

## The Queen of Strategy: The Road to Counterattack With \$1,000

### Summary

Nowdays, financial activities have become an important part of people's lives. In trading, some people end up making profits but some people lose money. The differences between the losers and the profiteers are not only for the choice of asset variety, but also for the trading strategy. In this paper, we focus on the development and evaluation of strategies. For this purpose, we designed two models to develop our strategies: Model I: Volatile Asset Price Forecasting Model; Model II: Strategy Improvement Model.

For Model I, as a forecasting model, the accuracy of its results is relevant to the formulation of our strategy and its returns. It is prudent to predict only one day ahead in a single run. We first use the ARIMA model to forecast the price and find that it can only predict well about the linear part of bitcoin price. Second, we take the LSTM model to forecast, and just the opposite of the ARIMA model, the LSTM model is able to capture the information of the non-linear part of the asset price relatively well; finally, we combine them to form ARIMA-LSTM model, and the final price prediction results almost overlap with the real results, with RMSE and MAPE were 0.0342 and 0.36, respectively.

Second, to develop the strategy, we used the dynamic programming method to split the five-year-long investment strategy problem into smaller problems of daily investment decisions. After programming using MATLAB, we calculated the value of the initial money as \$374,563.25 on 9/10/21 and discounted this value to five years ago to obtain \$360,595.32, which means that the value of \$374,563.25 on 9/10/21 is equivalent to the value of \$360,595.32 on 9/10/16. This results in a return of 37456.33%.

For Model II, we decided to use genetic algorithm to improve the original strategy as we found many opportunities were missing. We created 4 genes, namely trading portfolio, trading frequency, buying strategy, and selling strategy, hoping to improve our strategy from these four dimensions. After MATLAB solving, we learned that the initial capital was valued at \$442,697.25 on 9/10/21 and after discounting it was \$426,188.16, creating a total yield of 44269.73% ,an annualized return of 338.08%.

Again, we demonstrate that our strategy is optimal in two ways, one by changing the strategy itself and the other by observing different market environments. To change the strategy, we reset gene4 in the genetic algorithm. After running the results, we found that the yield after the change was reduced to 273%. To observe the strategy performance in different market environments, we extracted the bull, bear and oscillator markets that existed in five years, using bitcoin trading as an example. Comparing the strategy performance and the market performance during the periods respectively, we again find that the strategy outperforms the market performance by a wide margin.

Finally, we wrote a memorandum to investors detailing our modeling process, introducing our strategy, analyzing the strategy results, and informing about the risks and shortcomings of the strategy.

**Keywords: ARIMA-LSTM Model, Dynamic Programming, Genetic Algorithm**

## Contents

<b>1 Introduction</b> .....	<b>3</b>
1.1 Problem Background .....	3
1.2 Restatement of the Problem .....	3
1.3 Our Work.....	3
<b>2 Assumptions and Justifications</b> .....	<b>4</b>
<b>3 Definitions and Notations</b> .....	<b>4</b>
3.1 Definitions.....	4
3.2 Notations .....	5
<b>4 Model I: Volatile Asset Price Forecasting Model</b> .....	<b>5</b>
4.1 Volatile Asset Price Forecasting Theory .....	5
4.2 Volatile Asset Price Forecasting Methods.....	5
4.3 Volatile Asset Prediction Results .....	8
4.4 Comparative Analysis of Model Forecasting Performance .....	9
<b>5 Dynamic Planning Based Trading Strategy Determination</b> .....	<b>10</b>
5.1 Dynamic Programming to Determine the Trading Strategy Process.....	10
5.2 Dynamic Programming Results .....	11
<b>6 Model II: Strategy Improvement Model</b> .....	<b>11</b>
6.1 Introduction to Genetic Algorithm.....	11
6.2 Process of Genetic Algorithm Optimization Trading Strategy .....	12
6.3 Genetic algorithm to optimize trading strategy results .....	14
<b>7 Effectiveness Evaluation of The Trading Strategy</b> .....	<b>16</b>
7.1 Bull and Bear Market Strategies .....	16
7.2 Stop Loss on Capital .....	17
<b>8 Transaction Cost Sensitivity Analysis</b> .....	<b>18</b>
<b>9 Memorandum to Investor</b> .....	<b>19</b>
<b>10 Model Evaluation</b> .....	<b>20</b>
10.1 Strengths .....	20
10.2 Weaknesses .....	20
<b>11 Conclusion</b> .....	<b>20</b>
<b>References</b> .....	<b>21</b>
<b>Appendices</b> .....	<b>22</b>

# 1 Introduction

## 1.1 Problem Background

In real world of finance, all buying and selling behaviors are based on people's expectations for the future, so how to predict the future price of assets as accurately as possible has become one of the most important research contents in finance. The more accurate the investor's prediction of asset prices, the easier it is to achieve higher expected returns.

