International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 757-767

https://doi.org/10.70517/ijhsa46469

The Future of Digital Currency: Combining Time Series Forecasting Models with Artificial Intelligence Algorithms for Bitcoin Exchange Rate Forecasting

Ruiyao Liu[1,*]
[1] University of Nottingham Ningbo China, Ningbo, Zhejiang 315100, China
Corresponding authors: (e-mail: hmyrl5@nottingham.edu.cn).

Abstract Bitcoin, as the most representative cryptocurrency, has an extremely volatile price, and traditional financial theories are difficult to fully explain its market behavior. The high volatility and complexity of Bitcoin's price pose a great challenge to investment decisions. This study proposes a bitcoin exchange rate prediction method based on a combined ARIMA-LSTM model, which improves the prediction accuracy by combining traditional time series analysis with deep learning techniques. Methodologically, an LSTM neural network is first constructed to capture the nonlinear characteristics of the bitcoin price, then an ARIMA model is built to analyze the linear trend, and finally the prediction results of the two models are optimally combined by using the CRITIC weight assignment method. The experiment uses the bitcoin closing price data from September 1, 2021 to December 31, 2024 for validation. The results show that the combined ARIMA-LSTM model significantly outperforms the single model in terms of forecasting performance, with a mean absolute error (MAE) of 0.0002, a root mean square error (RMSE) of 0.0003, and a mean absolute percentage error (MAPE) of 0.0006, which are 0.0071, 0.004, and 0.0051 lower than that of the ARIMA model, respectively. Empirical analysis shows that the combined model can more accurately capture the changing law of bitcoin price by integrating the advantages of linear and nonlinear prediction methods, which provides effective technical support for digital currency investment.

Index Terms Bitcoin exchange rate prediction, ARIMA-LSTM combined model, time series analysis, deep learning technology, CRITIC weight assignment, prediction accuracy

I. Introduction

Digital currencies first appeared in the 1990s, but because blockchain technology had not yet been fully invented at that time, it did not produce a more complete decentralized digital currency like Bitcoin [1], [2]. Modern currencies are backed by governments and built on the credit of the state, and the credit system of traditional fiat currencies was challenged after the financial crisis, so certain professionals started working on decentralized currency technology, and it was in this environment that Bitcoin was born [3], [4].

As of May 1, 2025, Bitcoin still has the largest market capitalization in the digital currency market as measured by market capitalization, while its popularity is also ranked first in the field of digital currencies. In March 2020, due to the impact of the New Crown Pneumonia outbreak, global investor panic ensued, the U.S. stock market plummeted and triggered four meltdowns, and the price of the just-stabilized Bitcoin was not spared, dropping by more than 50% in a few days, with the price reaching as low as $3,700 [5], [6]. In order to avoid the serious harm to the social and economic development caused by the new coronary pneumonia epidemic, quantitative easing has become the common choice of the world's major central banks, in particular, the Federal Reserve has continued to reduce the deposit rate and issued a large number of dollars, which makes the market has a large amount of idle funds into the stock market, commodities, real estate and other fields, and some of these funds enter the digital currency market [7]-[9].

By 2021, digital currencies were the subject of discussion globally, with the highest price of Bitcoin approaching $65,000 [10], [11]. Research on bitcoin exchange rate prediction gradually expanded. However, many methods have exposed limitations, such as the accuracy is difficult to guarantee, the prediction performance is insufficient, and the model suffers from overfitting in contexts such as unexpected events and sideways periods [12]-[14]. Time series forecasting and artificial intelligence, on the other hand, provide new horizons for this purpose. Time series analysis can be useful for exploring the spatio-temporal correlation of data as well as for multi-scale feature extraction, while artificial intelligence can achieve multimodal data fusion and dynamic data optimization [15], [16].

In this study, firstly, preprocessing and smoothness test are performed on bitcoin historical price data, and an applicable ARIMA model is constructed to capture the linear trend and cyclical characteristics of the price series; secondly, the LSTM neural network architecture is designed to utilize its powerful sequence learning ability to mine the nonlinear patterns and long-term dependencies in the price data; then, the CRITIC objective weight assignment method is adopted to determine the optimal weight configuration of the two models based on the principles of contrast strength and conflict; finally, the combined model is comprehensively evaluated by a number of evaluation indicators and compared with the benchmark model to validate the scientific integration of the results; lastly, the prediction performance is comprehensively assessed through a number of evaluation indicators and compared with the benchmark model. Then, the CRITIC objective weight assignment method is adopted to determine the optimal weight configuration of the two models based on the principles of contrast strength and conflict, and to realize the scientific fusion of the prediction results; finally, the prediction performance of the combined model is comprehensively evaluated by a number of evaluation indexes, and compared and analyzed with the benchmark model to validate the effectiveness and superiority of the proposed method.

II. Exchange rate forecasting models combining time series forecasting and artificial intelligence algorithms

II. A. LSTM-based time series modeling

II. A. 1) Time series modeling

ARIMA model [17] is the most commonly used model in practical cases of forecasting time series, this model is mainly for smooth non-white noise series data. The full name of ARIMA model in Chinese is Integrated Autoregressive Sliding Average Model.

