These gradients guide the weight update step, typically executed using the adaptive moment estimation (Adam) optimizer, to minimize the error. The specific calculation for the convolutional layer gradients is defined in (5).

$$\frac{\partial L}{\partial w(i)} = \sum_t \frac{\partial L}{\partial y(t)} \cdot x(t+i) \quad (5)$$

Where $y(t) = \sum_{i=0}^{k-1} x(t+i) \cdot w(i)$ and $\frac{\partial L}{\partial y(t)}$ is the gradient of the loss with respect to the output.

3.3. Long short-term memory component for temporal modeling

In this hybrid model, the LSTM layer, composed of multiple LSTM cells, is designed to capture long-range dependencies within the gold price sequences and to help mitigate the vanishing gradient problem common in simpler recurrent networks. The specific internal structure of an LSTM cell is visualized in Figure 2. The core of an LSTM cell includes forget, input, and output gates that manage its state. The forget gate selectively removes information from the previous cell state (c_t), while the input gate adds relevant new information. The output gate determines the next hidden state (h_t) - the cell's output for the timestep - by filtering the updated c_t. These operations can be formally described by the tuple $\mathcal{L}_c = \langle f_t^c, i_t^c, o_t^c, c, \widetilde{c}, h_t \rangle$, whose components are detailed in Table 1.

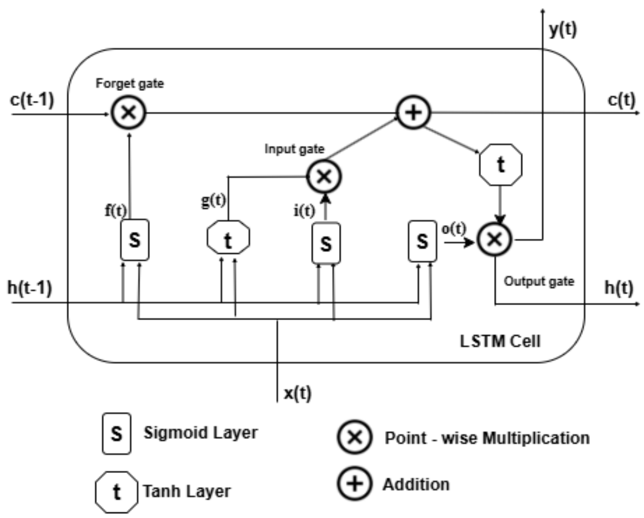

Figure 2. An LSTM cell structure

Table 1. Mathematical description of LSTM cell components

Math symbols	Meaning
$x(t)$	The input vector to the LSTM cell
W_d	The input-weight matrix where *d* can be *f, i, o,* or *c*.
U_d	The matrix of the recurrent connections where *d* can be *f, i, o* or *c*.
b_d	The bias vector where *d* can be *f, i, o,* or *c*.
σ	The sigmoid function
τ	The hyperbolic tangent function
$f_t^c = \sigma\left(W_f x(t) + U_f h(t-1) + b_f\right)$	The forget gate of the memory cell *c* at time *t*.
$i_t^c = \sigma(W_i x(t) + U_i h(t-1) + b_i)$	The input gate of the memory cell *c* at time *t*.
$o_t^c = \sigma(W_o x(t) + U_o h(t-1) + b_o)$	The output gate of the memory cell *c* at time *t*.
$\widetilde{c(t)} = \tau(W_c x(t) + U_c h(t-1) + b_c)$	The cell input activation vector at the time step *t*.
$c(t) = f_t^c \odot c(t-1) + i_t^c \odot \widetilde{c(t)}$	The c_t vector at the time step *t*.
$h(t) = o_t^c \odot \tau(c(t))$	The output vector of the LSTM cell (or the h_t vector) at the time step *t*.
$\odot$	This operator denotes the element-wise product.

