ISSN 2616-5775 Vol. 5, Issue 7: 79-84, DOI: 10.25236/AJCIS.2022.050713

Prediction on the Value Trends of Bitcoin and Gold-on Account of ARMA Time Series Forecasting Model

Ronghuan Li*, Weijie Chen, Wenhao Xu, Chen Li

School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu, China
**Corresponding author*

Abstract: *In this paper, we aimed to build a quantitative investment trading model based on a combination of a multivariate cycle ARMA model and Apriori. We first note that in order to have a sound investment strategy, a forecast for the next trading day needs to be made. To do this, a basic time series forecasting model was first built to predict the value of gold and bitcoin for the next day based on the market volatility of the previous 40 days. The next step is developing a trading strategy model with a stable rate of return and some risk tolerance. At the same time, we developed a fixed stop-loss strategy to protect the strategy's stability and improve the risk resistance performance. Ultimately, using this model, we calculated that on 10 September 2021, we will have a return of $4816941 in Bitcoin and $1129.0503 in gold.*

Keywords: *Quantitative Trading; Trading Strategies; Apriori; ARMA model; time series forecasting model*

1. Introduction

Market traders frequently buy and sell volatile assets with the aim of maximizing their total return. Two such assets are gold and bitcoin. Quantitative trading has a total of two objectives, one is to maximize the utility of the investment and the other is to minimize the value at risk.] In this context, this paper takes gold and bitcoin as the prediction object, and seeks excess returns by improving the time series prediction model to construct a trading strategy suitable for the characteristics of gold and bitcoin. According to the pricing data files LBMA-GOLD.csv and BCHAIN-MKPRU.csv, Bitcoin can be traded daily, but gold is only traded on the days the market is open. Based on this trading schedule, our model needs to calculate the value and demonstrate the development. Our mission is to make use of Apriori and ARMA Model based on Multivariate period to create quantitative transaction strategy model.

2. Establishment and Results of Model

We are asked to find the best daily trading strategy in order to maximize the benefits of investing in the two financial products given in the question, gold and bitcoin.

Model: Time Series Forecasting Model

2.1. Selection of predictive model

To construct forecasting models for gold and bitcoin, we have chosen time series analysis for forecasting purposes. We set the USD(PM) time series for gold to y_t and the value time series for Bitcoin to q_t . After observing and analyzing the data, we discovered that the data has a wide time span, a large volume of data and random fluctuations.

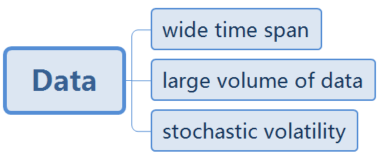

Figure 1: Data Characteristics

ISSN 2616-5775 Vol. 5, Issue 7: 79-84, DOI: 10.25236/AJCIS.2022.050713

Based on the above data characteristics, we choose Auto-regressive Moving Average Model (ARMA), which has a wide range of applicability in time series models.

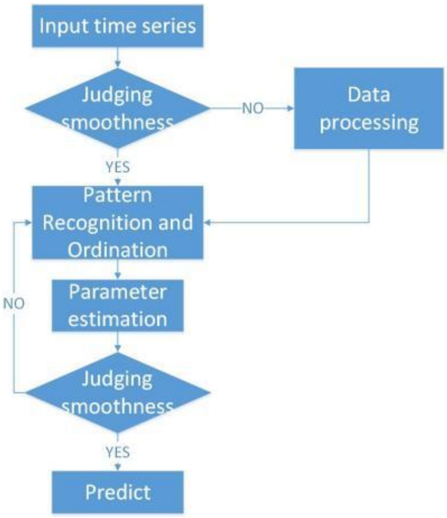

Figure 2: The specific process of ARMA

2.2. ADF testing and data processing

In order to determine whether the time series is smooth, a stability test is required, and we use the ADF test.

Step1: ARMA model requires the series to meet the smoothness, view the ADF test results, according to the analysis of t-value, analyze whether it can significantly reject the hypothesis that the series is not smooth (p<0.05 or 0.01).

Step2: View the data comparison chart before and after the difference, determine whether the smoothness (up and down fluctuations are not large), while the time series bias (auto-correlation analysis), according to the truncated tail Estimate its p, q value.

Step3: ARIMA model requires the model to have pure randomness, that is, the model residuals are white noise, view the model test table, according to the p value of the Q statistic (p value greater than 0.01 for white noise, strict then need to be greater than 0.05).

