Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 6, Issue 1: 14-18, DOI: 10.25236/AJCIS.2023.060103

Research on Multivariate Futures Price Forecasting Based on IMAT-LSTM Model

Wei Wei[1,a,*]

[1]*School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing, 400041, China*
[a]*leowwww@163.com*
[*]*Corresponding author: Wei Wei*

Abstract: *The study of futures price forecasting is of great significance to society and enterprises. The time series characteristics of futures prices are difficult to capture because of their non-stationary and nonlinear characteristics. This paper proposes an IMAT-LSTM model, which allocates the weight of the multivariable characteristics of futures prices through an improved attention mechanism, and comprehensively compares the prediction performance with LSTM and traditional AT-LSTM models to verify the effectiveness of the model proposed in this paper in futures price prediction. The performance of the model proposed in this paper is evaluated by fitting degree indicators (MAE, RMSE, R^2). The results show that, compared with LSTM and AT-LSTM models, the fitting accuracy of the proposed model is improved by 40.31% compared with the traditional LSTM model.*

Keywords: *Attention; LSTM; Futures price forecast; Hybrid model*

1. Introduction

The research on futures price forecasting has always been a hot topic in the financial field, which plays an important role in the government and society.

Futures price forecasting can obtain relatively good results through traditional time series methods, including regression conditional heteroscedasticity model (GARCH), differential integrated moving average autoregression model (ARIMA), seasonal ARIMA (SutteARIMA) and cubic exponential smoothing model (holt winter, HW). Faldzinski et al. [1] compared the prediction performance of the generalized autoregressive conditional heteroscedasticity (GARCH) model and the support vector regression (SVR) model for selected energy commodity futures contracts. The experimental results show that the prediction based on the asymmetric GARCH model is usually the most accurate. Rubio et al. [2] predicted the future value of Columbia Company through ARIMA and SVR hybrid model, and verified the effectiveness of the hybrid model. Ahmar, AS, etc. [3] conducted a prediction study on the seven day closing prices of the BRIC countries (China, Russia, Brazil, India) through ARIMA, Holt Winters, and SuttARIMA models. The study indicated that SuttARIMA is more suitable for forecasting the stock prices of Russia, India, and China, while Holt Winters is more suitable for forecasting the stock prices of Brazil. The traditional time series models used in futures price forecasting are all linear models. However, futures prices are highly nonlinear and highly volatile, so it is difficult for traditional forecasting models to capture the nonlinear characteristics of futures price data.

Long short term memory (LSTM) is a kind of recurrent neural networks (RNNs), which can capture the time series relationship and nonlinear characteristics of futures prices. Santos et al. [4] predicted the future price of ethanol through LSTM. In order to show the superiority of the model, they compared it with three benchmark algorithms: random forest, SVM linear and RBF. The results show that LSTM is superior to other models under the indicators MSE and MAPE. Based on the remarkable performance of the lstm model, researchers have conducted a lot of preprocessing on historical data, thus developing various lstm based models. HUANG et al. [5] used wavelet changes and singular value decomposition (SVD) methods to denoise the original data. The experimental results show that the model has a good prediction effect, and the prediction method has significant effectiveness. Attention mechanism is a weight allocation mechanism. Through supervised learning, adaptive weights can be allocated to multi feature data to strengthen the impact of key features on prediction results and eliminate the interference of non key features.

In response to the above problems, this paper proposes an improved LSTM model IMAT-LSTM based on ATTENTION mechanism, and uses gold futures prices to conduct empirical research on the model. Compared with LSTM and traditional AT-LSTM based on attention mechanism, the model proposed in this paper has achieved better results under MAE, RMSE and R^2 indicators.

2. Theoretical Research on IMAT-LSTM Model

2.1 Improved attention mechanism

The price of futures is affected by many factors. It is obviously not rigorous to assign the same weight to each feature that affects futures prices. Attention mechanism can capture the feature weight that affects futures prices, give more weight to key features, and weaken the feature information with low correlation to give it less weight. The Encoder Decoder structure after adding the Attention mechanism is shown in Fig.1 below.

X_i represents the input of the model, where C_i represents the weighted sum of hidden vector h in encoder encoder, and the specific formula is as follows:

$$C_i = \begin{cases} 0, \sum_{j=1}^{L_x} a_{ij} * h_j < 0.2 \\ \sum_{j=1}^{L_x} a_{ij} * h_j , else \end{cases} \tag{1}$$

Where, L_x represents the length of input data, a_{ij} represents the attention distribution coefficient of the jth word in the Decoder output sentence when the ith data is input in the Encoder layer, while h_j is the semantic code of the jth data in the Encoder input sentence. This paper overcomes the influence of irrelevant features in the traditional attention mechanism on the prediction results, and directly assigns the feature data with low weight to 0, without adding this feature information to the prediction of the model.