3.4. Short-term gold prediction network prediction procedure

Algorithm 2 presents the methodology for short-term gold price forecasting using the developed hybrid 1D-CNN-LSTM architecture. The procedure involves processing historical gold price data, applying a sliding window technique to generate input segments, and leveraging the trained model to predict prices across a designated forecast horizon. The algorithm encapsulates the necessary data preprocessing, model initialization, and prediction generation steps.

Algorithm 2. STGP-Net predicts next n-day gold prices
```
function predict_Next_N_Prices(file, window_size, forecast_size)
         timeseries_data ← read_data(file)
         tensors_data ← slideWindow(timeseries_data, window_size, forecast_size)
         model ← initialize_hybrid_model()
        results ← model.predict(tensors_data.X)
        return results
       end function
```

Specifically, Algorithm 2 begins by loading the historical gold price data. This data is then processed by the slideWindow function, using specified window_size and forecast_size parameters, to create appropriately shaped input tensors. Next, the pre-trained STGP-Net model is initialized (i.e., loaded). These prepared input tensors are fed into the model's predict method to generate the predicted gold price values for the next forecast_size time steps. Finally, the algorithm returns these predicted results, providing the short-term forecast for gold prices. This outlines an efficient sequence of data preparation, model execution, and forecast generation.

4. EXPERIMENT

The empirical basis for this research is a comprehensive dataset of historical gold prices obtained from the Yahoo Finance service at https://finance.yahoo.com/. Utilizing the yfinance Python library at https://pypi.org/project/yfinance/, daily records including open, high, low, close, and volume were collected for the period spanning August 30, 2000, to April 8, 2025. This dataset served as the foundation for training and evaluating the proposed short-term prediction models. All experiments were performed within the Google Colab environment at https://colab.research.google.com/, which was selected for its significant advantages, including free access to powerful computing resources like graphics processing units (GPUs), essential for accelerating deep learning model training, and its pre-configured setup with major data science and machine learning libraries readily available.

To prepare the raw time series data for input into the hybrid deep learning models, a sliding window technique was employed. This approach transforms the sequential data into overlapping windows of fixed size, where each window serves as an input sequence to predict a subsequent sequence of a specified forecast size. To investigate the impact of varying input lengths and prediction horizons, three distinct configurations were established. Configuration 1 utilized a window size of 10 days to predict the next 3 days. Configuration 2 used a window size of 15 days to forecast the subsequent 5 days. Finally, configuration 3 involved a window size of 20 days for predicting the next 7 days. This process resulted in three separate window-based datasets, each tailored to one of these configurations. For the purpose of model training, validation, and final evaluation, each of these three datasets was consistently partitioned into a training set (70%), a validation set (15%), and a test set (15%). This 70%-15%-15% split was applied uniformly across all three configurations to ensure comparable evaluation conditions. Figures 3 to 5 illustrate the training and validation loss curves for the three evaluated hybrid deep learning models (1D-CNN+RNN, STGP-Net, and 1D-CNN+BiLSTM) across configurations 1, 2, and 3, respectively.