The ARIMA model consists of three parts: one is the autoregressive model (AR), the second is the moving average model (MA), and the third is the difference model. The AR term is only used to predict the past value of the next value, which is defined by the p in the time series parameter. The MA term defines a number of past prediction errors in predicting the future value, which is denoted by the parameter q for the MA value. For the difference d, which is the order of the difference, the number of times the difference operation is performed on the sequence is specified. The purpose of performing difference operations on the data is to keep the time series smooth.

The ARIMA model, although due to its simple and easy to understand principle and easy to compute, the need for only autogenous variables makes it one of the most commonly used models for forecasting time series, it has some limitations. First of all, ARIMA model is based on linear data model, the input data must be smooth, the mean and variance cannot change over time, which means that there will be a great impact on the accuracy when dealing with nonlinear time series, and the unstable time series cannot be identified. Secondly, since ARIMA only uses past data to predict future values, the input time series data must be univariate.

II. A. 2) Principles of LSTM modeling

The original RNN network, with input x at different times on the left, hidden layer state s in the center, and network loss y on the right, LSTM adds a new time chain compared to RNN. Cell state is denoted by c, and this time chain is different from the hidden layer h state, where the hidden layer state update is rapid and the cell state update is slow. At the same time, LSTM increases the correlation between these two chains. Take moment t for example, compared with RNN, when calculating the Hidden Layer State h_t, in addition to the inputs x_i and the previous moment x_{i-1}, it also has to include the information of the current moment c_t. First the function f_t is like an eraser, the core of the composition of the forgetting gate, which decides which records are to be modified based on the memory of the previous moment h_{t-1} and the current input x_t, determining what records are to be erased from the previous state, described in mathematical language:

$$f_t = \sigma\left(W_f \cdot \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_f\right) \tag{1}$$

where σ is often a sigmoid function, taking values between 0 and 1. The multiplication of matrix elements erases those that take a value of zero, which is equivalent to selectively forgetting part of the memory, and is therefore often called a forgetting gate, like a valve that filters important features and ignores irrelevant information. The input gate is like a pencil, denoted here by f_i, that again decides which records to add based on the inputs from the previous moment and this moment, described in mathematical language as:

$$i_t = \sigma\left(W_i \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_i \right) \tag{2}$$

$$\tilde{c}_t = \tanh\left(W_c \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_f \right) \tag{3}$$

$$f_i = i_t * \tilde{c}_t = \sigma\left[W_i \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_i \right] * \tanh\left(W_c \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_f \right) \tag{4}$$

The sigmoid function, of which σ is often a part, again makes a selection of the contents, and the tanh function takes values between -1 and 1. This step is not an operation of forgetting, but rather amounts to generalizing and combing between these two points in time, and is therefore often referred to as an input gate. The records are first deleted with an eraser, and then the records are added with a pencil, and the two operations together, which we call Cell State, are represented by the formula: $c_i = f_i * c_{i-1} + i_i * \tilde{c}_i$, which results in the new record c_i, which processing will continue to pass down the line, and will also be used to update the current short-term memory h_i, and finally the output can be computed to get o_i, and h_i and o_i together make up the output gate, ie:

$$o_t = \sigma\left(W_o \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_o \right) \tag{5}$$

$$h_t = o_t * \tanh(c_t) \tag{6}$$

The short-term memory chain h_t and the long-term memory chain c_i are maintained at the same time and updated with each other, which is the principle of LSTM.

The LSTM model [18] introduces more parameter matrices, so it is a bit more cumbersome to train, but it can still be computed with the gradient descent algorithm. Due to the deep mining of interesting correlations in the temporal order of the data, LSTM mimics the brain in a way: focusing on the key pieces and ignoring irrelevant information, which greatly expands the applications of AI. Overfitting can easily occur when the network has too many and too large parameters while the data or number of simultaneous trainings is small, in Lstm computation dropout is one of the most effective methods to prevent overfitting of neural networks with regularization. Dropout takes a value between 0 and 1, and it is the proportion of neurons disconnected from the inputs that control the inputs to change linearly.

II. B.ARIMA model

II. B. 1) ARIMA model definition

ARIMA model is known as autoregressive difference moving average model, which as the most widely used linear time series forecasting model, contains three main processes, namely, the difference process (I), the autoregressive process (AR), and the moving average process (MA). In this section, AR, MA, ARMA, and ARIMA models are introduced in turn.

(1) p order autoregressive model $(AR_{(p)})$

For a set of random observation series, the traditional linear model can not directly describe the interdependence between the series, which gives rise to the autoregressive model that analyzes the correlation of data within the series, assuming that there exists a time series $\{x_1, x_2, ...x_t...\}$, the p-order autoregressive model can be expressed as:

$$x_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + ... + a_p x_{t-p} + \varepsilon_t \tag{7}$$

where $a_p \neq 0$, a_i is the autoregressive coefficient, and ε_i is the random error in order to ensure that the highest order is p.

