Financial Engineering and Risk Management (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2024.070216
ISSN 2523-2576 Vol. 7 Num. 2

Gold Price Prediction Based on CNN-LSTM

Jiaqi Li

School of Statistics, Southwestern University of Finance and Economics, Chengdu, 611130, China

Keywords: Gold Futures Price, CNN-LSTM, Time Series Forecast

Abstract: Gold plays a significant role in global economic markets, and its price forecasts are of great significance to investors, policymakers, and economic analysts. The serious nonlinearity and high noise characteristics of financial series make it difficult to accurately predict their trends by traditional statistical models. This paper will select the spot price of gold in London as the original data, and make prediction based on the CNN-LSTM combination model, and make comparative analysis with SVM, LSTM and Ridge Regression model. The results show that the MAE of CNN-LSTM model is 2.1758, MSE is 9.4822, $R^2$is 0.9487, and the model can capture the price trend more effectively, so as to provide more accurate prediction results.

1. Introduction

Gold is a valuable metal with good electrical conductivity, ductility and long-term preservation, which has a wide range of uses in jewelry, electronics and other industrial fields[1]. At the same time, it has the triple attributes of money, investment and commodity. As a result, gold prices are seen as a barometer of economic health. As a reserve asset of countries, gold helps to maintain currency stability, so it plays an important role in monetary policy and international trade[2]. At the same time, as a safe haven asset[3], gold becomes a safe haven for funds when the stock market is in turmoil, and is an important part of portfolio risk diversification. In recent years, the price of gold has shown great volatility due to the influence of various factors [4]. Therefore, accurate prediction of the price of gold is of great significance to investors, policy makers and economic analysts, and has also become the focus of the majority of scholars and researchers.

The volatility of gold market price makes the accurate prediction of gold price trend become a key topic for scholars at home and abroad. At present, domestic and foreign scholars have made a lot of exploration on the prediction of gold spot price, from linear model to nonlinear model [5], from traditional econometric model [6] to single machine learning model [7] and even complex deep learning model[8].In an experiment, Yifan[9]and colleagues used the ARMA model to predict the gold price in the future, but it was unable to explain the changes. Ferdinandus[10]and colleagues utilized the LSTM model to forecast the price of gold, and it was able to provide a more accurate prediction than the ARIMA model. Through empirical analysis and comparison of the performance of ARIMA model, GM (1,1) model and multi-factor BP neural network model in predicting gold price, Cheng Ming [11] found that the combined prediction model based on genetic algorithm had the best effect in predicting gold price. According to Han Xu et al. [12], the model delivers better predictions than the traditional method by combining quotient space theory and support vector machines. Wavelet neural networks were used by Zhang Long et al. [13] to forecast gold spot prices using additional momentum, and they found that the wavelet neural networks had good prediction efficiency. The GM (1, N) grey model is improved by Huang et al. [14] using a neural network model coupled with BP to view the characteristics of gold price fluctuations. The model was proved to have a higher application value than the GM (1, N) grey model alone. In summary, in order to find a more accurate way to predict the spot price of gold, this paper will establish a CNN-LSTM model and compare the model with SVM, LSTM and Ridge Regression models.

2. Introduction to the theory

In this paper, gold price prediction is performed using CNN-LSTM deep learning model, a combination of convolutional neural networks (CNN) and long short-term memory models (LSTM).

2.1 CNN model

CNN models, or Convolutional Neural Networks, are an efficient artificial neural network structure in deep learning that can be used to recognize and process images, as well as predict continuous data. The CNN model mainly consists of an input layer, a convolution layer, an activation function, a pooling layer, a fully connected layer and an output layer. The data passes through the input layer and is then stored in the convolution layer through an activation function in the convolution layer. Then the pooling layer performs feature extraction, and the spatial dimension of the feature map is reduced in the extraction process, that is, downsampling is performed. This helps reduce the number of parameters in the network, prevents overfitting, and makes feature detection insensitive to small position changes. After several convolution and pooling layers, the fully connected layer maps the high-level features in the network to the label space of the samples, i.e. H.i.e., for the classification task as shown in Figure 1.