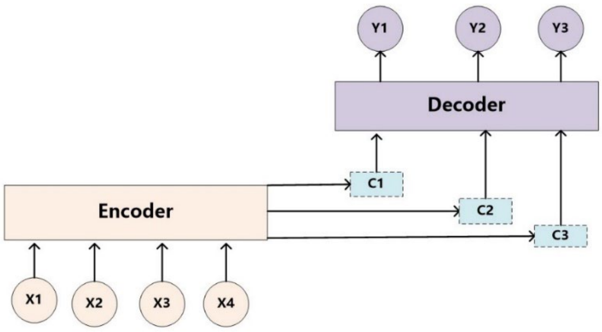

Figure 1: Attention mechanism

2.2 LSTM

Long term and short term memory network (LSTM) improves the cyclic neural network (RNN) through the 'gate' mechanism, solves the problem that once the RNN time series is too long, it is easy to cause gradient propagation, and improves the long term prediction performance. Figure 2 abstractly represents the basic structural unit of LSTM, which includes forgetting gate, input gate, output gate and storage unit. The forgetting gate is used to determine how much historical information is retained. The input gate is used to extract important information from the current input information and remove non important information. The output gate is used to determine how much current information can be output. The storage unit is responsible for storing important information. The basic steps of LSTM are as follows:

Suppose a set of financial time series $x = x_1, x_2, \dots, x_{T-1}, x_T (T > 0)$, where the data at time t is expressed as x_t. Input of LSTM includes current data x_t. Cell state at last momentc_{t-1} and the output value h at the previous time_(t-1)。 First, the t-1 time data in the input data passes through the forgetting

ISSN 2616-5775 Vol. 6, Issue 1: 14-18, DOI: 10.25236/AJCIS.2023.060103

gate f_t Determine the degree of retention.

$$f_t = \sigma(\omega_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

The input gate is composed of two parts. The first part passes the output data h of the previous time h_{t-1} and the current input data generate a new cell state $\tilde{c}_t$, The second part determines the cell state of the first part $\tilde{c}_t$ Reserved weight i_t

$$\tilde{c}_t = \tanh(\omega_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

$$i_t = \sigma(\omega_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

The retention degree of cell state c_{t-1} at the last moment is determined by the forgetting gate, and is determined by the input gate $\tilde{c}_t$ And finally add up to get the cell state c_t at time t:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{5}$$

Finally, the output level of the current cell state c_t is determined by the output gate o_t, and the result is taken as the output h_t at the current time.

$$o_t = \sigma(\omega_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{7}$$

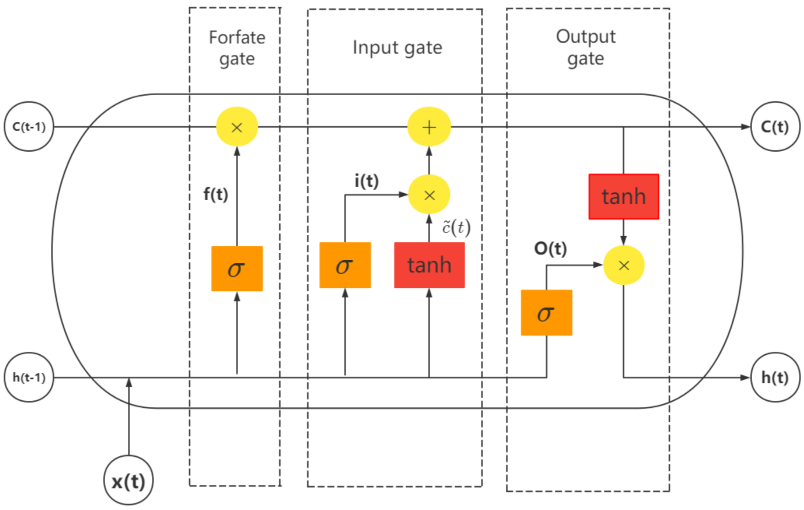

Figure 2: LSTM Model Structure

3. Experiment and analysis

3.1 Data description and preprocessing

In order to verify the effectiveness of the model proposed in this paper in futures price forecasting, the gold futures price is used as the empirical research object. The time range is from January 2010 to May 2022. The variables include opening price, highest price, lowest price, closing price and trading volume. The above data are obtained through the wind platform.

For the experiment, the data is divided into training set and test set according to 8:2. In order to reduce the impact of outliers in the nonlinear model training stage, the data is normalized to the range of [0,1]:

$$x(t)' = \frac{x(t) - \min x(t)}{\max x(t) - \min x(t)} \tag{8}$$

Finally, the actual predicted value is obtained through the inverse normalization method and compared with the actual value intuitively. The inverse normalization process is as follows:

$$x(t) = x(t)'[\max x(t) - \min x(t)] + \min x(t) \tag{9}$$

ISSN 2616-5775 Vol. 6, Issue 1: 14-18, DOI: 10.25236/AJCIS.2023.060103

3.2 Evaluation index

In order to evaluate the performance of VMD-LSTM-HW model, mean absolute error (MAE), root mean square error (RMSE) and determination coefficient (R^2) are used.