(2) q -order moving average model $(MA_{(q)})$

For the time series $\{x_1, x_2, ...x_i, ...\}$, there exists a perturbation term $\{\varepsilon_1, \varepsilon_2, ...\varepsilon_i, ...\}$, if the moment t corresponds to the value x_i with its previous time $\{t-1, t-2, ...\}$ corresponding to the values $\{x_{i-1}, x_{i-2}, ...\}$ are not directly correlated, while there is some correlation with the perturbation $\{\varepsilon_{i-1}, \varepsilon_{i-2}, \ldots\}$ corresponding to the previous moments, and a model that fits into this scenario is known as a moving average model. If the highest order is q -order, it is also called a q -order moving average model, denoted as $MA_{(q)}$, and its structure can be expressed as:

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (8)$$

where, to ensure that the highest order is q, it is necessary to satisfy $\theta_q \neq 0, \theta_1, \theta_2, \ldots \theta_q$ are the moving coefficients.

(3) Autoregressive Moving Average Model $(ARMA_{(p,q)})$

For time series $\{x_1, x_2, \ldots x_i \ldots\}$, there is a perturbation term $\{\varepsilon_1, \varepsilon_2, \ldots \varepsilon_t, \ldots\}$, if the value x_i corresponding to moment t is related not only to the value $\{x_{i-1}, x_{i-2}, \ldots\}$ at the moment before it, but also to the perturbation $\{\varepsilon_{i-1}, \varepsilon_{i-2}, \ldots\}$ also have a certain correlation, and the model that satisfies this situation is called an autoregressive moving average model, and its structure is as follows:

$$x_t = a_0 + a_1 x_{t-1} + \ldots + a_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_p \varepsilon_{t-p} \quad (9)$$

where $a_p \neq 0$ and $\theta_q \neq 0$ are needed to be satisfied in order to ensure that the highest order is p and q, a_i is the autoregressive coefficient, and θ_i is the movement coefficient.

The ARIMA model refers to the fitting of the ARMA model series using the ARMA model series based on the differential treatment of the time series, which has the following structure:

(4) Autoregressive Differential Moving Average Model $(ARIMA_{(p,d,q)})$:

$$\omega(B)\nabla^d = \varphi(B)\varepsilon_t \quad (10)$$

where $\nabla^d = (1-B)^d$ denotes the d th order difference operation, $\omega(B)$ is the autoregressive coefficient polynomial of $ARMA_{(p,q)}$, and $\varphi(B)$ is the corresponding moving average coefficient polynomial.

II. B. 2) ARIMA model modeling process

The establishment of ARIMA model needs to go through three stages, which are, in order, smoothness test, model identification and order fixing, and white noise test.

(1) Model identification and order setting, mainly to determine the values of the three parameters of autoregressive term order (p), differential order (d) and moving average term order (q), the differential order can be determined after the time series for smoothness test. For the same time series prediction, there may be more than one significantly effective model, by determining the value of p and q parameters to select the model with the best prediction effect, p and q fixed order can be estimated by the value of the sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF).

① Autocorrelation function (ACF)

The $ACF(k)$ represents the linear correlation between the current time series values and the actual series values for the past k time intervals, and is calculated as follows:

$$ACF(k) = p(k) = \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t+k})}} \quad (11)$$

where k represents the lag order.

② Partial Autocorrelation Function (PACF)

For a smooth model, the lagged k autocorrelation coefficient $p(k)$ is not only the correlation between x_r and x_{t-k}, but also intermediate random variables, so the autocorrelation coefficient $p(k)$ is actually mixed with the influence of other variables on x_t and x_{t-k}. What $PACF(k)$ represents is the linear correlation between the current time series values and the actual series values for the past k time intervals, with the effects of the intermediate variables removed.

The ACF and PACF of the $ARIMA(p,d,q)$ model are characterized by a rapid convergence to zero after a certain order or remain consistently taken by non-zero values, a phenomenon referred to as truncation or trailing off, and is often used to determine the type of model. The values of the orders p and q can be selected by plotting the ACF and PACF as the maximum values of p and q at their initial truncation.

The minimum informatization criterion (AIC) establishes a system of indicators to determine the goodness of the model, which is an important criterion for evaluating the goodness of the model fit value, and the fit of the model can be judged based on the value of the AIC. The mathematical expression is as follows:

$$AIC = -2\ln(L) + 2N \tag{12}$$

where L is the likelihood function and N is the number of parameters.

When the number of samples is too large, the model selected using the AIC criterion is prone to overfitting, so Schwartz modified it on the basis of the AIC criterion and proposed the Bayesian Criterion (BIC), which changes the weight of the number of unknown parameters in the indicator from 2 to $\ln(N)$, so that the weight of the number of unknown parameters becomes a variable related to the sample capacity. According to the BIC criterion, to determine the end of the model, it is sufficient to choose the p,q order corresponding to the model with the smallest BIC value. The expression is as follows:

$$BIC = -2\ln(L) + N\ln(N) \tag{13}$$

(2) White noise test, also known as the randomness test, is to test whether there is a correlation between the data of a time series, if it is not correlated, on behalf of the sequence is random, then the sequence is white noise sequence.

