

Research on Quantitative Trading Model

Ying Ge^{1,*}, Yue Cao¹, Zi Yang¹, Yuyao Wang¹, Xin Ge²

¹DongWu Business School, Soochow University, Suzhou, Jiangsu, 215021

²Huaiyin Institute of Technology, Huaian, Jiangsu, 223002

*Corresponding author: geying1373145611@126.com

Abstract. With the rapid development of data science, quantitative trading models have become prevalent in financial markets. Based on Apriori algorithm, we analyze the rise characteristics of gold and bitcoin. Based on the principle of financial investment, BiLSTM with attention mechanism and VaR, the transaction decision model was established. We also establish a exponential model to identify how much money should be invested each time. Simultaneously, we apply a modified greedy algorithm to decide our investing decisions and it turns out to be quiet successful. Generally, the trading model established in this paper has good sensitivity to adapt to market changes and has strong risk resistance.

Keywords: Apriori algorithm; BiLSTM; Attention Mechanism; Greedy Algorithm.

1. Introduction

In recent years, quantitative investment has gradually become prevalent. Quantitative investment has been widely used in Europe and American countries for more than 30 years and its investment performance is stable. So far, more and more traders have recognized the importance of quantitative trading models. Quantitative trading refers to the transaction that takes a large number of investment-related data as a sample, establishes appropriate mathematical models and formulas in a quantitative way, writes efficient programs using computer technology, studies and analyze the future returns and risks of financial products, judge various market trends and issue buying and selling instructions to realize investment [1]. Market transactions often buy and sell volatile assets with the goal to maximize their total returns. Quantitative transactions have the advantages of high computational efficiency, strong discipline, high return probability, and rapid analysis and optimization process. Using quantitative trading models can help investors make better decisions.

In order to establish a reasonable model, we first analyzed and processed the pricing data of gold and bitcoin. Then, based on the Apriori algorithm, we analyze the distribution of the consecutive rise or fall times of gold and bitcoin. After fully understanding the characteristics of gold and bitcoin as short-term financial products, we established a trading strategy model. Our trading strategy model mainly consists of two parts, one is the forecasting model based on BiLSTM, the other is the position management model in the form of index based on the short-term financial market investment law of financial market. At last, we use the greedy algorithm, VaR and adjusted the transaction costs to prove the optimal.

2. Prediction model

As shown in Figure 1, assuming that when the given window size is 100, taking 1 as the step, we divide the historical series into multiple price series for prediction In the real prediction process, the window size needs to be set according to the actual situation, and all the data will be used to train the prediction model.

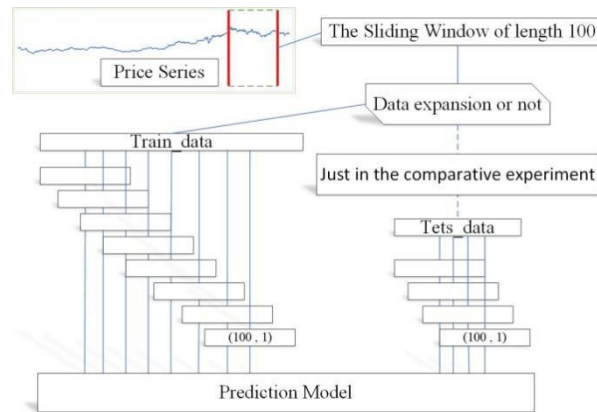


Figure 1. Display of division method of price sreies

Bi-directional long short-term memory (BiLSTM) is a classical model of recurrent neural network [2]. As an outstanding variant of LSTM [3], it can be competent for generation, prediction, classification and other work, and is widely used in natural language processing, computer vision, data mining and other fields.

Figure 2 shows the structure of a LSTM cell, where f_t , i_t , o_t represent the forgetting gate, input gate and output gate respectively. The forgetting gate indicates how much information is retained at the last time, the input gate indicates how much information is retained, and the output gate indicates how much information is output. The calculation process of LSTM can be summarized as forgetting and memorizing the information in c_t and $\tilde{c}_t$ by controlling three gates, so that the key information can be retained in the subsequent calculation.

