Financial Asset Volatility Forecasting using LSTM with Intraday High-Low Price Information

Yaohui Bai*[1a], Ping Bai[2b], Xinyan Zhang[3c], Huayang Li[1d]

[*] Corresponding author: [a]baiyaohui@jxufe.edu.cn
E-mail addresses: [b]bai7510@163.com, [c]1200600390@jxufe.edu.cn, [d]dariusli@163.com

[1]School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang, China;
[2]Library of Jiangxi University of Finance and Economics, Nanchang, China
[3]School of Finance and Public Administration, Jiangxi University of Finance and Economics, Nanchang, China

Abstract: In recent years, predicting the volatility of financial assets has received increasing attention due to the continuous development and increased volatility of financial markets. In this paper, we propose a volatility prediction model based on the Long Short-Term Memory (LSTM) model in deep learning, which considers the intraday high and low prices in financial asset sequences. To compare the performance of the proposed model, we also use the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model for comparison. By analyzing the data of the Shanghai Composite Index from January 1, 2015 to December 31, 2019, the results show that the proposed model outperforms the GARCH model, which is also reflected in the root-mean-square error of the two models. The proposed method exhibits promising results in predicting asset volatility and highlights the potential of LSTM models in the field of financial asset volatility prediction.

Keywords: volatility forecasting, GARCH, LSTM, financial asset, high-low price

1. INTRODUCTION

The prediction of financial asset volatility is a crucial issue in risk management and investment decision-making. Volatility can be understood as the range of fluctuation in asset prices. High volatility implies a significant change in asset prices, while low volatility indicates small price variations. Predicting volatility is important for investors to assess the risk level of their portfolio, make informed investment decisions, and accurately price options and financial derivatives.

In recent years, numerous methods have been developed for predicting financial asset volatility, and statistical models have long been the primary tools used for this purpose. Among the various statistical models, generalized autoregressive conditional heteroskedasticity (GARCH) models such as ARCH[1], GARCH[2], and EGARCH[3] have gained widespread popularity due to their ability to capture the time-varying nature of volatility. GARCH models estimate the conditional variance of the asset returns based on historical data and are particularly suitable for modeling the persistence and clustering of volatility. Moreover, GARCH models can handle various types of financial data, including

MSEA 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2334349

high-frequency data, and have been widely applied in various fields, such as risk management, option pricing, and portfolio optimization. Despite the success of GARCH models, they also face several challenges, including the difficulties in modeling the complex and non-linear patterns of volatility and the computational burden of estimating the model parameters. Therefore, there is a growing need for developing more advanced and efficient models to improve the accuracy and efficiency of financial asset volatility prediction.

With the recent advancements in machine learning and deep learning, researchers have been exploring various non-linear models for predicting volatility. These models include neural networks[4], support vector machines[5], random forests[6], and deep learning models such as convolutional neural networks[7] and recurrent neural networks[8]. These models have shown great potential in predicting volatility and have become the focus of extensive research in the field of finance.

Among these models, the Long Short-Term Memory (LSTM) model[9,10] has emerged as a popular deep learning model for predicting volatility. The LSTM model has shown remarkable results in capturing the complex temporal dependencies and non-linear patterns that exist in financial time series data. In comparison to other models, LSTM has a unique ability to learn long-term dependencies in data, making it well-suited for the prediction of volatility, which is inherently a complex and non-linear problem. One of the key advantages of the LSTM model is its ability to incorporate past data to make future predictions. This is especially important in financial markets, where historical data plays a crucial role in forecasting future trends. Moreover, the LSTM model is highly customizable and can be adapted to a wide range of applications in finance, including stock market forecasting, risk management, and portfolio optimization. This flexibility has made the LSTM model a popular choice among financial analysts and investors.