For the established ARIMA model available $LB(Ljung-Box)$ test statistic for χ^2 test, the original hypothesis and alternative hypothesis are:

$H_0 : \rho_1 = \rho_2 = ... = \rho_m, \forall m \geq 1$

H_1: there exists at least some $\rho_k \neq 0, \forall m \geq 1, k \leq m$

The LB test statistic expression is as follows:

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\rho_k^2}{n-k}\right) \sim \chi^2(m), \forall m > 0 \tag{14}$$

where ρ_k denotes the sample residual autocorrelation function, n denotes the number of samples, and m is $\sqrt{n}$ or $\frac{n}{10}$ by calculating the probability value of the LB statistic, and if it is more than the value at the confidence level, the original hypothesis does not hold true, which indicates that there is no correlation of the residuals between the model predicted values and actual values , the model has adequately extracted the correlation information between the sequences, otherwise, the model is considered to have failed the test.

II. C.ARIMA-LSTM Combined Modeling

II. C. 1) CRITTC weight assignment methods

The CRITTC assignment method is an objective weight assignment method, and its basic idea is to determine the objective weights of the indicators on the basis of contrast and conflict. The contrast strength is expressed in the form of standard deviation, which indicates the size of the gap between the values of different evaluation methods in the same indicator, and the larger the standard deviation, the larger the gap between the values of each; the conflict between the indicators is based on the correlation between the indicators, and the stronger the positive correlation between the two indicators, the smaller the conflict.

The conflictivity of the j th indicator and other indicators is: $\sum_{k=1}^{K}(1-r_{k,j}), j = 1,2,\cdots,J$. Let C_j be the amount of information contained in the j th evaluation indicator, then C_j can be expressed as:

$$C_j = \sigma_j \sum_{k=1}^{K}(1-r_{k,j}), j = 1,2,\cdots,J \tag{15}$$

where σ_j is the standard deviation of the j th indicator, which indicates the contrast of the indicator; $r_{k,j}$ is the correlation coefficient between the evaluation indicators k and j. The larger C_j is, the more information the j th indicator contains, so the objective weight W_j of the j th indicator should be:

where σ_j is the standard deviation of the j th indicator, which indicates the contrast of the indicator; $r_{k,j}$ is the correlation coefficient between the evaluation indicators k and j. The larger C_j is, the more information the j th indicator contains, so the objective weight W_j of the j th indicator should be:

$$W_j = \frac{C_j}{\sum_{j=1}^{J} C_j}, j = 1, 2, \cdots, J \tag{16}$$

II. C. 2) Combined LSTM and ARIMA models

The combined model uses LSTM model prediction as the main form and ARIMA model prediction as the supplementary form for prediction. Firstly, LSTM is used to establish the prediction model of latitude, longitude and altitude; then, ARIMA is used to establish the altitude prediction model alone; finally, CRITTC method is used to fuse the altitude prediction values of LSTM and ARIMA, and the altitude prediction value and latitude/longitude prediction value of LSTM are composed of the final prediction results after the fusion. The prediction process is shown in Figure 1.