The CNN model has the characteristics of weight sharing, which means that a number of connections can have the same weight, and local connection, which means that the neurons in each layer are only connected to some neurons in the next layer. These two features help improve the prediction accuracy of the model.

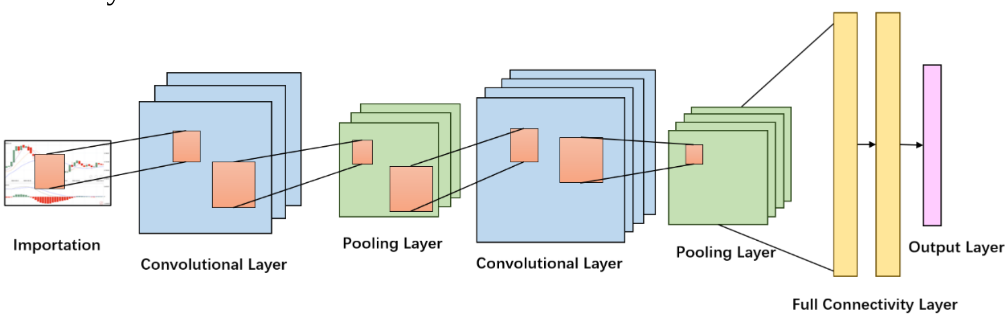

Figure 1: Structure of the CNN model

2.2 LSTM model

The LSTM model is a special type of RNN model that is commonly used when it is necessary to predict a time series over a longer period of time. The LSTM model consists of three types: an input gate, an amnesia gate, and an output gate, as well as a cellular state C_t for transmitting information, as shown in Figure 2. The incoming data first enters the input gate, and then some of the data is filtered out of the cell state through the forgetting gate. Here, the activation function in the Forgetting Gate ensures that the Forgetting Gate outputs a function value between 0 and 1. When the forgetting gate receives new information, the forgetting gate will superimpose a part of the previous information with the new information with a certain probability to form a new input information, and then multiply it with the new cell state one by one to determine whether its value is 1. If it is 1, its information is retained, otherwise, its information is discarded. To a certain extent, this can reduce the amount of data in the training process of the model. The next step in the LSTM model is to determine what information needs to be stored in the cell state, which is a two-step process. Firstly, we need to identify the information that needs to be updated, and then we need to determine what information needs to be added to the cell state. The formulas for these two processes are shown in (1), (2), and (3):

$$i_t = \sigma(W_{ii}h_{t-1} + U_{ix}x_t + b_i) \tag{1}$$

$$\tilde{C}t = tanh(Wcih_{t-1} + U_{cx}x_t + b_c) \tag{2}$$

$$C_t = C_{t-1} \otimes f_t + i_t \otimes \widetilde{C_t} \tag{3}$$

Finally, we need to determine the h_t of the output information, as shown in (4) and (5):

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{4}$$

$$h_t = o_t \otimes tanh(C_t) \tag{5}$$

In this process, we can see that the LSTM model can effectively solve the problem of gradient vanishing and explosion compared with the RNN model.