Of course, it is rogue to talk about returns without risk. Theoretically, in the market, varieties with higher risk tend to have greater profit. In order to adapt to people's different risk preferences, a wealth of financial products have emerged. Among them, Bitcoin is favored by investors due to its limited total amount, safety, free transaction, wide circulation range and so on. However, due to high price volatility of bitcoin, gold, which is considered as a safe-haven capital, is paired with it to form an investment portfolio.

The "best daily trading strategy" described in the subject must have a specific measurement. It is often said in finance that if investor A is a risk averter and investor B is a risk seeker, we cannot consider investor B's strategy to be superior one even if investor A's return is lower than B's return. Therefore, in this question, our strategy must be based on a given risk preference so that our strategy comparable.

## 1.2 Restatement of the Problem

Considering the background information and restricted conditions identified in the problem statement, we need to solve the following problems:

- If you have \$1000 on 9/10/2016, with known historical prices of the asset, create a model to make an optimal strategy, calculate how much is it worth on 9/10/2021.
- Prove that your strategy is optimal in different situations.
- Perform a sensitivity analysis of strategy results between trading costs and derive how trading costs affect strategy and investment results.
- Writing a letter with traders to summarize your strategy, model, and trading results.

## 1.3 Our Work

The flow chart of our idea is shown in Figure 1.

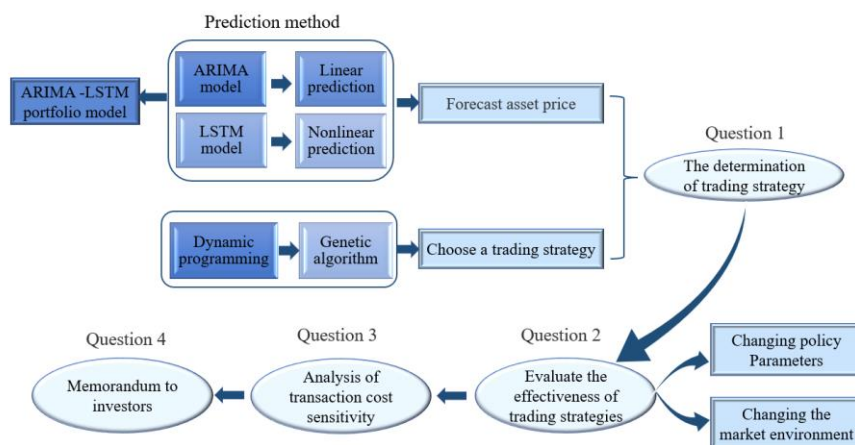


Figure 1: Flow Chart

## 2 Assumptions and Justifications

**Assumption 1:** Investor who will use this strategy is a risk-seeker who willing to take arbitrary risk for the sake of return.

**Justification:** As explained in Section 1.1: strategies with different risks are not comparable. In order to evaluate strategies better, we must assume the risk of strategy, that is, the risk preferences of the investors implementing the strategy. However, since individual risk preferences are difficult to quantify, we make the above assumption that investors are extreme risk seekers, that is, investor will ignore risk to achieve higher expected return.

**Assumption 2:** Once we decide to trade, we can achieve the deal instantly, regardless of the volume we need to achieve.

**Justification:** In the real world, if there are no enough sellers, then even investors trading at current prices may still not be able to close the deal. To simplify the circumstance, we need to make this assumption to rule out this contingency.

**Assumption 3:** No more positions are allowed to be added to an asset while it is held.

**Justification:** Position control is a major challenge in investment strategy formulation and is difficult to implement in code. Since Assumption 1 assumes that the investor is a risk-seeker, it is reasonable for us to have a perception that such investor will tend to invest as much money as possible in profitable assets in order to maximize profits, regardless of the risk involved in doing so.

**Assumption 4:** Assume that the minimum transaction unit for bitcoin is 0.1 bitcoins, and the minimum transaction unit for gold is 1 ounce.