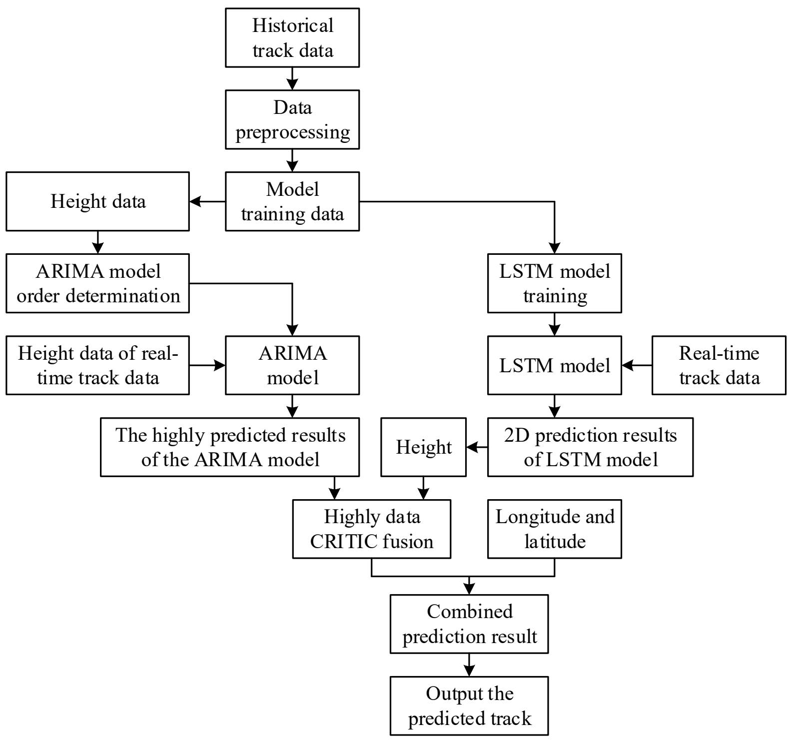

Figure 1: Prediction process of the combined model

III. Bitcoin exchange rate prediction results based on ARIMA-LSTM combination model

III. A. Analysis of single-model Bitcoin price prediction results

Bitcoin price prediction has been a hot topic of concern for many scholars and practitioners, and if we can accurately predict its trend and price, we can seize the first opportunity in this market. In this section, for example analysis, two comparison models, LSTM and ARIMA, and this paper's model, ARIMA-LSTM, are selected to model and predict the price of bitcoin, and finally the optimal model is selected based on the size of the prediction error value. The experimental dataset comes from the closing price of bitcoin in GD province in M website between 2021.9.1-2024.12.31.

III. A. 1) Univariate LSTM Neural Networks

To normalize the data before constructing the LSTM neural network, this paper chooses the min-max normalization method to process the data. The input variable is the closing price, and the output variable is also the closing price. After a number of experiments to determine the number of cycles is 120 times, the 130 samples will be packed into a group for training, that is, the number of samples obtained in one training is 130, and disrupt the training set data.

The model learning rate is 0.001 and is optimized using Adam's algorithm for training with the RMSE indicator as the loss function of the model. After 120 iterations, the loss function metric RMSE decreases quickly and converges, indicating that the univariate LSTM neural network prediction model is reasonable.

After the model is trained, it can be predicted, the test set of 200 days of data into the model for prediction, and compared with the actual data, the univariate LSTM neural network prediction results are shown in Figure 2, you can see that the univariate LSTM neural network model for the bitcoin closing price trend prediction is more accurate, and the gap between predicted value and the actual value is relatively small.

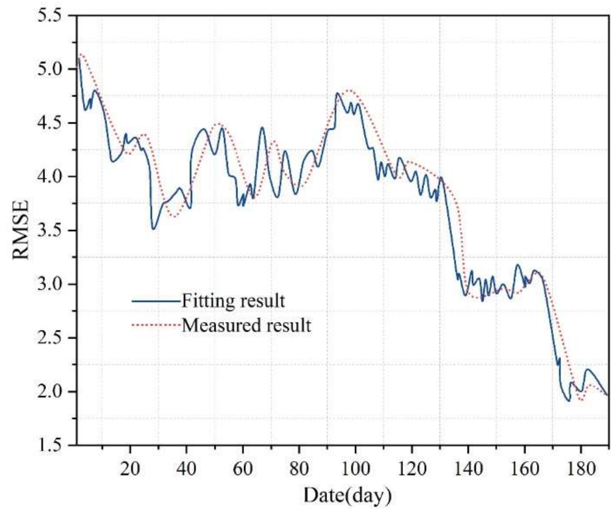

Figure 2: The prediction of the single variable LSTM neural network

III. A. 2) ARIMA model

The first step is to draw a time series of bitcoin's closing prices from September 1, 2021 to December 31, 2024 the bitcoin closing price time series results are shown in Figure 3. Initially, the bitcoin closing price is judged to be a non-stationary sequence.

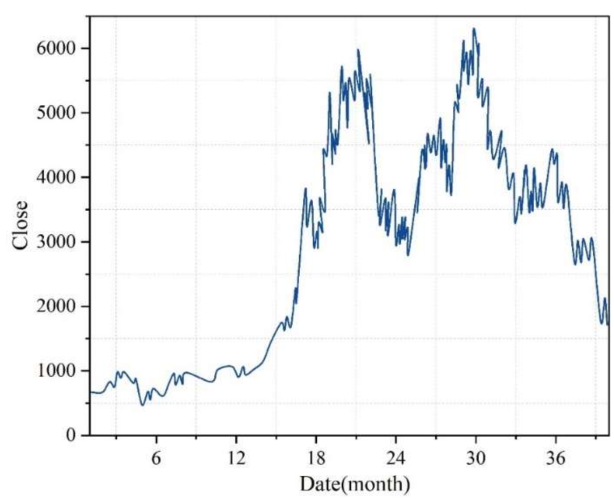

Figure 3: The result of the closing price of bitcoins

The timing diagram of bitcoin's closing price after first-order differencing is shown in Figure 4. It can be seen that its closing price sequence after first-order differencing is basically uniformly distributed on the upper and lower sides of the 0 scale, and it is initially judged that the sequence after first-order differencing is smooth.