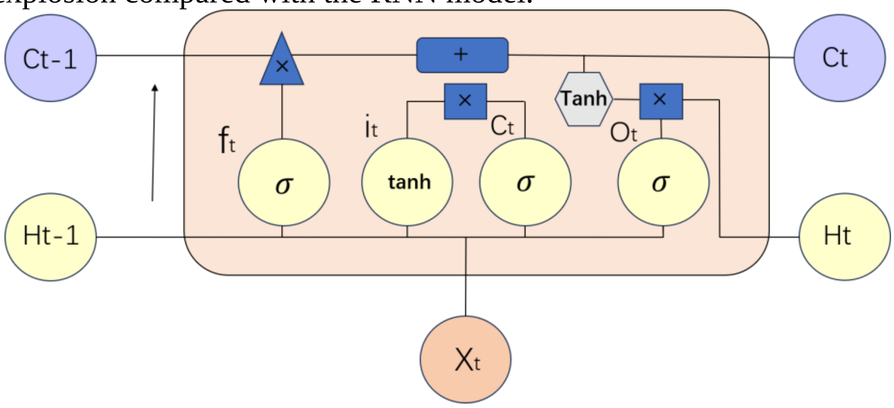

Figure 2: Structural diagram of the LSTM

2.3 CNN-LSTM combination model

CNN is used to extract features from time series, and the convolutional layer in CNN can effectively capture features in the input information, and then pass these features to the LSTM, which is then used for modeling to further extract valid features in the time series data. The concrete model structure is shown in figure 3.

CNN-LSTM steps:

1) Data extraction and preprocessing

First, the model samples the input data through a window, and then normalizes the sampled data, i.e., averages

The value variance is normalized as follows:

$$X' = \frac{X-\mu}{\sigma} \tag{6}$$

where μ represents the average value of the data set, while σ denotes the data's standard deviation.

2) Capture the spatial features of time series on the CNN layer

Convolution kernels slide across time series, and each convolution kernel is used to extract a different feature, the size of the window (convolutional kernel).and quantity are important parameters on the CNN layer. The size of the window determines the range of captured features, and the number of windows determines the kind of captured feature.

3) CNN output data goes into the LSTM

The CNN output data goes through some operations and is then transmitted to the LSTM as time series data, which the LSTM can receive.

4) Further processing of LSTM

The LSTM performs time-dependent modeling on the CNN output data as input data for further processing.

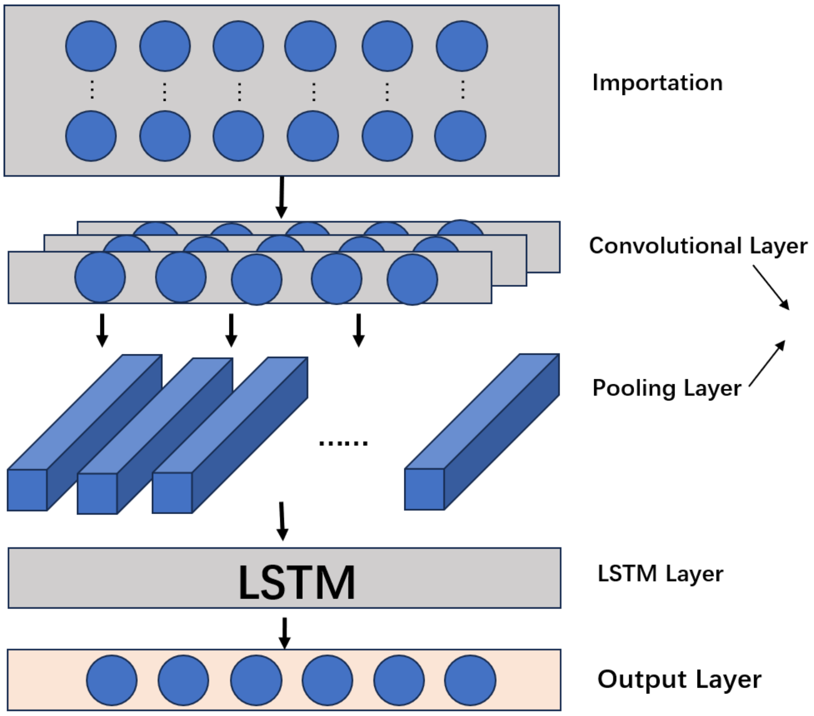

Figure 3: Structure diagram of CNN-LSTM model

3. Prediction analysis and comparison

3.1 Data preprocessing

This paper takes the spot price of gold in Shanghai as the research object, and obtains the gold price (denominated in yuan) from the wind database from January 4, 2021 to February 2, 2024 for forecast analysis, The time series diagram of the data is shown in Figure 4, and the descriptive statistics are shown in Table 1.