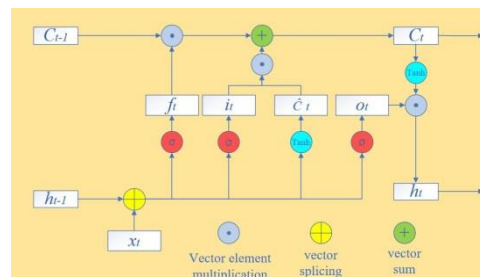


Figure 2. Structure of a LSTM cell

Attention mechanism [4] is used by human beings in processing information, which was born from cognitive science. Its mathematical expression can be expressed as:

$$S(x_i, q) = x_i^T * q \quad (1)$$

$$a_i = softmax(S(x_i, q)) \quad (2)$$

$$att(q, x) = \sum_{i=1}^N a_i x_i \quad (3)$$

where $S(x_i, q)$ indicates the attention scoring mechanism, this paper adopts the commonly used point product model. a_i is called attention distribution. In the mathematical sense, $att(q, x)$ is a kind of weighted average of information, which can be interpreted as the degree of attention of different information in context query q .

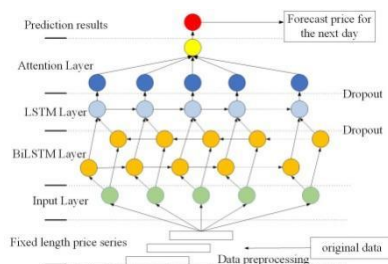


Figure 3. Structure of BiLSTM with attention mechanism

Figure 3 shows the structure of our prediction model. Based on the previous explanation of LSTM model, the prediction model in this paper actually transmits the relationship between prices in different time periods through the input price sequence, and provides back propagation to optimize various parameters in the model by outputting a result and the real value. The measurement method of the model in this paper is that the loss in the neural network is MAE.

In order to prevent over-fitting of the model, which means excellent prediction performance in training data but poor performance in practical application, we use the Dropout layer to ignore 20% of each incoming data. Meanwhile, the training data is randomly arranged in advance to avoid the characteristics observed by the model when the data is continuously entered. In addition, we chose a smaller batch size to avoid falling into sharp minima [5].

RMSprop is used as the activation function in this paper. Both RMSprop and Adam are commonly used activation functions in deep learning. Figure 4 shows the comparison of the price of Bitcoin in the last 200 days predicted by the BiLSTM model with the attention mechanism with the real data.

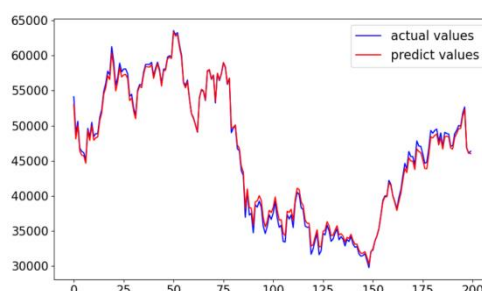


Figure 4. A comparison of the predicted price of bitcoin in the last 200 days with the real price

Compared with LSTM model, gated recurrent unit (GRU) only modifies the gating mechanism [6]. As shown in Figure 5, it adopts a simpler gating scheme to accelerate the prediction speed.

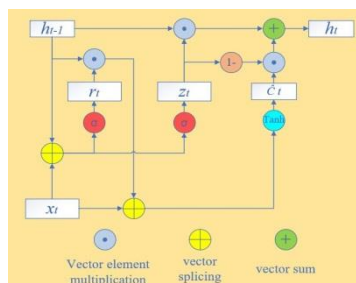


Figure 5. Structure of a GRU cell

Autoregressive comprehensive moving average model (ARIMA) is a method commonly used in econometrics to process time series [7-8]. In this paper, we use Python's auto_arima library to complete price series prediction. Hidden Markov model (HMM) is a probability model about time series in the field of statistics [9]. In this paper, we use the Viterbi algorithm to calculate the state transition matrix based on the given data, and uses the dynamic programming algorithm to take the price sequence as the observation matrix and calculate the optimal path to get the predicted value.

XGBoost (extreme gradient boosting) [10] is an outstanding integrated learning model based on decision tree. In the prediction task, XGBoost takes the square error as the loss function and performs the first derivative and second derivative gradient descent through the second order Taylor expansion. At the same time, it reduces the amount of computation through random forest and uses regular terms to prevent overfitting. Holt-winters method is an effective method to predict trend and seasonal non-stationary series [11].