**Justification:** In order to match reality as much as possible, and to simplify the difficulty of programming, we change the minimum transaction unit for bitcoin from 0.01 bitcoins to 0.1 bitcoins, and the minimum transaction unit for gold to be consistent with reality.

## 3 Definitions and Notations

### 3.1 Definitions

**Bull Market:** A market with a long-term upward price trend. The overall trend is upward, with some declines, but one wave higher than the other.

**Bear Market:** A market that has a long-term downward price trend. The overall trend is downward, although there are rallies, but one wave is lower than the other.

**Oscillating Market:** A market situation in which share prices fluctuate and the future of the market is uncertain. It is characterized by an increase in short term investment, unstable market popularity and large ups and downs in prices.

**Position control:** It is your control over the size of your position, and your control over opening, adding, reducing and cutting positions.

**Risk-seeker:** People who prefer to receive the expected income from a risk rather than the expected value of the risk. For a risk taker, the utility of the expectation is greater than the expected utility of the risk itself.

**time value of money:** A certain amount of money currently held has a higher value

than an equivalent amount of money acquired in the future.

### 3.2 Notations

The key mathematical notations used in this paper are listed in Table 1.

**Table 1: Notations used in this paper**

Symbol	Description
$Y_T$	Asset price in time T
$\hat{Y}_T$	Predicted asset price at time T
$p$	The autoregressive term
$d$	The number of differences when the time series is stationary
$q$	The number of moving average
$\mu$	The constant term
$\gamma_i$	The autocorrelation coefficient
$\varepsilon_t$	The error term
$\theta_i$	The coefficient of error term
$CF_{real_d}$	real value by dynamic programming
$CF_{nominal_d}$	nominal value by dynamic programming
$r_f$	risk-free yield

## 4 Model I: Volatile Asset Price Forecasting Model

### 4.1 Volatile Asset Price Forecasting Theory

According to the question, we need to develop a strategy based on price movements to maximize returns by buying in a lower price and selling in a higher price. Therefore, we need to forecast the daily price in the future first.

In the financial market, different asset prices have their own fluctuation trends and patterns. In order to obtain higher returns, people often build reasonable time series models to predict the future development of volatile assets. Price series forecasting is something that based on historical data, using some scientific method to estimate the price of assets in a certain period in the future. Knowing the time series  $\{Y_1, Y_2, Y_3, \dots, Y_T\}$ , find the asset price  $Y_{T+1}, Y_{T+2}, Y_{T+3}, \dots, Y_{T+m}$ . The formula is shown in equation 1.

$$\hat{Y}_{T+1}, \hat{Y}_{T+2}, \hat{Y}_{T+3}, \dots, \hat{Y}_{T+m} = f(Y_1, Y_2, Y_3, \dots, Y_T) \quad (1)$$

### 4.2 Volatile Asset Price Forecasting Methods

The forecasting methods for volatile assets can be divided into linear forecasting models and nonlinear forecasting. Financial markets contain many uncertainties and can be influenced by economic, political, social factors and so on. The changes caused by these factors have a strong disorderly nature, so it is difficult to say exactly whether it is a purely linear or nonlinear system. Therefore, we need to build a model that contains both linear and nonlinear features.

Traditional time series models can extract linear features, while neural network models

have strong mapping properties for nonlinearity, so our study combines the linear time series forecasting algorithm ARIMA and the nonlinear forecasting algorithm LSTM neural network together for volatile asset price forecasting. the specific structure of ARIMA-LSTM is shown in Figure 2.