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y_t - \hat{y}_t\right| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \tag{11}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{N}(\bar{y} - \hat{y}_t)^2} \tag{12}$$

3.3 Experimental evaluation

In this part, in order to verify the effectiveness of the model in futures price forecasting, other models are used for comparison, including the traditional single forecasting model LSTM and the unimproved Attention LSTM. The experimental results are shown in Table 1. The experimental results verify the effectiveness of IMAT-LSTM model in futures price forecasting, and it has superior forecasting ability.

Table 1: Prediction performance evaluation of gold futures model

Models	MAE	RMSE	R^2(%)
LSTM	0.0210	0.0387	92.5351
AT-LSTM	0.0198	0.0311	97.4667
OURS	**0.0127**	**0.0231**	**98.5606**

Table 1 shows the forecasting performance of different models on gold futures prices. Compared with other models, the models proposed in this paper have the best results under the mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2) of 0.0127, 0.0231 and 98.5606%, respectively. The experimental results show that the model proposed in this paper can be effectively used in the prediction of futures prices.

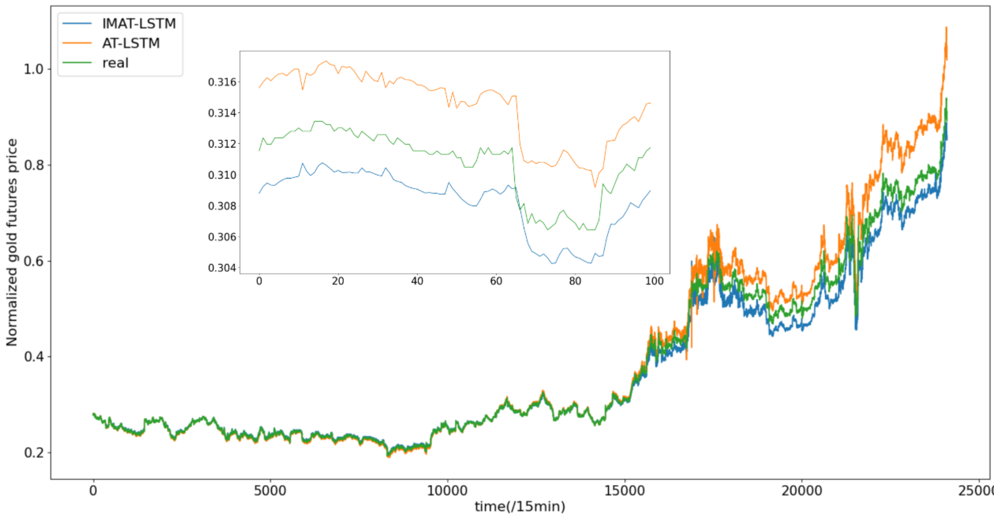

Figure 3: Prediction results of gold futures prices by different models

Figure 3 shows the prediction results of IMAT-LSTM model and other comparative models on gold futures prices. It can be seen that the predicted value of IMAT-LSTM is usually the closest to the true value, and it also shows superior performance in the prediction direction, which means it has better prediction accuracy.

ISSN 2616-5775 Vol. 6, Issue 1: 14-18, DOI: 10.25236/AJCIS.2023.060103

4. Conclusion

In this study, the improved Attention mechanism is used to assign weight to multiple features of empirical data. In combination with LSTM's ability to capture sequence features, the IMAT-LSTM model presented in this paper has better comprehensive prediction ability in horizontal prediction indicators (MAE, RMSE, R^2) than LSTM, AT-LSTM and other models. The experimental results show that the improved Attention mechanism can improve the prediction accuracy of the model, the empirical study of gold futures price shows the effectiveness of the model proposed in this paper.

References

[1] Faldzinski, M, Fiszeder P, Orzeszko W. Forecasting volatility of energy commodities: Comparison of GARCH models with support vector regression [J]. Energies, 2020, 14(1): 6.

[2] Rubio L, Alba K. Forecasting Selected Colombian Shares Using a Hybrid ARIMA-SVR Model [J]. Mathematics, 2022, 10(13): 2181.

[3] AHMAR A S, SINGH P K, VAN T N, et al. Prediction of BRIC stock price using ARIMA, SutteARIMA, and holt-winters[J]. Computers, Materials and Continua, 2021: 523-534.

[4] Santos G C, Barboza F, Veiga A C P, et al. Forecasting Brazilian Ethanol Spot Prices Using LSTM [J]. Energies, 2021, 14(23): 7987.

[5] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 1998, 454(1971): 903-995.