Table 1: ADF Test Results of Gold

Variables	Differential orders	t	p	AIC	Threshold values		
					1%	5%	10%
1807.7	0	-2.344	0.158	233.862	-3.621	-2.944	-2.610
	1	-4.875	0.000***	229.804	-3.621	-2.944	-2.610
	2	-5.459	0.000***	231.277	-3.639	-2.951	-2.614

Note: *** represents the 1% level of significance.

The results of the golden series test in Table 2, based on field 1807.7, show that at order 0 of the difference, the significance P-value is 0.158, which is not significant at the level of the difference, and the hypothesis cannot be rejected, the series is not a smooth time series.

Table 2: ADF Test Results of Bitcoin

Variables	Differential orders	t	p	AIC	Threshold values		
					1%	5%	10%
35365.2	0	-1.883	0.340	482.797	-3.616	-2.941	-2.609
	1	-6.815	0.000***	471.833	-3.621	-2.944	-2.610
	2	-4.723	0.000***	463.956	-3.661	-2.961	-2.619

Note: *** represents the 1% level of significance.

2.3. Steps of Autoregressive Moving Average Model (ARMA)

Step1: Fitting Model $AR(P)$

ISSN 2616-5775 Vol. 5, Issue 7: 79-84, DOI: 10.25236/AJCIS.2022.050713

We use gold as an example for our analysis and assume that x_t is a smoothed time series of gold after data processing. Assume that the p-order auto-regressive model that we are going to fit is:

$$y_t = \mathrm{x}_t - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} - \cdots - \varphi_p x_{t-p} \tag{1}$$

Where *p* represents the order of the auto-regressive model $AR(P)$.

We then fitted the coefficients of the model $\varphi_1, \varphi_2, \cdots \varphi_p$ using the least squares method. For model $AR(P)$, the following equation is given by:

$$\begin{aligned} x_t &= I_1 X_{t-1} + I_2 X_{t-2} + \cdots + I_p X_{t-p} + y_t \\ &= \sum_{j=1}^{\infty} I_j X_{t-i} + y_t \end{aligned} \tag{2}$$

Comparing equations (1) and (2) it follows that:

$$\varphi_j = I_j \tag{3}$$

Where φ_j is the inverse function of the model $AR(P)$.

Step2: Estimating the initial values of the parameters of model ARMA (2,1)

For model *ARMA (2,1)*,

$$y_t - \theta_1 y_{t-1} = x_t - \varphi_1 x_{t-1} - \varphi_2 x_{t-2} \tag{4}$$

Let *n = 1*, then for model $ARMA(2,1)$, the following expression is available.

$$I_j - \theta_1 I_{j-1} = 0, j > 2 \tag{5}$$

In the above equation, let *j = 3, then* we get $\theta_1 = I_3 / I_2$. Replace I_1, I_2, I_3 in model *ARMA (2,1)* with an approximation of I_1, I_2, I_3 in model *AR (3).*

If $|\theta_1| > 1$, then $\theta_1' = 1/\theta_1$. The relationship satisfies the reversibility condition.

For I_1, I_2, I_3, θ_1, then there is the following relationship.

$$\begin{aligned} \varphi_1 &= I_1 + \theta_1 \\ \varphi_2 &= \theta_2 - \theta_1 I_1 + I_2 \end{aligned} \tag{6}$$

Step3: Testing the accuracy of the predicted values

Using $\varphi_1, \varphi_2, \theta_1$ from step 2 as the initial value, the final value of N and the confidence interval are obtained using least squares and the sum of squared residuals (RSS) is found. Secondly, let *n = n + 1* and fit model $ARMA(2n, 2n-1)$ in the same way.

We test the applicability of the model using the F-test. If the F-test is significant, the model needs to be re-parameterized.

Assuming that A_0 is the sum of squares of the residuals of model $ARMA(n_1, m_1)$ and A_1 is the sum of squares of the residuals of model $ARMA(n_2, m_2)$, then there is the following equation.

$$\begin{gathered} F = \frac{A_1 - A_0}{S} \bigg/ \frac{A_0}{N-r} \sim F(\mathrm{S}, \mathrm{N-r}), \\ (r = n_2 + m_2, s = n_2 + m_2 - (n_1 + m_1)) \end{gathered} \tag{7}$$

ISSN 2616-5775 Vol. 5, Issue 7: 79-84, DOI: 10.25236/AJCIS.2022.050713

Where *N* represents the number of data collection.