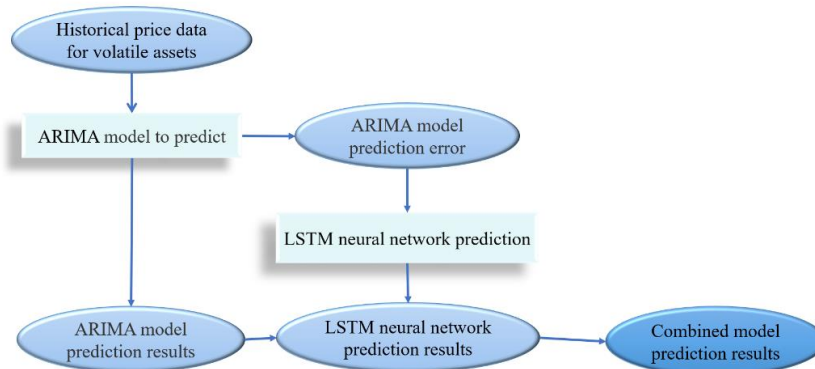


Figure 2: Model ARIMA-LSTM Overview

### 4.2.1 ARIMA Linear Prediction Model

The ARIMA  $(p, d, q)$  model is known as the Autoregressive Integrated Moving Average Model, where  $p$  is the autoregressive term;  $d$  is the number of differences when the time series is stationary;  $q$  is the number of moving average items. This model is a combination of autoregressive (AR) and moving average (MA), which can transform a non-stationary time series into a stationary time series, and then regress the lagged values of the dependent variable, the present and lagged values of the random error term to the model established. The formula is given in equation 2.

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \tag{2}$$

Where  $y_t$  is the current value;  $\mu$  is the constant term;  $\gamma_i$  is the autocorrelation coefficient;  $\varepsilon_t$  is the error term;  $\theta_i$  is the coefficient of error term. The ARIMA  $(p, d, q)$  model process for forecasting asset price is shown in Figure 3.

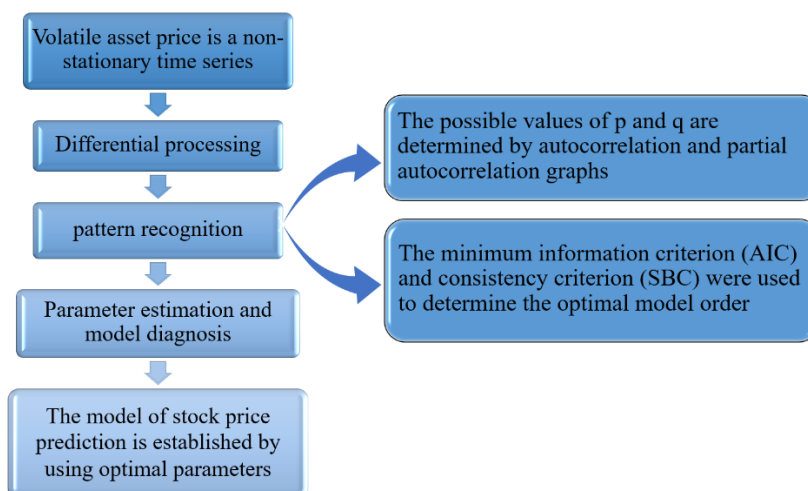
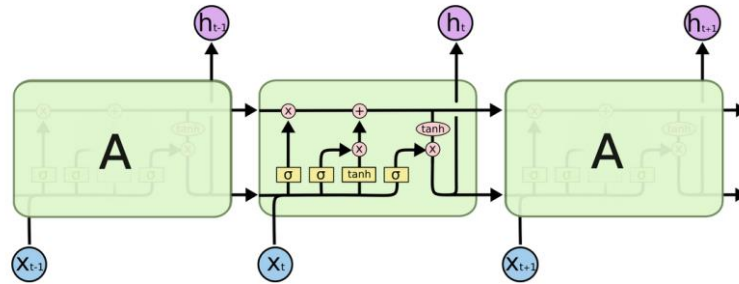


Figure 3: ARIMA Forecast Asset Price Flow Chart

### 4.2.2 LSTM Neural Network Nonlinear Prediction Model

Long Short Term Memory Network (LSTM), a modified recurrent neural network, can handle the problem of long-range dependencies.

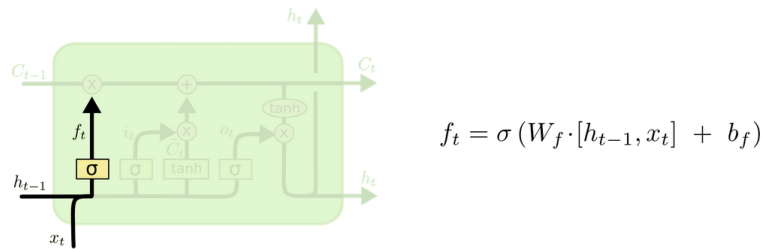


**Figure 4:** LSTM Working Mechanism Flow Chart

The key to the LSTM is the rectangular box in the second elliptical rectangle in Figure 4, which is called the memory block and contains three main gates (forget gate, input gate, output gate) and a memory cell. The horizontal line at the top of the box is called cell state, which is like a conveyor belt that controls the transfer of information to the next moment. The two tanh layers in the diagram above correspond to the input and output of the cell.