We select LSTM, GRU, BiLSTM, Holt-Winters, ARIMA, HMM, XGB (XGBoost) for comparison. We divide all price sequences into training sets and test sets in a 7:3 ratio, and test sets are strictly not involved in the training process. In order to ensure the fairness of the experiment, all parameters of the deep learning models are consistent, and the results of other models are debugged to the optimal solution.

Table. 1 shows our experimental results. The experimental results show that our model is obviously better than other domain models. Figure 6 shows the prediction results of each model on the gold and Bitcoin datasets respectively. In short, BiLSTM with attention mechanism has a stronger ability to mine the potential semantics of price sequences, and its prediction effect is better than other models.

Table 1. Comparison results of prediction experiments

Gold	LSTM	GRU	BiLSTM	Holt-Winters	ARIMA	HMM	XGB	Att-BiLSTM
RMSE	29.39	28.06	27.80	33.76	28.01	40.03	34.62	19.01
MAE	21.80	21.12	20.89	24.63	20.50	28.27	25.350	13.98
MAPE	1.1912	1.1534	1.1404	1.3425	1.1177	1.5620	1.3286	0.7626
R Squared	0.8671	0.8702	0.8732	0.8500	0.8782	0.8128	0.8506	0.9392
Bitcoin	LSTM	GRU	BiLSTM	Holt-Winters	ARIMA	HMM	XGB	Att-BiLSTM
RMSE	2138	2005	1901	2534	2886	5609	3623	1479
MAE	1357	1299	1240	1568	1809	2887	2187	906
MAPE	4.2663	4.0406	3.9142	5.0309	5.5572	7.7862	6.2358	2.8775
R Squared	0.9859	0.9869	0.9884	0.9800	0.9738	0.9246	0.9592	0.9929

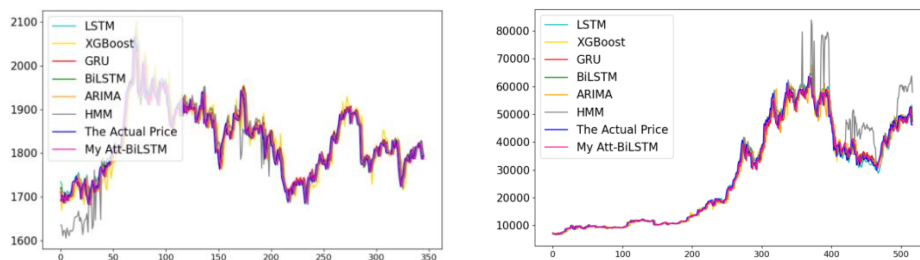


Figure 6. The broken line charts of gold and bitcoin price forecast by each model

3. Position management model

According to the investment law, in order to establish a model, it is necessary to process data on the existing historical data, and count the pricing increase law of gold and bitcoin respectively. The continuous increase and fall law is an important basis for the formulation of scientific and reasonable investment strategies.

Data were performed on historical data given in MATLAB. The rise and fall of financial products can be expressed as:

$$M_i = \frac{p_i - p_{i-1}}{p_i} \quad (4)$$

where M_i is the rise or fall degree, p_i is the current price of financial products, p_{i-1} is the prices of the previous issue of financial products.

Using Apriori algorithm, the number of subsets of 2 consecutive, 3,4,5,6,..., j , m_j represents the number of j consecutive rises, n_j represents and the number of j consecutive drops. The number of all consecutive rises or falls above 90% is u_n . The 90th percentile of all gains is $M_u0.9$, the 10th percentile of all declines was $M_d0.1$.

Ideally, if each drop is the same, when the number of drops is u_n , Just down $M_d0.1$, Each decline will choose to add positions, If the next day the rise is $M_u0.5$ then can make a profit. The commission cost per transaction (sale) is a% of the transaction amount.

The amount of the first added position is p_1 . The second add warehouse added amount is p_2 . We have $p_2 \times (1 - a\%) \times M_u0.5 + p_1 \times (1 - a\%) \left(\frac{M_d0.1}{u_n}\right) = 0$.