In this paper, we propose an LSTM-based approach that incorporates daily high and low prices to predict asset volatility. Specifically, we use a sequence-to-sequence architecture that takes in the daily high, low and closing prices as input and predicts the corresponding asset volatility. By including daily high and low prices, we are able to capture more information about the underlying market dynamics and improve the predictive performance of our model. To evaluate the performance of our proposed model, we compare it with the GARCH estimation model using Shanghai Composite Index from January 1, 2015 to December 31, 2019, totaling 1,219 data points. Our empirical results demonstrate that our proposed model outperforms the GARCH model used for comparison in terms of out-of-sample forecasting accuracy. The proposed method exhibits promising results in predicting asset volatility and highlights the potential of LSTM models in the field of financial asset volatility prediction.

The remainder of this paper is organized as follows. Section 2 provides an overview of relevant theories, including the GARCH model and LSTM model. In Section 3, we present the experimental results and analysis. Finally, Section 4 summarizes the findings and outlines potential avenues for future research.

2. FORMATTING OF MANUSCRIPT COMPONENTS

2.1 GARCH model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is a popular approach for modeling financial asset volatility. The model was introduced by Bollerslev in 1986 and has since been widely used in financial research due to its ability to capture the time-varying nature of volatility.

The GARCH model is based on the assumption that the variance of an asset's returns is not constant over time but is instead a function of past returns and past variances. Specifically, the GARCH model is defined by the following equations:

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= \sigma_t \cdot z_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \end{aligned} \tag{1}$$

where r_t is the return on the asset at time t, μ_t is the expected return at time t, ε_t is the error term at time t, σ_t^2 is the variance of the error term at time t, and z_t is a standardized residual with zero mean and unit variance.

The GARCH model is a time-series model that estimates the variance of the asset's returns based on its own past returns and past variances, as well as the past returns and past variances of other assets. The model uses two sets of parameters: the first set, $\alpha_1, \cdots, \alpha_p$, measures the short-term impact of past returns on the variance, while the second set, $\beta_1, \cdots, \beta_q$, measures the long-term impact of past variances on the variance. The parameter ω is a constant term that represents the long-run average variance.

The GARCH model is estimated using maximum likelihood estimation (MLE) and can be used to forecast future volatility. The model has been widely used in finance and has been found to be effective in modeling the time-varying nature of financial asset volatility.

In summary, the GARCH model is a popular and effective approach for modeling financial asset volatility, based on the assumption that the variance of an asset’s returns is a function of past returns and past variances.

The model uses two sets of parameters to measure the short-term and long-term impact of past returns and past variances on the variance, and can be used to forecast future volatility.

2.2 LSTM model

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) that is designed to address the problem of vanishing gradients in traditional RNNs. LSTM has become increasingly popular in various fields, including finance, for time-series analysis and prediction.

The LSTM model is based on a memory cell, which can remember values for an extended period of time.

The memory cell is connected to three gates: the input gate, output gate, and forget gate. The input gate controls the information that enters the memory cell, while the output gate controls the information that exits the memory cell. The forget gate controls the information that is forgotten from the memory cell.

The LSTM model is trained using backpropagation through time, where the gradients of the loss function are propagated backwards through time to update the weights of the network. The architecture of the LSTM model is shown in Figure 1.

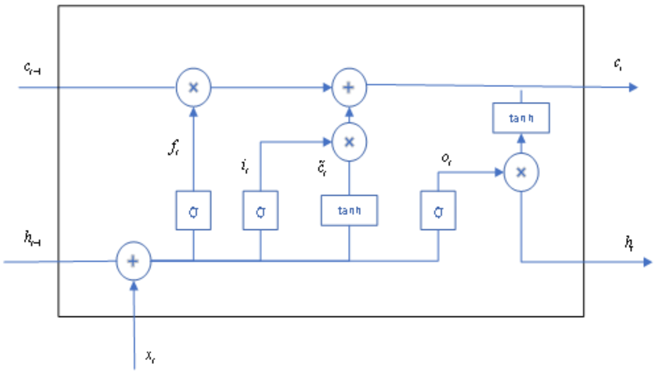


Figure 1. LSTM model architecture

Mathematically, the LSTM model can be represented as follows:

$$
\begin{aligned}
i_t &= \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + b_i\right) \\
f_t &= \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \\
o_t &= \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + b_o\right) \\
g_t &= \tanh\left(W_{xg}x_t + W_{hg}h_{t-1} + b_g\right)t \\
c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\
h_t &= o_t \odot \tanh\left(c_t\right) \\
\hat{y}_t &= W_{hy}h_t + b_y
\end{aligned}
\tag{2}
$$

where x_t is the input at time t, h_{t-1} is the hidden state at time t-1, i_t, f_t, o_t are the input, forget, and output gates, respectively, g_t is the candidate memory cell, c_t is the memory cell at time t, h_t is the hidden state at time t, $\hat{y}_t$ is the predicted value at time t, and σ is the sigmoid activation function.

In the context of asset volatility prediction, we use a sequence-to-sequence LSTM architecture that takes in the daily high, low, and closing prices as input and predicts the corresponding asset volatility. The input sequence is fed into the LSTM network, which outputs a hidden state at each time step. The hidden states are then used to make predictions about the future volatility of the asset.

The LSTM model is trained using the mean squared error (MSE) loss function, which measures the difference between the predicted and actual values of the asset volatility.

3. EXPERIMENTAL RESULTS AND ANALYSIS

This paper selected data from the Shanghai Stock Exchange Composite Index (SSECI) ranging from January 1, 2015, to December 31, 2019, totaling 1219 data points.To facilitate further analysis, the data is transformed into log-returns as shown in Figure 2a.

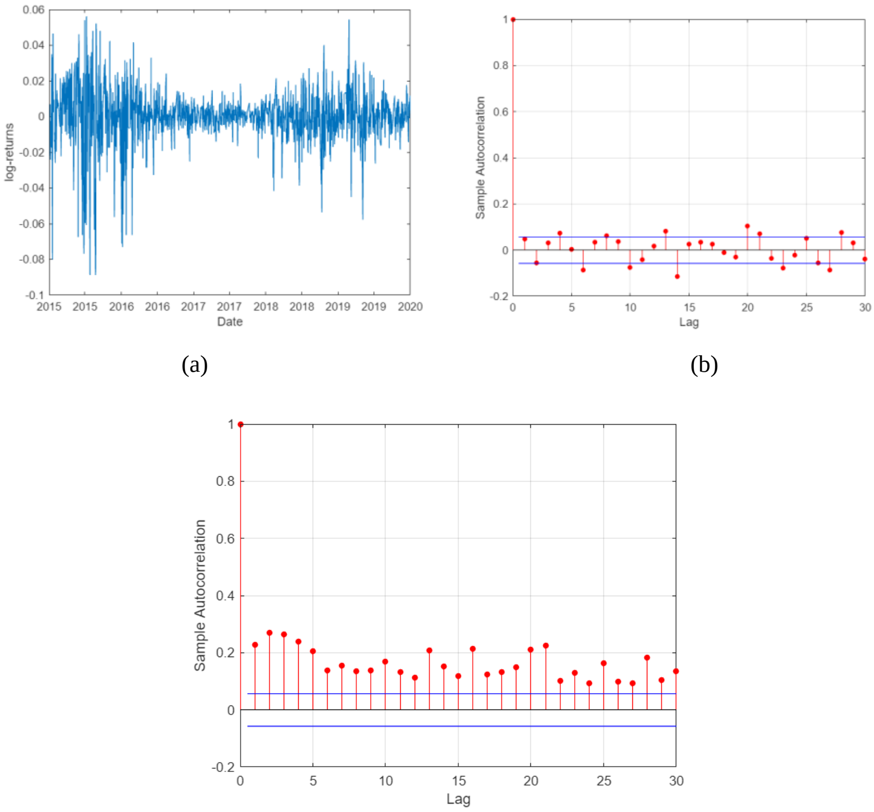
Table 1 presents the descriptive statistics of the SSECI log-returns sample. In addition to the mean, standard deviation, maximum, minimum, skewness, and kurtosis of the log-returns, the table also provides the q-statistic with a lag of 10 to test for autocorrelation in the log-returns and squared log-returns. The q-statistic for the 10th-order autocorrelation cumulative effect of the SSECI log-returns and squared log-returns is highly significant, indicating the presence of autocorrelation in the first and second moments of the log-return distribution.