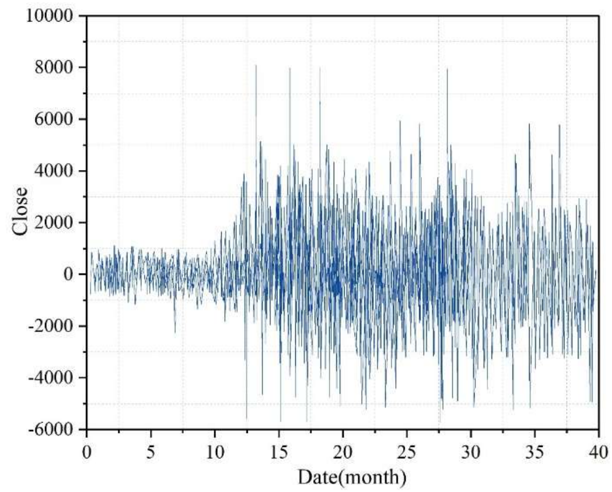

Figure 4: The sequence of the first order difference of bitcoins

After the first-order differencing, the autocorrelation (ACF) and partial correlation (PACF) plots of the sequence plot are within the confidence interval, which is consistent with the smoothness test, at this time the sequence has been smooth. The following to determine the parameter p and parameter q, it can be seen that after the difference of the ACF plot and PACF plot can be seen that p and q in 1 after convergence, so can establish ARIMA model. ARIMA model fitting prediction results are shown in Figure 5. The results show that the model is fitted quite well, and the white noise test is done on the residuals of the fitted model. The residual series does not have autocorrelation and the model passes the test.

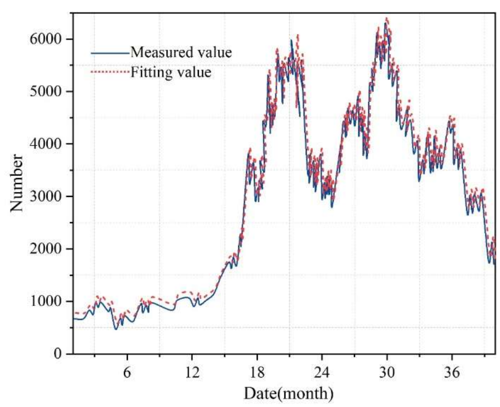

Figure 5: The fitting prediction of the ARIMA model

The ARIMA model prediction results are shown in Figure 6. Therefore, the ARIMA (1, 1, 1) model is reasonable and the fitted this model is used to predict the Bitcoin closing price for 190 days of data. The value of prediction error RMSE is 1763.

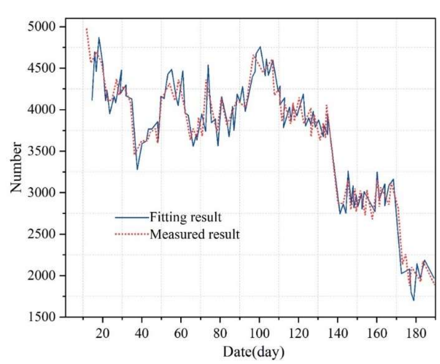

Figure 6: The ARIMA model predicts the results

III. B. Experimental analysis of ARIMA-LSTM

III. B. 1) CRITIC weight settings

The training set was divided into 7200 data as the validation set, and the improved LSTM, ARIMA, and Open were tested, and the weights of the three were determined after the information entropy calculation of CRITIC, which were 0.39489, 0.20522, and 0.39399, respectively.

III. B. 2) ARIMA-LSTM experimental results

ARIMA-LSTM uses July 18, 2021 to February 11, 2023 as the training set to train the model. The prediction results of the combined ARIMA-LSTM model for Close are shown in Fig. 7, where the horizontal axis represents the time and the vertical axis represents the close of the foreign exchange EURUSD Close, the solid line represents the true close, and the dashed line represents the value predicted by ARIMA-LSTM for Close. LSTM for the predicted value of Close. It can be seen that the real value and the predicted value fluctuation is basically the same, there is only a small difference, indicating that ARIMA-LSTM for the real foreign exchange trend of the fitting effect is better.

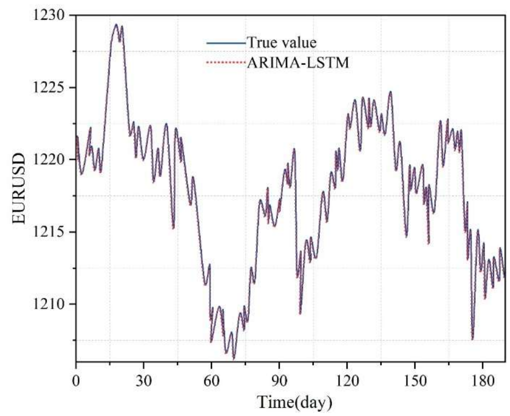

Figure 7: Predictive results of the ARIMA-LSTM model

III. B. 3) Comparative experimental analysis

In order to prove that ARIMA-LSTM is suitable for foreign exchange trend prediction, in the same computer operating environment, the use of 2 groups of comparative experiments, respectively, LSTM, ARIMA, ARIMA-LSTM model, of which ARIMA-LSTM model is the model chosen in this paper.

LSTM predicted value and the real value of the comparison results shown in Figure 8, respectively, for part of the real foreign exchange trend value and LSTM model and ARIMA model prediction results comparison chart, view the chart can be seen that the LSTM model predicted value of the trend and the real value of the difference between the relatively small, and finally the ARIMA model. ARIMA model in the overall trend of fluctuations close to the real value of the trend, the best fitting effect, but part of the interval there are still some gaps.