Table 1: Descriptive statistics of gold spot prices

Variable name	Sample size	maximum	minimum	mean	median	variance	kurtosis	skewness	Coefficient of Variation (CV)
price	730	483.33	354.2	407.8	394.56	1369.974	-0.998	0.652	0.091

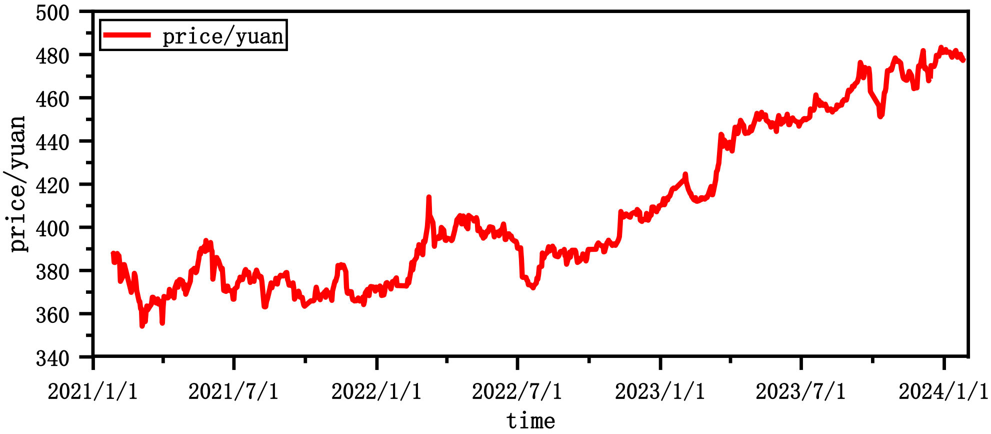

Figure 4: Timing chart of gold spot prices

To eliminate the influence of dimensional differences between variables in the model, the data were first normalized. According to the characteristics of the data, we choose the mean variance normalization method to solve the comparison problem between data, accelerate the convergence speed of the training network, and improve the prediction accuracy of the model.

$$x' = \frac{x - mean}{std} \tag{7}$$

where x is the original data and x' is the normalized data

Before establishing the CNN-LSTM model, to ensure that the model can effectively learn the characteristics of the data, it is essential to convert the time series data into a supervised learning problem. The detailed model flow is shown in Figure 5.

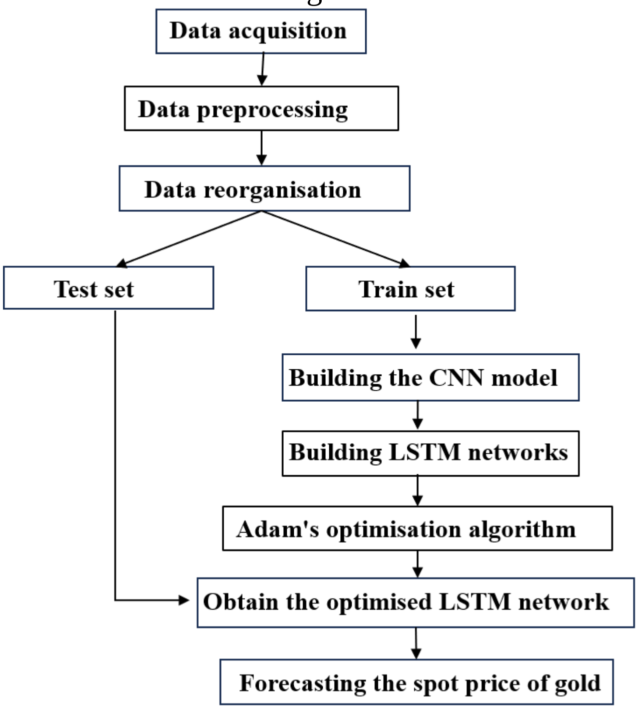

Figure 5: Flow chart of CNN-LSTM algorithm

3.2 Model evaluation index

The paper employs the Mean Square Error (MSE), the Mean Absolute Error (MAE), and the coefficient of determination (R^2) as performance metrics for the model, with their respective formulas provided as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y_i} - y_i)^2 \quad (8)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y_i} - y_i| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y_i})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (10)$$

where n is the number of samples, Y_i is the true value of the sequence, $\hat{Y_i}$ is the predicted value of the sequence, and $\bar{Y}$ is the mean value of the sequence.