The third add warehouse added amount is:

$$p_3 \times (1 - a\%) \times M_u0.5 + p_2 \times (1 - a\%) \times M_u0.5 \left(\frac{M_d0.1}{u_n}\right) + p_1 \times (1 - a\%) \left(\frac{M_d0.1}{u_n}\right)^2 = 0 \quad (5)$$

When the number of positions is n, added the amount is:

$$p_n \times (1 - a\%) \times M_u0.5 + p_{n-1} \times (1 - a\%) \times M_u0.5 \times \left(\frac{M_d0.1}{u_n}\right) + \dots + p_1 \times (1 - a\%) \left(\frac{M_d0.1}{u_n}\right)^{n-1} = 0 \quad (6)$$

Using a linear regression model, the parameters were fitted to obtain the formula:

$$s_i = ae^t \quad (7)$$

If the maximum number of drops is x, the amount of each addition is:

$$y = \frac{PRO}{\sum_{i=1}^{u_n} s_i} s_i \quad (8)$$

According to formula, we can get the gold rise curve and the rise curve of bitcoin, which is shown in Figure 7. Separating the rise and fall, and get the rise curve and fall curve of gold and bitcoin, which is shown in Figure 8.

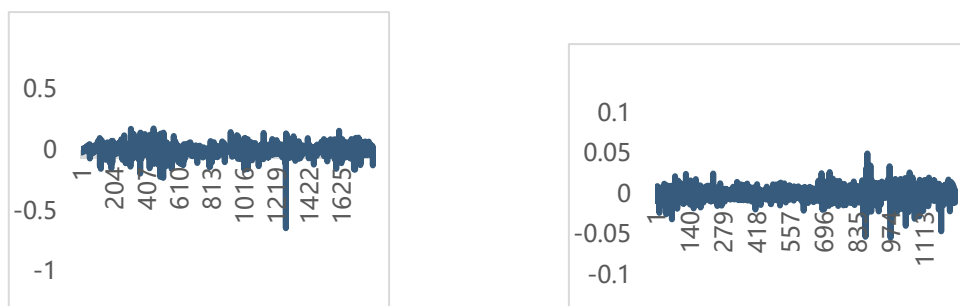


Figure 7. The gold rise curve and the rise curve of bitcoin

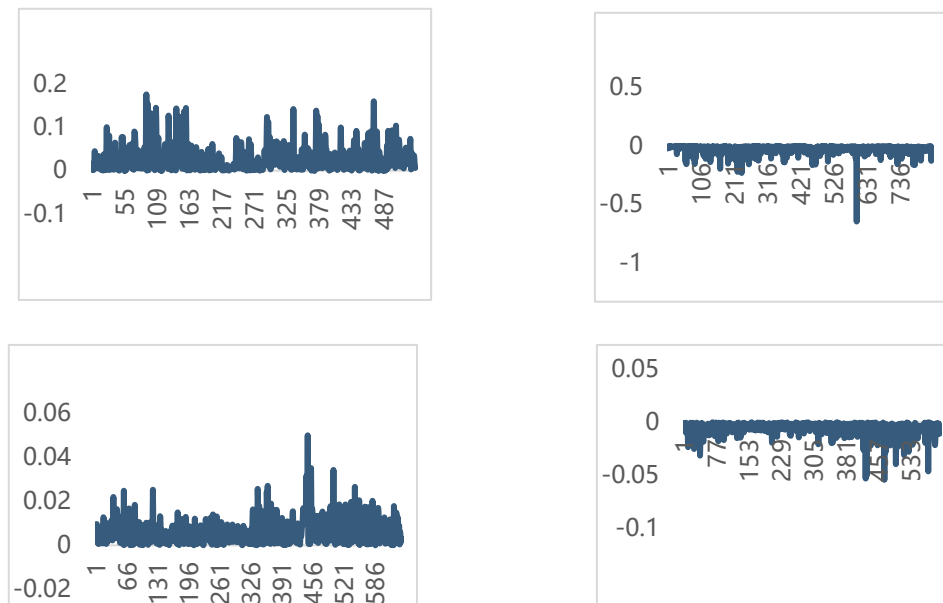


Figure 8. The rise curve and fall curve of gold and bitcoin

4. Quantitative investment model and income analysis

The quantitative investment model consists of a predictive model and a position management model. Predictive model to determine when to trade, and position management model to determine the amount of investment. Invest according to the above strategies and rules to get the returns of gold and Bitcoin. The calculation eventually costs \$646 for investing with \$500 in gold, and investing bitcoin with \$500 for \$215487. The initial \$1000 investment worth on 9/10/202 is \$216133. The results is shown in Figure 9 and Figure 10.