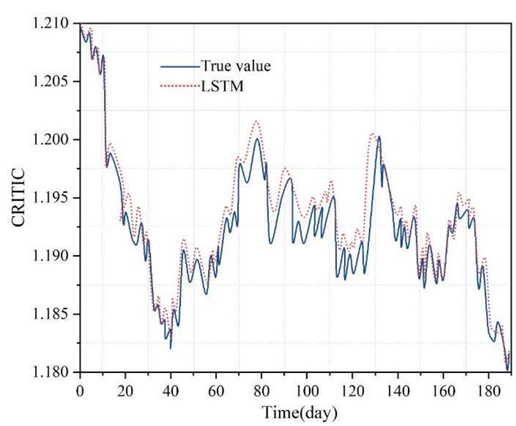

Figure 8: The LSTM prediction is compared to the real value

The results of the comparison between the predicted and real values of ARIMA model and ARIMA-LSTM model are shown in Figure 9 and Figure 10. It can be seen that the ARIMA model and ARIMA-LSTM model are basically in line with the real value of the data trend, and the data difference is small, but the ARIMA model in individual positions of the prediction results are not as good as the ARIMA-LSTM model. Therefore, it can be seen that the prediction performance of ARIMA-LSTM is better than the prediction performance of ARIMA model.

Compared to the other models, the ARIMA-LSTM model predicted the highest fit to the true value, followed by the ARIMA model and the LSTM model, respectively, with the LSTM model having the lowest fit.

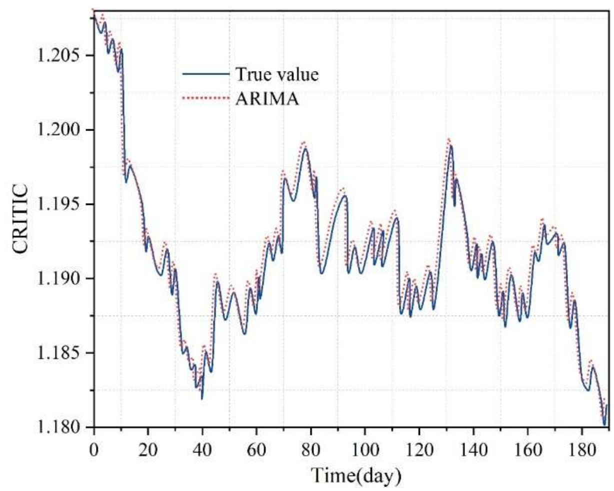

Figure 9: The predictive value of the ARIMA model is compared to the real value

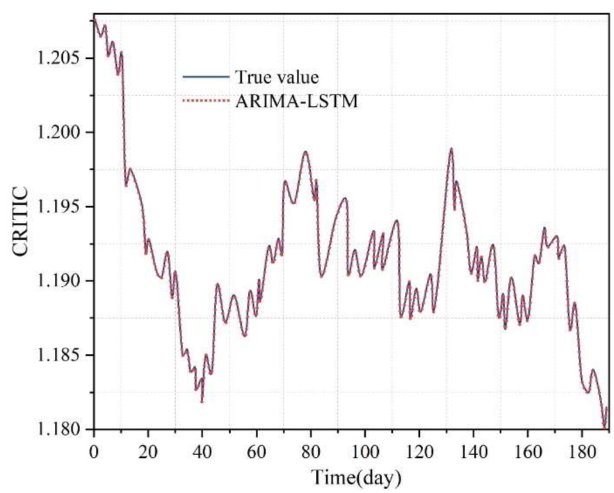

Figure 10: The predictive value of the ARIMA-LSTM is compared to the real value

The results of the error comparison between the predicted and true values of different models are shown in Table 1. In order to evaluate the model fitting effect more intuitively, the experimental results were evaluated using three metrics, which are mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE). From the table, it can be seen that MAE, RMSE and MAPE are minimum for ARIMA-LSTM and maximum for LSTM. The order of MAE, RMSE and MAPE from high to low is LSTM, ARIMA and ARIMA-LSTM, respectively. The MAE, RMSE and MAPE of ARIMA in comparison to the LSTM model have been reduced, indicating that ARIMA also has some advantages in time series. In addition, combining the ARIMA model and the improved LSTM model with each other can improve the prediction accuracy. The MAE, RMSE and MAPE errors of the ARIMA-LSTM model are reduced by 0.0071, 0.004 and 0.0051, respectively, compared with the ARIMA model. It can be seen that the prediction results of combining the ARIMA model and the improved LSTM model can improve the forecast accuracy. It can be seen that the prediction performance of ARIMA-LSTM is better than the other models, no matter for MAE, RMSE or MAPE, so the ARIMA-LSTM model is more superior in this experiment.