Next, test $\varphi_{2n}, \theta_{2n-1}$ .If its value is small and the confidence interval contains zero, fit model $ARMA(2n-1, 2n-2)$ and conversely apply model $ARMA(2n, 2n-1)$. Finally, test model $ARMA(2n-1, 2n-2)$ and model $ARMA(2n, 2n-1)$ using the F-test.

-If F is not significant, discard the known MA parameters, fit model $ARMA(2n-1, \mathrm{m})$ with $m<2n-2$ and test using the F criterion until a suitable model with the smallest parameters is obtained.

-If F is significant, the small MA parameters are discarded, model is fitted and step 2 is repeated.

2.4. Selection of the period of Multivariate Time Series Forecasting

Taking gold as example, we take the time period $T_1 = 1, T_2 = 2, T_3 = 3$ as the data sets D_{T1}, D_{T2}, D_{T3}, respectively, and obtain the forecast value of the next day as y_{T1}, y_{T2}, y_{T3}. The predicted values are calculated from the following expressions.

$$\hat{y}_T = \frac{\sum_{i=1}^{n} y_{T_i}}{n}, n = 3 \tag{8}$$

Since the characteristics of the data set exhibit stochastic volatility, we use 20, 30 and 40 trading days as the base data set for forecasting with a multi-period time series model respectively, and set the following constraints to ensure the quality of the base data set:

$$D_{Ti} < 10, i = 1, 2, 3 \tag{9}$$

When D_{Ti} does not satisfy the above conditions, the data set is ignored.

Model accuracy of gold is measured in terms of relative error β_i .

$$\beta_i = \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \tag{10}$$

Where $\hat{y}_i$ represents the predicted average of the *i-th* periodicity, y_i represents the actual average of the *i-th* periodicity.

The relative error α_i of the model for Bitcoin can be obtained in the same way.

$$\alpha_i = \frac{|\hat{q}_i - q_i|}{q_i} \times 100\% \tag{11}$$

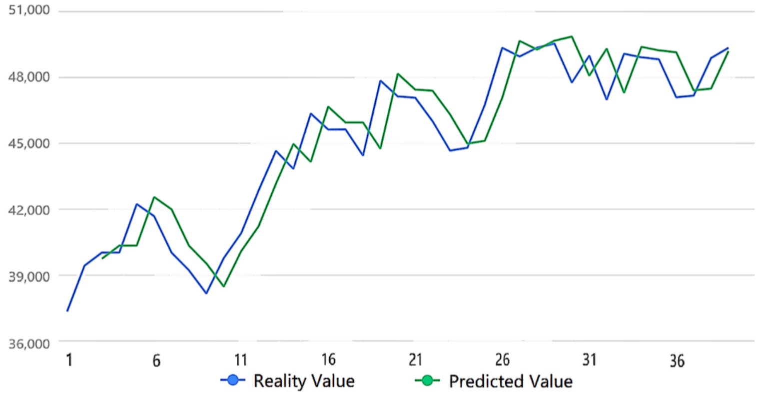

Figure 3: Bitcoin Forecast with a 40-day Cycle

ISSN 2616-5775 Vol. 5, Issue 7: 79-84, DOI: 10.25236/AJCIS.2022.050713

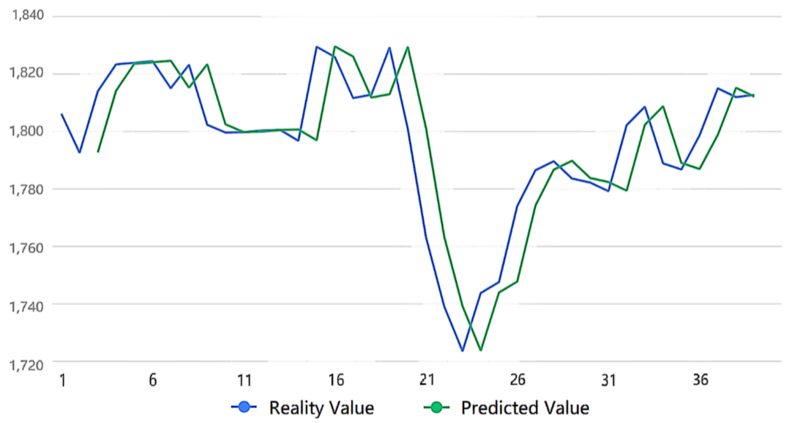

Figure 4: Gold Forecast with a 40-day Cycle

The figure above visualizes that the predicted time series is in good agreement with the actual time series. The model is validated and conclusions are solved for using a known period of historical data as an example.