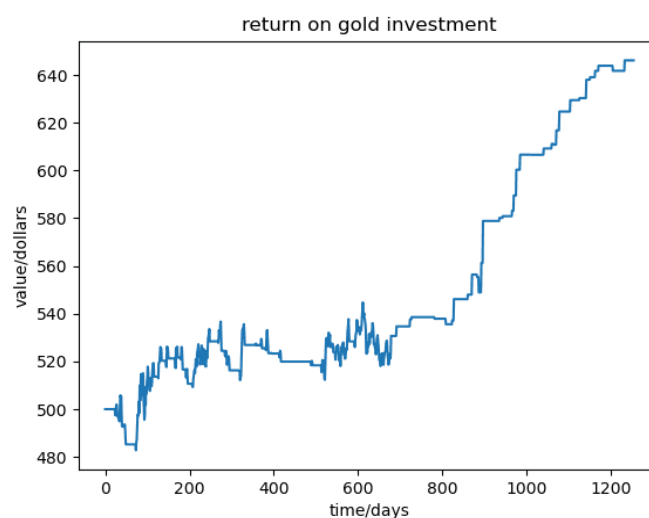


Figure 9. Return on gold investment

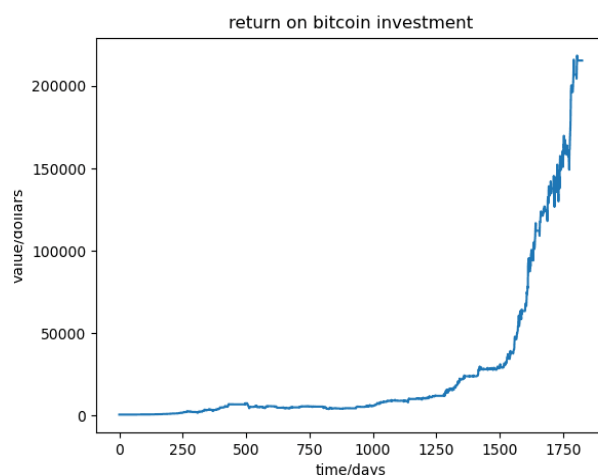


Figure 10. Return on bitcoin investment

From Figure 9 and Figure 10, it can be seen that the return of investment on both gold and bitcoin is quite ideal. The return of gold increases constantly with the moving of time. The return of bitcoin increases sharply in the last 500 days and the profit is really attractive. Further it is showed that the trading strategy constructed by this model strike a balance between short term interests and long term interests. At the same time, the model avoids certain transaction risks.

According to the results predicted by the LSTM neural network, the average predicted increase of gold is 0.43%, and the average predicted value of the increase is 0.41%. The initial transaction fee for this model is as high as 1%, which is much higher than average increasing rate. We set the value of transaction fee to be 0.2%, 0.4%, 0.6%, 0.8%, 1.0%, 1.2% respectively. We put these values into the model and get the sensitivity between transaction fees and ultimate benefit of gold, which is shown in Figure 11.