3.3 Prediction and comparison of models

Firstly, the CNN architecture is constructed to derive characteristics from the time series data, and subsequently, the extracted attributes are fed into an LSTM network. Then, the LSTM network model is established, which consists of a hidden layer, a dropout layer, and a full connection layer; The architecture of the LSTM component consists of 2 layers, with each layer having 64 hidden units; The normalized MSE of the predicted value and the true value is used as the loss function; The Adam optimization algorithm is used to update the network parameters, and the initial learning rate is 0.0001, the attenuation rate based on the first-order moment of the gradient is 0.6, and the attenuation rate based on the second-order moment of the gradient is 0.999; The grid search was used to optimize the batch_size and training times, and the optimal batch_size was 63 and the optimal training times were 12,000. In order to further verify the effectiveness of the model, the LSTM model, support vector machine model (SVR) and Ridge Regression model were used to predict, and the comparative experiments were carried out, and the evaluation indexes on the test set were calculated.

The predictive effects of the different models are shown in Table 2.

Table 2: Evaluation indicators of different prediction models

	CNN-LSTM	LSTM	SVR	Ridge Regression
MAE	2.1758	5.1034	16.3863	2.5065
MSE	9.4822	34.5230	291.8686	12.3962
R^2	0.9487	0.8131	-0.5801	0.9329

In comparison to the LSTM, the MAE and MSE of the CNN-LSTM model are smaller, and R^2 is higher, The outcomes indicate that the CNN-LSTM model achieves a superior fitting performance over the LSTM; The MAE of the SVR model is 16.3863 and the MSE is 291.8686, indicating that the fitting effect of the CNN-LSTM model is better than that of the SVR model; The MAE of Ridge Regression is 2.5065 and 2.1758, and the MSE is 12.3962 and 9.4822, indicating that the fitting effect of the CNN-LSTM model is better than that of the Ridge Regression model. In summary, the prediction effect of the CNN-LSTM model is relatively accurate.

To provide a clearer visualization of the disparities in predictive outcomes among the various models, this paper presents a comparison of the predicted results against the actual values on the test set for all four models. As can be seen in Figure 6, these four models can predict the general trend of the data, but when the data volatility is large, the performance of the CNN-LSTM, LSTM and ridge regression models is significantly better than that of the SVM model, and the prediction data of the CNN-LSTM is closer to the true value. Through further comparison, it can be found that compared with other models, the CNN-LSTM model's predictions are more aligned with the true values, the error is relatively smaller, and the fitting effect of the model is better. The combination of the two can further illustrate that the CNN-LSTM model is better than the LSTM model, the ridge regression model and the SVM model in the prediction of gold price.