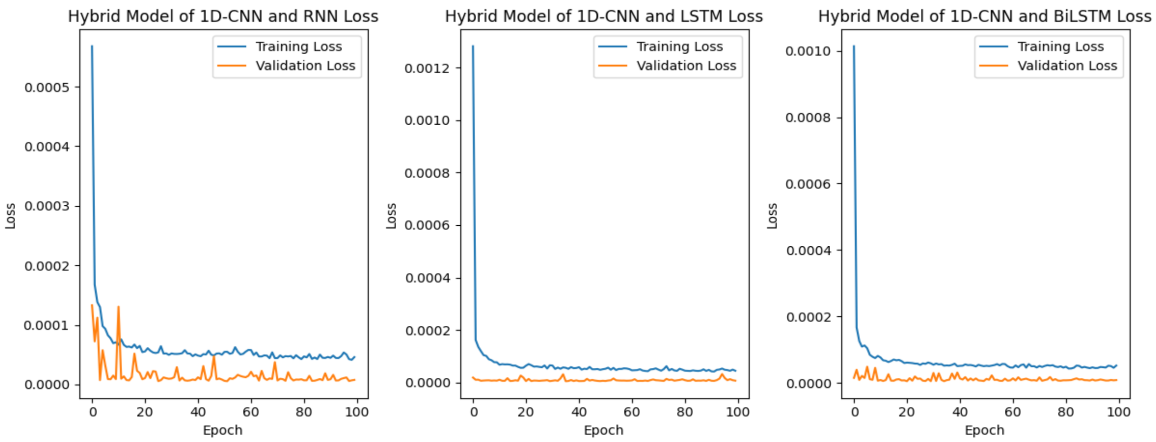

Figure 3. Loss curves of configuration 1

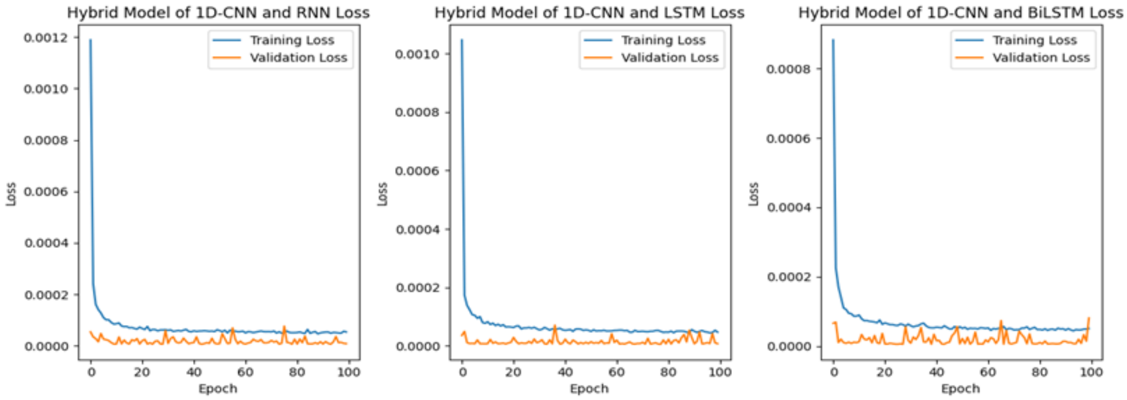

Figure 4. Loss curves of configuration 2

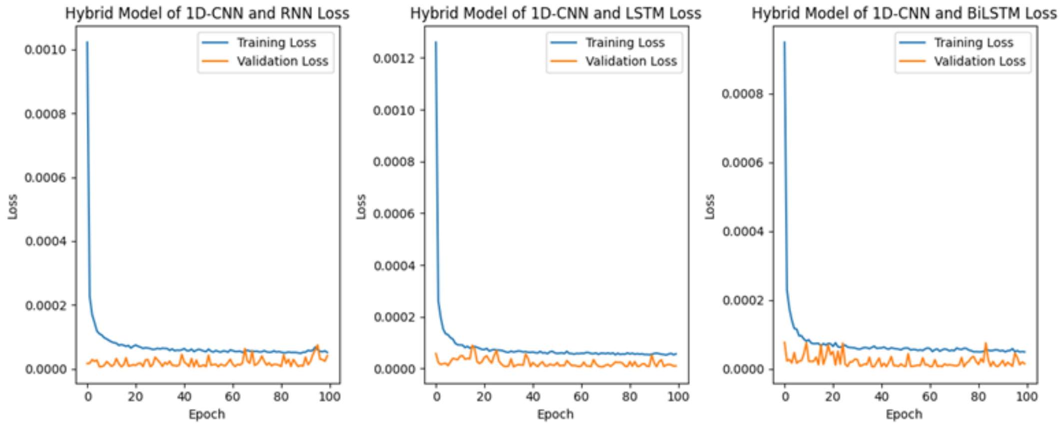

Figure 5. Loss curves of configuration 3

Observing the loss curves presented in Figures 3 to 5, all three hybrid models exhibit a characteristic rapid decrease in both training and validation loss during the initial epochs across all configurations, indicating successful learning initiation. Subsequently, the loss values for both training and validation tend to stabilize, suggesting convergence towards a minimum. Notably, the gap between the training and validation loss remains relatively small and does not significantly widen after the initial stabilization phase for most models and configurations, implying that substantial overfitting did not occur within the 100 training epochs shown. This general behavior holds across the different window and forecast settings represented by the three configurations, demonstrating consistent learning dynamics for the evaluated models under these experimental conditions.