Table 1. Descriptive statistics of the Shanghai Stock Exchange Composite Index.

mean	min	max	std	skewness	kurtosis	Q-test(10)	Q2-test(10)
0.0007	-0.0887	0.0560	0.0151	-1.2018	9.9652	38.3240	490.5336

Figures 2b and 2c display the lag-30 autocorrelation and squared autocorrelation graphs of the index series, respectively. Figure 2b shows the sample ACF of the log-return series, indicating no significant serial correlations except for a minor one at lag 11. Figure 2c shows the sample ACF of the squared log-returns, suggesting that the log-returns are not serially independent. Combining the two plots, it appears that the log-returns are indeed serially uncorrelated but dependent, indicating the presence of ARCH effects in the index series. Therefore, volatility models are required to capture this dependence in the return series.

To compare and analyze the modeling performance of the deep learning LSTM model, this study employs the GARCH (1, 1) model as the benchmark model. The fitting results are shown in Table 2, which includes the T-statistics and corresponding p-values of the model fitting.



(a) (b)

(c)

Figure 2. Time plot and statistics. (a) Time plot of daily log returns. (b) Sample ACF of log returns. (c) Sample ACF of the squared log returns.

Table 2. The fitting results of the GARCH model.

	Value	StandardError	TStatistic	PValue
Constant	5.008e-07	4.1208e-07	1.2153	0.22426
GARCH{1}	0.94302	0.0045315	208.1	0
ARCH{1}	0.054814	0.0055167	9.9361	2.8993e-23

Subsequently, this study uses the LSTM model to analyze the same SSECI data. First, the true volatility is calculated using a time window of length 30. Then, the SSECI data is split into two parts: one for training the model, forming the training set, and the other for testing the model, forming the testing set. 80% of the data is used for both training and testing sets, with 951 data points, and the remaining 20% of data is used for the validation set, comprising 237 data points. This study establishes a 5-layer LSTM network model, including a sequence input layer with an input dimension of 3, corresponding to the intraday high, low, and closing prices, two LSTM layers, a fully connected layer, and a regression layer. The number of hidden units in each LSTM layer is set to 100. During the network training process, the Adam algorithm is adopted for the stochastic gradient algorithm with a gradient threshold of 1. The mini-batch size is set to 64, the initial learning rate is 0.01, and every 125 training iterations, the learning rate is multiplied by a factor of 0.2 to reduce the learning rate. The maximum training iteration step is set to 500.

In this paper, we analyze the selected Shanghai Stock Exchange Composite Index data using the established GARCH(1,1) and LSTM models, and use the last 237 data points for comparison of predictive performance. The computed results are shown in Figure 3.

As shown in Figure 3, the actual volatility exhibits significant fluctuations at the beginning, followed by a relatively stable period between 50 and 150, and then increased volatility afterwards. The predicted curve of the LSTM model presents a similar trend, although with smaller fluctuations. In contrast, the predicted curve of the GARCH model fails to capture the same trend and shows an upward trend after 200. The root mean square error (RMSE) is calculated for both models, and the GARCH model is 0.0092, while the LSTM model is 0.0059, indicating that the LSTM model outperforms the GARCH model significantly. Based on the above results, it is evident that the LSTM model used in this paper performs better than the GARCH model in predicting volatility. This suggests that considering the intraday high and low prices can improve the predictive performance of volatility.