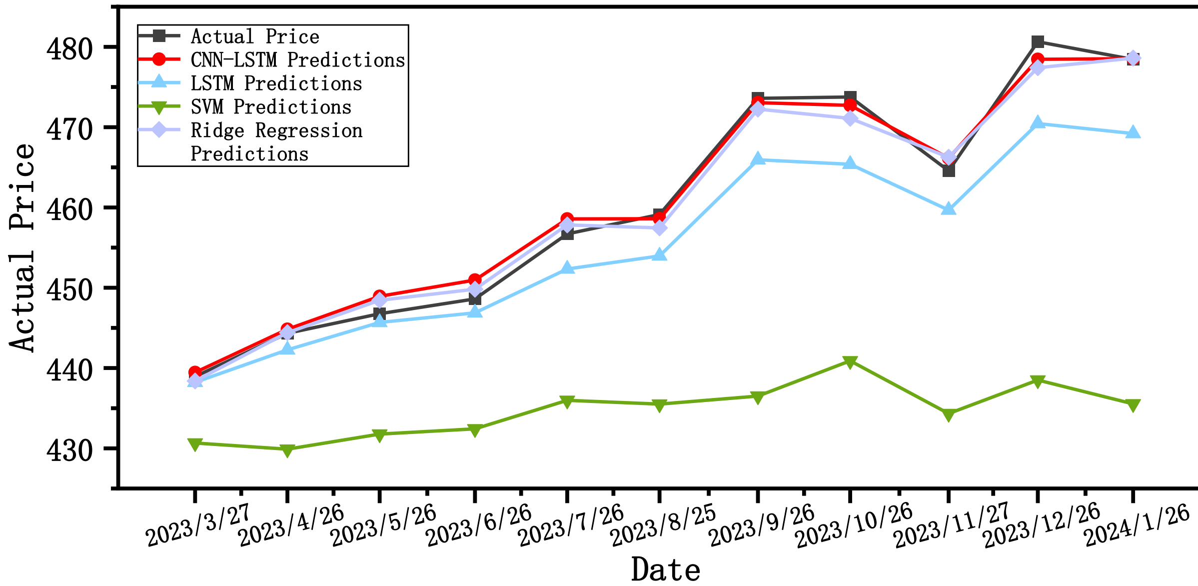

Figure 6: Comparison of prediction results of different models

4. Conclusions

In this paper, the Shanghai gold spot price is used as the prediction object, and the mean variance is normalized to the dataset. In view of the severe nonlinearity and high noise characteristics of financial series, this paper adopts the CNN-LSTM to predict the spot price, and optimizes the CNN-LSTM model through the Adam optimizer and network search method, in order to minimize the loss function and enhance the precision of the model. Simultaneously, compared with LSTM, SVR and Ridge Regression, it is confirmed that the CNN-LSTM has a smaller prediction error and a higher degree of fitting, and its accuracy is better than that of LSTM, SVR and Ridge Regression models, which can better predict the spot price of gold.

References

[1] Ding Lei. Application of Bayesian network in short-term prediction of gold price rise and fall [D]. Shanghai University of Finance and Economics, 2023.

[2] Mo Y, Ruotong W, Zixun Z, et al. Improved prediction of global gold prices: An innovative Hurst-reconfiguration-based machine learning approach [J]. Resources Policy, 2024, 88

[3] Gil C, Avishay A. Forecasting gold price using machine learning methodologies [J]. Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena, 2023, 175(P2)

[4] Zeng Li, Li Chun. Research on gold price prediction model by integrating linear and nonlinear models[J]. Gold, 2013, 34(03): 7-10.

[5] Hao Yajuan, Zhang Rongwang. Gold investment continues to heat up due to safe-haven demand and the "Spring Festival effect"[N]. China Business News, 2024-02-05 (B02).

[6] Si Jin. Short-term forecast of gold price based on ARIMA model [J]. Times Finance, 2016(17): 221-222.

[7] Cao X. Gold price prediction model based on SVM and its parameter optimization [D]. Shandong University, 2017.

[8] Li Mengze, Zhu Dongyuan. Empirical analysis of gold price trend prediction based on deep learning integrated approach (SDAE-B) model [J]. China Asset Appraisal, 2021(01):12-19.

[9] Fei Yifan, Tao Yuqian. Short-term forecast analysis of gold price based on ARMA model[J]. Times Finance, 2018(30): 236-237.

[10] Ferdinandus, Yudha Randa Madhika et al. Gold Price Prediction Using the ARIMA and LSTM Models. Sinkron, 2023.

[11] Cheng Ming. Research on the forecasting method of international gold price [D]. Shandong University, 2020.
[12] Han Xu, Yang Shan, Wang Ximei. SVM model prediction of gold price based on quotient space[J]. Gold Science and Technology, 2020, 28(01):148-157.
[13] Zhang Long, Zhang Xiaoxue. Spot gold price prediction by wavelet neural network based on additional momentum [J]. Journal of Changzhi College, 2022, 39(02): 124-127.
[14] Huang Qian, Liu Sheng. Research on gold price prediction based on IGM-BP model[J]. China Price, 2022, (10): 87-89.