To quantitatively assess the predictive accuracy and goodness-of-fit of the different hybrid models, four standard evaluation metrics were employed: MAE in (6), mean squared error (MSE) in (7), RMSE in (8), and the R^2 in (9).

$$MSE(y,\hat{y}) = \frac{1}{N}\sum_{i=0}^{N-1}(y_i - \hat{y}_i)^2 \tag{6}$$

$$MAE(y,\hat{y}) = \frac{1}{N}\sum_{i=0}^{N-1}|y_i - \hat{y}_i| \tag{7}$$

$$RMSE(y,\hat{y}) = \sqrt{MSE(y,\hat{y})} \tag{8}$$

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=0}^{N-1}(y_i - \hat{y}_i)^2}{\sum_{i=0}^{N-1}(y_i - \bar{y}_i)^2} \tag{9}$$

Where N is the total samples; y and $\hat{y}$ are true and predicted values, respectively; $\bar{y} = \frac{1}{N}\sum_{i=0}^{N-1} y_i$.

Lower values for MAE, MSE, and RMSE indicate smaller prediction errors, while an R^2 value closer to 1 signifies a better fit of the model to the actual data. The computed values for these metrics, comparing the performance of the 1D-CNN+RNN, STGP-Net (1D-CNN+LSTM), and 1D-CNN+BiLSTM models across all three experimental configurations (configurations 1, 2, and 3) on the test set. The results are summarized in Table 2.

Table 2. Model performances across three configurations

Configuration	Model	MAE	MSE	RMSE	R^2
Configuration 1	1D-CNN and RNN	0.001899	0.005799	0.043583	-0.002297
	STGP-Net	0.001815	0.007254	0.042602	0.042427
	1D-CNN and BiLSTM	0.001763	0.005955	0.041984	0.07879
Configuration 2	1D-CNN and RNN	0.00117	0.004272	0.034207	0.18883
	STGP-Net	0.001137	0.004445	0.033715	0.198889
	1D-CNN and BiLSTM	0.001253	0.01016	0.035393	-0.329788
Configuration 3	1D-CNN and RNN	0.000914	0.009388	0.030232	-0.705088
	STGP-Net	0.000872	0.003265	0.029532	0.3419
	1D-CNN and BiLSTM	0.00087	0.004031	0.029496	0.236335

An analysis of the quantitative results presented in Table 2 reveals distinct performance differences among the evaluated hybrid models across the three configurations. While the 1D-CNN+BiLSTM model showed the lowest MAE and RMSE in configuration 1, its performance significantly degraded in configuration 2, exhibiting the highest errors and a negative R^2 value. Conversely, the 1D-CNN+RNN model performed poorly in configurations 1 and 3, indicated by negative R^2 scores, although it achieved reasonable metrics in configuration 2. The proposed STGP-Net (1D-CNN+LSTM) model consistently demonstrated strong performance; it achieved the lowest MAE/RMSE and highest R^2 in configuration 2, the highest R^2 in configuration 3 (with MAE/RMSE values very close to the best), and competitive results in configuration 1. Based on this comprehensive evaluation across multiple metrics and configurations, the STGP-Net model emerges as the most robust and effective predictor, displaying superior overall performance and greater consistency compared to the 1D-CNN+RNN and 1D-CNN+BiLSTM hybrid models.