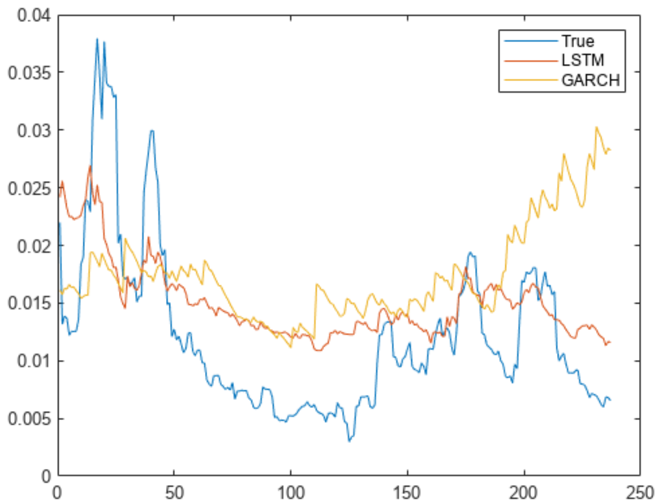


Figure 3. The predictive results of the LSTM and GARCH models

4. CONCLUSION

Financial asset volatility analysis and prediction have always been hot topics in the economic and financial research field. In this paper, we adopt the LSTM model, which is specifically designed for analyzing sequential data behavior in deep learning, and consider the intraday high and low price information to analyze and predict the volatility of the Shanghai Stock Index data. In addition, to compare the performance of the model, we selected the widely used GARCH model in volatility analysis for comparison. The calculation results show that the proposed LSTM model can effectively predict the volatility trend of time series in volatility prediction and its performance is better than that of the GARCH model. The root-mean-square error (RMSE) results of the two models also demonstrate this point. Finally, how to consider other trading information to improve the prediction performance of volatility will be the further work of this paper.

ACKNOWLEDGMENTS

This work was supported in part by the Humanities and Social Sciences Research Project of Jiangxi Colleges and Universities under Grant JC18101, and Grant GL19107.

REFERENCES

[1] Engle, Robert F.. “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.” Econometrica 50 (1982): 987-1007.

[2] Bollerslev, Tim. “Generalized autoregressive conditional heteroskedasticity.” Journal of Econometrics 31 (1986): 307-327.

[3] Nelson, Daniel B.. “CONDITIONAL HETEROSKEDASTICITY IN ASSET RETURNS: A NEW APPROACH.” Econometrica 59 (1991): 347-370.

[4] Su, Jung-bin. “How to Promote the Performance of Parametric Volatility Forecasts in the Stock Market? A Neural Networks Approach.” Entropy 23 (2021).

[5] Geng, Liyan et al. “Forecasting Range Volatility Using Support Vector Machines with Improved PSO Algorithms.” TELKOMNIKA Telecommunication Computing Electronics and Control 14 (2016): 208-216.

[6] Luong, Chuong and Nikolai Dokuchaev. “Forecasting of Realised Volatility with the Random Forests Algorithm.” Journal of Risk and Financial Management (2018).

[7] Dimitroff, Georgi et al. “Volatility Model Calibration With Convolutional Neural Networks.” Econometrics: Econometric & Statistical Methods - Special Topics eJournal (2018).

[8] Petneházi, Gábor and József Gáll. “Exploring the predictability of range-based volatility estimators using recurrent neural networks.” Intell. Syst. Account. Finance Manag. 26 (2019): 109-116.

[9] Hochreiter, Sepp and Jürgen Schmidhuber. “Long Short-Term Memory.” Neural Computation 9 (1997): 1735-1780.

[10] Greff, Klaus et al. “LSTM: A Search Space Odyssey.” IEEE Transactions on Neural Networks and Learning Systems 28 (2015): 2222-2232.