2.5. Results of Model

Table 3: Relative Errors for Different Forecasting Cycles

<table>
<tr><td></td><td colspan="2">Predicted value</td><td colspan="2">Real value</td><td colspan="2">Relative Error</td></tr>
<tr><td>period</td><td>Bitcoin</td><td>Gold</td><td>Bitcoin</td><td>Gold</td><td>Bitcoin</td><td>Gold</td></tr>
<tr><td>0</td><td>49645.086</td><td>1812.722</td><td rowspan="5">50035.330</td><td rowspan="5">1823.700</td><td rowspan="4"></td><td rowspan="4"></td></tr>
<tr><td>1</td><td>51377.579</td><td>1819.241</td></tr>
<tr><td>2</td><td>48786.42</td><td>1810.714</td></tr>
<tr><td>3</td><td>50495.079</td><td>1812.256</td></tr>
<tr><td>Average Value</td><td>50076.043</td><td>1813.598</td><td>0.08%</td><td>0.05%</td></tr>
</table>

From the table above, the average USD(PM) of the predicted gold on September 3, 2021 is 1823.7 and the average value of the predicted bitcoin on September 4, 2021 is 50035.33. In addition, the relative errors of the gold and bitcoin forecasts are 0.05% and 0.08% respectively. For predicting future trading markets, the forecasting model is highly reliable.

3. Model Evaluation

Strengths

● Our models use Apriori in conjunction with predictive models to provide multi-factor assistance to traders' investment strategies, as well as providing constructive guidance on the quantification of trade amounts, guaranteeing traders stable returns.The ARMA model is very simple and requires only endogenous variables without the need for other exogenous variables. Our trading strategy takes account of capital retractions and has a stop-loss strategy in place as a means of avoiding significant losses and effectively improving the model's resilience to risk.

Weaknesses

● The generalizability of our model has not been validated as we have only tested the model on specific financial products gold and bitcoin, however it may not be applicable to other financial products. The ARMA model requires that the time-series data is stable, or is stable through differing.

Model Extension and Promotion

● Given the similarity in characteristics of other types of financial products, the model can be extended to other financial product portfolios such as foreign exchange, securities and other investment issues. This model can be extended to the problem of formulating materials, i.e. how to formulate raw materials to give the best product performance, and also to the problem of transferring materials to multiple points of sale.

● Finding new measures of asset returns and risks in an uncertain environment. Constructing portfolio selection models that meet different environmental constraints. Proposing various efficient algorithms for optimal portfolio selection, and conducting case studies and analysis.

4. Memo to the Market Traders in the U.S

In the short term the general market environment is stable with little difference in the historical and political context. The market environment can be divided into three phases from selected historical data and the relevant financial background, namely the stable phase (Sept 2016 to Jan 2017), the stable up phase (Sept 2017 to Dec 2017) and the market turbulence phase (Jan 2018 to Apr 2018).

● Stable phase

In a smooth financial market environment, financial products do not rise or fall much and are only traded when the predicted rise or fall is greater than the transaction fee, the investment returns are not large but can be smoothly profitable.

● Stable up phase

In times of steadily rising financial markets, depending on the trading strategies offered, investors can make profits and receive annualized returns in the region of 328.16%, exceeding 90% of investment traders.

● Market turbulence phase

In a period of constant financial market volatility, the use of our trading strategies effectively discourages traders from investing incorrectly due to subjective and emotional influences, maximizing investment returns in a volatile environment.

This quantitative trading decision model has the ability to achieve stable returns and hedge risk for both traditional and non-traditional financial products in different financial environments. The annualized return is consistently above 25%. It is suitable for traders who seek stable returns and do not seek high returns with high risk.

References

[1] Guo Sihan. Research on bitcoin price prediction and trading strategy based on improved recurrent neural network [D]. Shanghai Normal University, 2021.
[2] Li Yi. Research on the application of quantitative trading strategy model [D]. Yunnan University of Finance and Economics, 2014.
[3] An Xiaoxiao. ARMA correlation model and its application [D]. Yanshan University, 2008.
[4] Liu H. T., Guo R. C., Jiang H. Research and improvement of Apriori algorithm for association rule mining[J]. Computer Applications and Software, 2009, 26(01):146-149.
[5] Zhu JL, Fu YZ, Chen YI. Portfolio model and its application [J]. Computer Technology and Development, 2013, 23(12): 168-170+174.
[6] Deng Xue. Research on portfolio optimization model [D]. South China University of Technology, 2010.