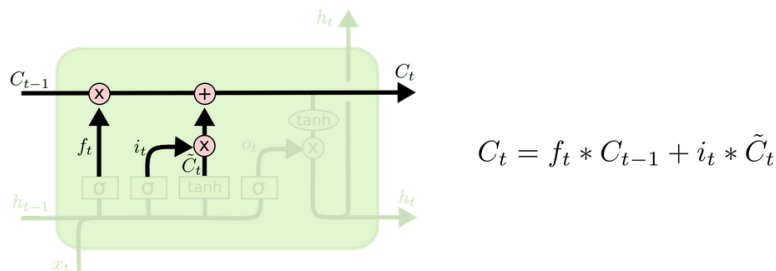
Looking at Figure 4 we can see that the work of LSTM can be divided into 3 main steps.

**Step 1:** Decide what information can pass through the cell state. This decision is controlled by the "forget gate" layer through sigmoid, it will pass or partially pass according to the output of the previous moment. See Figure 5 for details:



**Figure 5:** Step 1 of LSTM

**Step 2:** Generate the new information we need to update. This step consists of two parts, the first is an "input gate" layer that determines which values to update by sigmoid, and the second is a tanh layer that generates new candidates and adds them to the previous candidates to obtain the final candidate values for this part. The two steps are combined to discard the unwanted information and add the new information.



**Figure 6:** Step 2 of LSTM

**Step 3:** Decide the output of the model. The first step is to get an initial output from the

sigmoid layer, then use tanh to scale the value to between -1 and 1, and then multiply the output with the sigmoid pair by pair to get the final output of the model.

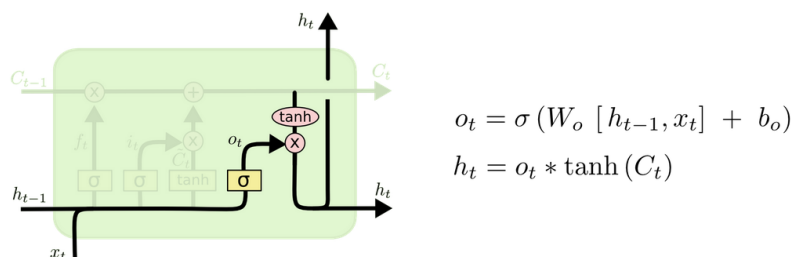


Figure 7: Step 3 of LSTM

### 4.3 Volatile Asset Prediction Results

Taking the asset price prediction of Bitcoin as an example, the ARIMA model prediction and the LSTM neural network model prediction of Bitcoin’s price are performed respectively. The results are obtained as shown below.

#### 4.3.1 ARIMA Model Prediction Results

For this problem, we use the one-step prediction method to predict the asset price, i.e., the first i-1 data are used as the training set when predicting the price of bitcoin at the ith moment, while the ith sample is added to the training set when predicting the i+1th sample.

The prediction results of the ARIMA model are shown in Figure 8. From Figure 8, we can see that the ARIMA model is not very accurate in predicting the price of Bitcoin, which is especially obvious in the upward trend of Bitcoin price. However, the ARIMA model is still able to capture the trend of bitcoin price changes well, which means that it can predict well about the linear part of bitcoin price.