To further assess the predictive capabilities, Figures 6 to 8 plot the forecasted values from the 1D-CNN+RNN, STGP-Net, and 1D-CNN+BiLSTM models alongside the true historical gold prices for each of the three configurations. The visual results presented in Figures 6 to 8 further support the quantitative findings. While all models attempt to capture the trend of the gold price, the plots visually confirm the inconsistencies observed in the metrics for the 1D-CNN+RNN and 1D-CNN+BiLSTM models, particularly evident in configurations 2 and 3 where their predictions sometimes deviate significantly from the true values. The STGP-Net (1D-CNN+LSTM) model's predictions generally appear to track the actual price movements more closely across the configurations, aligning with its stronger performance metrics.

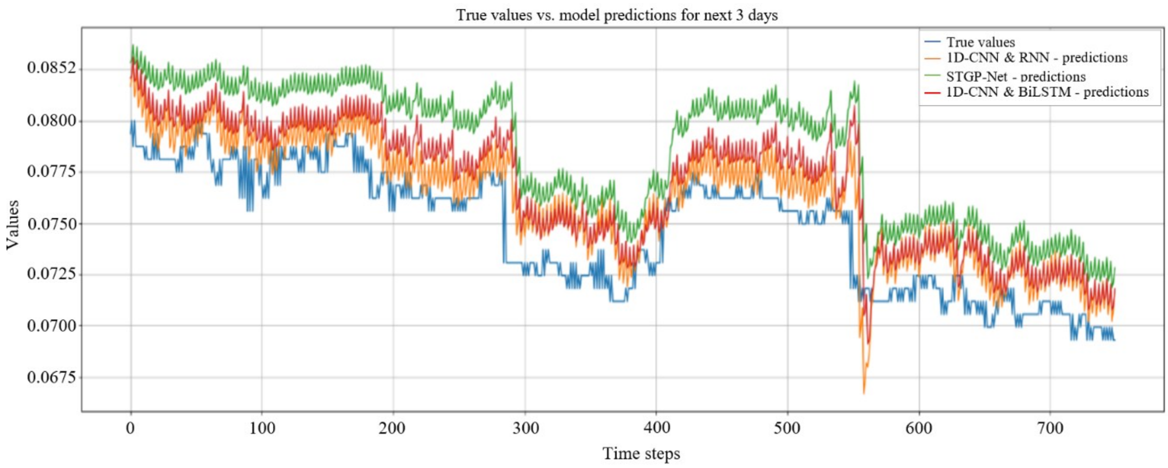

Figure 6. True values vs. predicted values of configuration 1

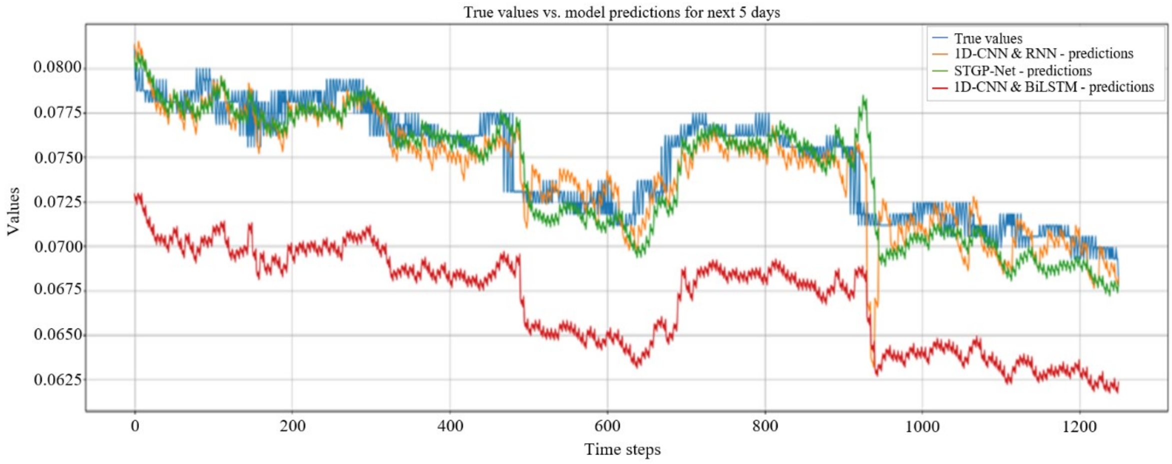

Figure 7. True values vs. predicted values of configuration 2

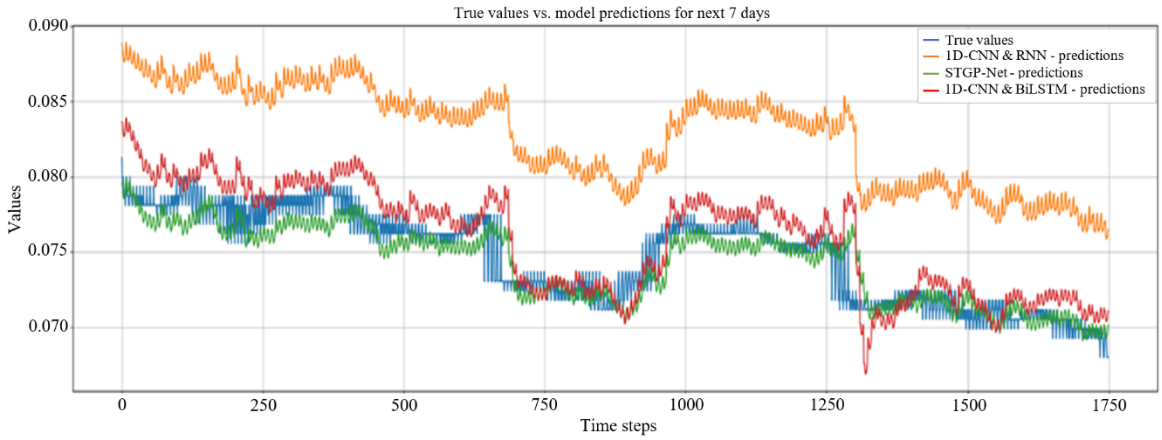

Figure 8. True values vs. predicted values of configuration 3

In summary, this experimental section detailed the data preparation using sliding windows, the evaluation framework including metrics and loss analysis, and the comparative performance assessment of three hybrid models. The comprehensive results consistently indicate that the proposed STGP-Net model provides enhanced and more reliable short-term gold price forecasting under the tested conditions. Compared to the 1D-CNN hybrid models utilizing standard RNN or BiLSTM layers, STGP-Net demonstrated superior performance.