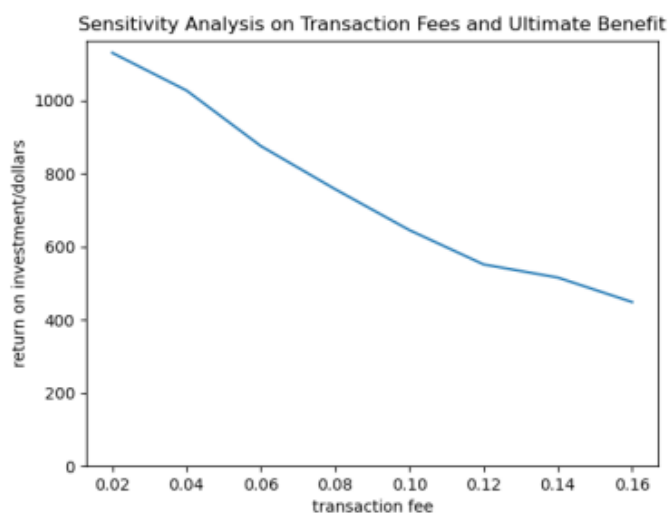


Figure 11. Sensitivity analysis on transaction fees and ultimate benefit

At the same time, the single-day value of bitcoin can increase by more than 20% at the highest, and the minimum decline is also more than 20%. From this, it can be seen that the variation range of the handling fee in Bitcoin transactions can be relatively large. Therefore, we set the value of transaction fee to be 1.5%, 2%, 2.5%, 3%, 3.5% respectively. We put these values into the model and get the the sensitivity between transaction fees and ultimate benefit of bitcoin, which is shown in Figure 12.

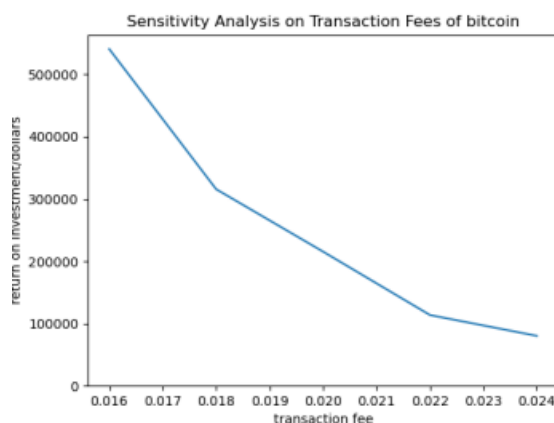


Figure 12. Sensitivity analysis on transaction fees of bitcoin

5. Conclusion

In this paper, multiple models are used for horizontal comparison to produce the optimal model. When considering the decision making, through the forecast of the next three days as trading conditions. We calculate VaR and apply it to the decision model to avoid the potential grey swan event in the market, so as to improve the risk resistance of the model. By comparing with econometrics, statistics, deep learning and other models, we can prove the accuracy of BiLSTM with attention mechanism in predicting the price of gold and bitcoin.

References

- [1] Liu Lijun, Liang Guopeng, Wang Haitao. System model of quantitative trading strategy [J]. Shanghai Business, 2021 (09): 158-159.
- [2] Siami-Namini S, Tavakoli N, Namin A S. The performance of LSTM and BiLSTM in forecasting time series [C]. 2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019: 3285-3292.
- [3] McNally S, Roche J, Caton S. Predicting the price of bitcoin using machine learning [C]. 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP). IEEE, 2018: 339-343.
- [4] Zhang Y, Zhang H, Zhang J, et al. Power grid stability prediction model based on BiLSTM with attention [C]. 2021 International Symposium on Electrical, Electronics and Information Engineering. 2021: 344-349.
- [5] Keskar N S, Nocedal J, Tang P T P, et al. On large-batch training for deep learning: Generalization gap and sharp minima [C]. 5th International Conference on Learning Representations, ICLR 2017. 2017.
- [6] Dai G, Ma C, Xu X. Short-term traffic flow prediction method for urban road sections based on space-time analysis and GRU [J]. IEEE Access, 2019, 7: 143025-143035.
- [7] Poongodi M, Vijayakumar V, Chilamkurti N. Bitcoin price prediction using ARIMA model [J]. International Journal of Internet Technology and Secured Transactions, 2020, 10(4): 396-406.
- [8] Hua Y. Bitcoin price prediction using ARIMA and LSTM [C]. E3S Web of Conferences. EDP Sciences, 2020, 218: 01050.
- [9] Zhang H, Zhang W, Palazoglu A, et al. Prediction of ozone levels using a Hidden Markov Model (HMM) with Gamma distribution [J]. Atmospheric environment, 2012, 62: 64-73.
- [10] Chen T, He T, Benesty M, et al. Xgboost: extreme gradient boosting [J]. R package version 0.4-2, 2015, 1(4): 1-4.
- [11] Liu W, Qin Y, Dong H, et al. Highway passenger traffic volume prediction of cubic exponential smoothing model based on grey system theory [C]. 2nd International Conference on Soft Computing in Information Communication Technology. Atlantis Press. 2014.