Table 1: The comparison of the predicted value and the real value

Model	MAE	RMSE	MAPE

LSTM	0.0195	0.0055	0.0062
ARIMA	0.0073	0.0043	0.0057
ARIMA-LSTM	0.0002	0.0003	0.0006

IV. Conclusion

By constructing the ARIMA-LSTM combined prediction model, this paper has achieved significant results in bitcoin exchange rate prediction. The experimental results fully prove the superior performance of the combined model, and its average absolute error is only 0.0002, which is much lower than that of the LSTM model (0.0195) and the ARIMA model (0.0073), and the prediction accuracy has been greatly improved. Meanwhile, the root mean square error and the average absolute percentage error also reach the excellent levels of 0.0003 and 0.0006, respectively, showing good prediction stability. The application of the CRITIC weight assignment method makes the model adaptive to adjust the contribution of different prediction methods, in which the weights of the improved LSTM, ARIMA, and Open are 0.39489, respectively, 0.20522 and 0.39399, realizing a scientific and reasonable model fusion. Compared with the single ARIMA model, the combined model achieves significant improvements in the three key indicators, and the MAE, RMSE and MAPE are reduced by 0.0071, 0.004 and 0.0051 respectively, which fully reflects the effectiveness of the combination of linear and nonlinear methods. The combined model successfully overcomes the limitations of traditional statistical models in dealing with nonlinear relationships, while avoiding the overfitting problem that may occur in a single deep learning model, and enhances the generalization ability of the model while maintaining the prediction accuracy.

This study provides a new technical path for digital currency price prediction, which is of great practical significance for investors to formulate trading strategies, financial institutions to control risks, and regulators to formulate policies. Future research can further explore the integration of diversified features and dynamic weight adjustment mechanism to cope with the increasingly complex changing features of the digital currency market.

References

[1] Dahdal, A. M. (2023). Monetary History and Digital Currencies: Towards a Conceptual Framework. Available at SSRN.

[2] Baldwin, J. (2018). In digital we trust: Bitcoin discourse, digital currencies, and decentralized network fetishism. Palgrave Communications, 4(1).

[3] Belke, A., & Beretta, E. (2020). From cash to private and public digital currencies: The risk of financial instability and “modern monetary Middle ages”. Economics and Business Letters, 9(3), 189-196.

[4] Panda, S. K., Sathya, A. R., & Das, S. (2023). Bitcoin: Beginning of the cryptocurrency era. In Recent advances in blockchain technology: Real-world applications (pp. 25-58). Cham: Springer International Publishing.

[5] Shu, M., Song, R., & Zhu, W. (2021). The ‘COVID’crash of the 2020 US Stock market. The North American journal of economics and finance, 58, 101497.

[6] Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the COVID-19 bear market. Finance Research Letters, 35, 101607.

[7] Echarte Fernández, M. Á., Náñez Alonso, S. L., Jorge-Vázquez, J., & Reier Forradellas, R. F. (2021). Central banks’ monetary policy in the face of the COVID-19 economic crisis: Monetary stimulus and the emergence of CBDCs. Sustainability, 13(8), 4242.

[8] Beckworth, D. (2020). COVID-19 pandemic, direct cash transfers, and the Federal Reserve. Mercatus Center Research Paper Series, Special Edition Policy Brief (2020).

[9] Lindmayer, R. (2021). The Liquidity Pandemic: A Recent History of the Federal Reserve and Economic Implications of Historically Aggressive Actions during the COVID-19 Pandemic. Syracuse L. Rev., 71, 19.

[10] Cong, L. W., & Mayer, S. (2022). The coming battle of digital currencies. The SC Johnson College of Business Applied Economics and Policy Working Paper Series, 4.

[11] Marthinsen, J. E., & Gordon, S. R. (2022). The price and cost of bitcoin. The Quarterly Review of Economics and Finance, 85, 280-288.

[12] Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. International Journal of Forecasting, 37(1), 28-43.

[13] Feng, W., & Zhang, Z. (2023). Currency exchange rate predictability: The new power of Bitcoin prices. Journal of International Money and Finance, 132, 102811.

[14] Radityo, A., Munajat, Q., & Budi, I. (2017, October). Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods. In 2017 international conference on advanced computer science and information systems (ICACSIS) (pp. 433-438). IEEE.

[15] Zhang, R., & Hao, Y. (2024). Time series prediction based on multi-scale feature extraction. Mathematics, 12(7), 973.

[16] Wu, X., & He, A. (2025). Multimodal information fusion and artificial intelligence approaches for sustainable computing in data centers. Pattern Recognition Letters, 189, 17-22.

[17] T.E. Arijaje,T.V. Omotosho,A.P. Aizebeokhai,S. A. Akinwumi & K. D. Oyeyemi. (2025). Modelling and Prediction of Satellite Signal Path Loss using the ARIMA models at Ku-band in Lagos State, South Western Nigeria. IOP Conference Series: Earth and Environmental Science,1492(1),012042-012042.

[18] Lijian Wei,Sihang Chen,Junqin Lin & Lei Shi. (2025). Enhancing return forecasting using LSTM with agent-based synthetic data. Decision Support Systems,193,114452-114452.