5. CONCLUSION

This study investigated hybrid deep learning models for short-term gold price forecasting, proposing STGP-Net (a 1D-CNN-LSTM architecture) and comparing it against 1D-CNN combined with standard RNN and BiLSTM across various window configurations. Comprehensive experimental results consistently showed that STGP-Net outperformed the baseline models in terms of predictive accuracy and robustness based on MAE, RMSE, and R^2 metrics. These findings establish the proposed 1D-CNN-LSTM combination as an effective approach for this challenging forecasting task. Future research could extend this work by incorporating additional features like macroeconomic indicators or market sentiment, exploring more advanced hyperparameter optimization methods, or adapting the model for longer-term prediction horizons.

FUNDING INFORMATION

This study was funded by Hue University, Vietnam (grant number DHH2025-01-225).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Hoang Ha Nguyen	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓
Minh Duc Nguyen		✓				✓		✓	✓	✓	✓			
Cuong H. Nguyen-Dinh	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

This research did not involve human participants or animals.

DATA AVAILABILITY

The data that support the findings of this study are available at https://finance.yahoo.com/.

REFERENCES

[1] G. Meena, K. K. Mohbey, S. Kumar, and K. Lokesh, "A hybrid deep learning approach for detecting sentiment polarities and knowledge graph representation on monkeypox tweets," *Decision Analytics Journal*, vol. 7, Jun. 2023, doi: 10.1016/j.dajour.2023.100243.

[2] S. Tan, B. Chen, J. Zeng, B. Li, and J. Huang, "Hybrid deep-learning framework for object-based forgery detection in video," *Signal Processing: Image Communication*, vol. 105, Jul. 2022, doi: 10.1016/j.image.2022.116695.

[3] R. Akter, M. R. Islam, S. K. Debnath, P. K. Sarker, and M. K. Uddin, "A hybrid CNN-LSTM model for environmental sound classification: Leveraging feature engineering and transfer learning," *Digital Signal Processing: A Review Journal*, vol. 163, Aug. 2025, doi: 10.1016/j.dsp.2025.105234.

[4] D. Deshwal, P. Sangwan, N. Dahiya, U. K. Lilhore, S. Dalal, and S. Simaiya, "COVID-19 detection using hybrid CNN-RNN architecture with transfer learning from X-rays," *Current Medical Imaging Reviews*, vol. 20, Jan. 2023, doi: 10.2174/1573405620666230817092337.

[5] W. Chen, W. Hussain, F. Cauteruccio, and X. Zhang, "Deep learning for financial time series prediction: a state-of-the-art review of standalone and hybrid models," *CMES - Computer Modeling in Engineering and Sciences*, vol. 139, no. 1, pp. 187–224, Dec. 2023, doi: 10.32604/cmes.2023.031388.

[6] P. Peng, Y. Chen, W. Lin, and J. Z. Wang, "Attention-based CNN–LSTM for high-frequency multiple cryptocurrency trend prediction," *Expert Systems with Applications*, vol. 237, 2024, doi: 10.1016/j.eswa.2023.121520.

[7] H. Harbaoui and E. A. Elhadjamor, "Enhancing stock price prediction: LSTM-RNN fusion model," *Procedia Computer Science*, vol. 246, no. C, pp. 920–929, Jan. 2024, doi: 10.1016/j.procs.2024.09.511.

[8] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Computing and Applications*, vol. 32, no. 23, pp. 17351–17360, Dec. 2020, doi: 10.1007/s00521-020-04867-x.

[9] A. O. Ige and M. Sibiya, "State-of-the-art in 1D convolutional neural networks: a survey," *IEEE Access*, vol. 12, pp. 144082–144105, 2024, doi: 10.1109/ACCESS.2024.3433513.

[10] D. Makala and Z. Li, "Prediction of gold price with ARIMA and SVM," in *Journal of Physics: Conference Series*, 2021, vol. 1767, no. 1, doi: 10.1088/1742-6596/1767/1/012022.

[11] I. A.-Doush, B. Ahmed, M. A. Awadallah, M. A. Al-Betar, and A. R. Rababaah, "Enhancing multilayer perceptron neural network using archive-based Harris hawks optimizer to predict gold prices," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 5, May 2023, doi: 10.1016/j.jksuci.2023.101557.

[12] I. Ghosh and R. K. Jana, "A granular machine learning framework for forecasting high-frequency financial market variables during the recent black swan event," *Technological Forecasting and Social Change*, vol. 194, Sep. 2023, doi: 10.1016/j.techfore.2023.122719.

[13] X. Song and Z. S. Chen, "Enhancing financial time series forecasting in the shipping market: a hybrid approach with light gradient boosting machine," *Engineering Applications of Artificial Intelligence*, vol. 136, Oct. 2024, doi: 10.1016/j.engappai.2024.108942.

[14] A. Bhambu, R. Gao, and P. N. Suganthan, "Recurrent ensemble random vector functional link neural network for financial time series forecasting," *Applied Soft Computing*, vol. 161, Aug. 2024, doi: 10.1016/j.asoc.2024.111759.

[15] G. Cohen and A. Aiche, "Forecasting gold price using machine learning methodologies," *Chaos, Solitons and Fractals*, vol. 175, Oct. 2023, doi: 10.1016/j.chaos.2023.114079.

[16] S. Das, J. Nayak, B. K. Rao, K. Vakula, and A. R. Routray, "Gold price forecasting using machine learning techniques: review of a decade," in *Computational Intelligence in Pattern Recognition*. Singapore: Springer, 2022, pp. 679–695, doi: 10.1007/978-981-16-2543-5_58.

[17] Y. C. Chen and W. C. Huang, "Constructing a stock-price forecast CNN model with gold and crude oil indicators," *Applied Soft Computing*, vol. 112, 2021, doi: 10.1016/j.asoc.2021.107760.

[18] K. Huang, X. Li, F. Liu, X. Yang, and W. Yu, "ML-GAT: a multilevel graph attention model for stock prediction," *IEEE Access*, vol. 10, pp. 86408–86422, 2022, doi: 10.1109/ACCESS.2022.3199008.

[19] J. Wang, Y. Li, T. Wang, J. Li, H. Wang, and P. Liu, "A gold futures price forecast model based on SGRU-AM," *IEEE Access*, vol. 9, pp. 146745–146754, 2021, doi: 10.1109/ACCESS.2021.3122140.

[20] M. M. Billah and S. Das, "Analysis and prediction of gold price using CNN and Bi-GRU based neural network model," in *24th International Conference on Computer and Information Technology, ICCIT 2021*, Oct. 2021, pp. 1–6, doi: 10.1109/ICCIT54785.2021.9689880.

[21] M. Salim and A. Djunaidy, "Development of a CNN-LSTM approach with images as time-series data representation for predicting gold prices," *Procedia Computer Science*, vol. 234, pp. 333–340, Jan. 2024, doi: 10.1016/j.procs.2024.03.007.

[22] A. Varela, M. K. Siam, A. A. Maruf, H. Gu, J. Q. Cheng, and Z. Aung, "An analytical framework for real-time gold trading using sentiment and time-series forecasting," *Decision Analytics Journal*, vol. 17, Dec. 2025, doi: 10.1016/j.dajour.2025.100633.

[23] Y. Liang, Y. Lin, and Q. Lu, "Forecasting gold price using a novel hybrid model with ICEEMDAN and LSTM-CNN-CBAM," *Expert Systems with Applications*, vol. 206, 2022, doi: 10.1016/j.eswa.2022.117847.

[24] Y. Song, J. Huang, Y. Xu, J. Ruan, and M. Zhu, "Multi-decomposition in deep learning models for futures price prediction," *Expert Systems with Applications*, vol. 246, Jul. 2024, doi: 10.1016/j.eswa.2024.123171.

[25] W. Yang, H. Zhang, J. Wang, and Y. Hao, "A new perspective on non-ferrous metal price forecasting: an interpretable two-stage ensemble learning-based interval-valued forecasting system," *Advanced Engineering Informatics*, vol. 65, May 2025, doi: 10.1016/j.aei.2025.103267.

BIOGRAPHIES OF AUTHORS

Hoang Ha Nguyen is currently the head of the Faculty of Information Technology, University of Sciences, Hue University, Vietnam. He received the doctor degree of Computer Science in 2017 from Hue University. His bachelor and master degrees in Computer Science were also issued by Hue University in 1999 and 2005, respectively. His research interests include (but not limited to): swarm optimization, financial data analysis, and deep learning. He can be contacted at email: nguyenhoangha@hueuni.edu.vn.

Minh Duc Nguyen is a lecturer at the Department of Economic Information Systems, University of Economics, Hue University, Hue, Vietnam. He received his M.Eng. and Ph.D. degrees in Information Systems from Inje University, Korea, in 2017 and 2021, respectively. His research interests include data mining, ontology engineering, semantic technologies, knowledge management, and blockchain systems. He can be contacted at email: nguyenminhduc@hueuni.edu.vn.

Cuong H. Nguyen-Dinh received his doctorate in Computer Science from Khon Kaen University in 2016. Currently, he serves as the lecturer of the Department of Data Science at University of Finance - Marketing, Vietnam. His primary research interests encompass deep learning, text mining, knowledge management systems, and recommender systems. He can be contacted at email: cuongndh@ufm.edu.vn.