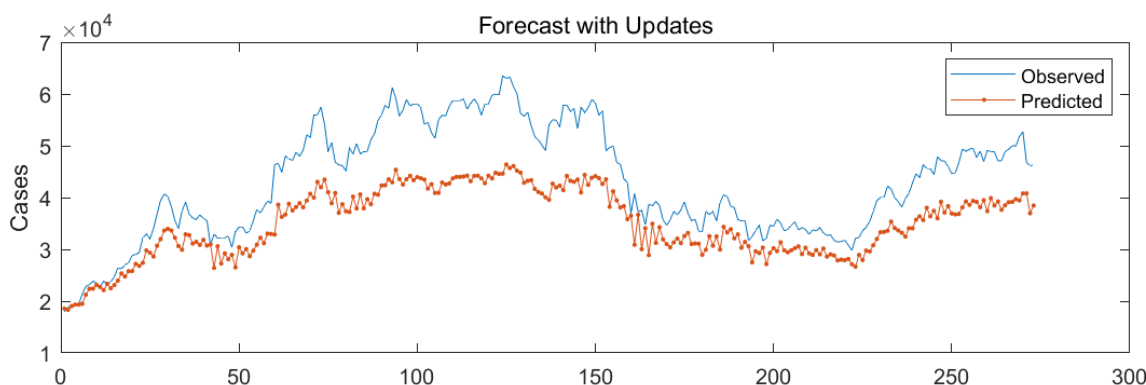


Figure 8: ARIMA Model - Bitcoin Price Forecast Chart

#### 4.3.2 LSTM Model Prediction Results

The results obtained by using LSTM neural network for bitcoin price prediction are shown in Figure 9. It can be found that the LSTM neural network model has improved the predict accuracy of the bitcoin price compared to the ARIMA model, and it can capture the fluctuation of the bitcoin price well. Therefore, we can conclude that the LSTM model is able to capture the information of the non-linear part of the asset price relatively well, and we can also find that with the increase of the training set its prediction accuracy will get higher and higher over time.



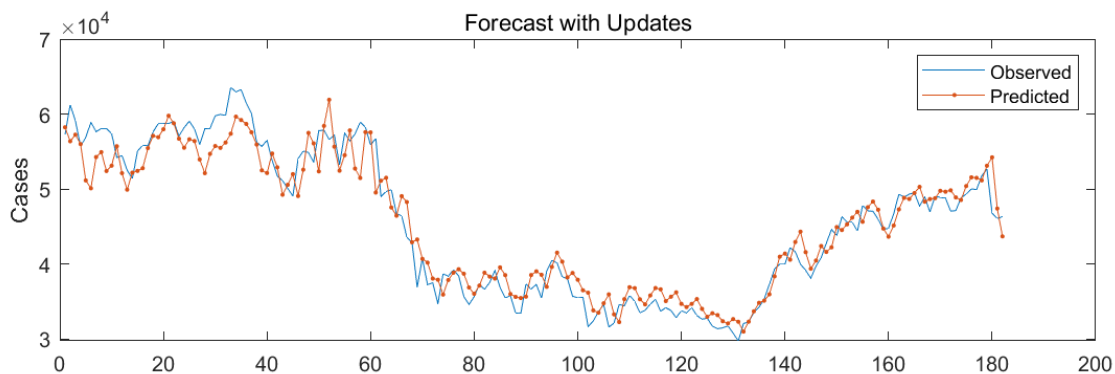


Figure 9: LSTM Model - Bitcoin Price Forecast Chart

### 4.3.3 ARIMA-LSTM Model Prediction Results

To further improve the prediction accuracy, we adopt a combined linear and nonlinear model for prediction. For this purpose, firstly, we based on the ARIMA prediction results and the actual bitcoin price worth to the residual sequence of bitcoin price, which is used as expected output of the LSTM neural network; secondly, we phase space reconstruct the original data with a delay time of 1, and finally determine the optimal price number as 7; thirdly, use the data after reconstructing with the optimal order as the LSTM input; fourthly, input the training set to the LSTM neural network for learning modeling and predicting the residual series test set to obtain the ARIMA residual series prediction value; finally, the ARIMA and LSTM neural network model prediction results are summed to obtain the final prediction result of stock price. The prediction results are shown in Figure 10.



Figure 10: ARIMA-LSTM Model - Bitcoin Price Forecast Results

Accroding to Figure 10, it is easy to find that our prediction results are already very accurate relative to the real values and can be used as the basis for investment decisions.

### 4.4 Comparative Analysis of Model Forecasting Performance

In order to verify the superiority of the ARIMA-LSTM-based asset price forecasting model, we use Root Mean Squared Error (RMSE) and Mean Absolute Percent Error (MAPE) as the model performance evaluation indexes. The evaluation results are obtained in Table 1.

Table 1: Model performance evaluation table

Model	RMSE	MAPE
ARIMA	0.7034	7.21
LSTM	0.6183	6.25
ARIMA-LSTM	0.0342	0.36

By observing the Table 1, it is not difficult to find that: the RMSE and MAPE of ARIMA-LSTM model are very close to 0. Its prediction accuracy is much higher than the prediction accuracy of single ARIMA or LSTM neural network models, and the prediction error is greatly reduced. This is because ARIMA-LSTM combines the advantages of ARIMA model and LSTM model to portray the change pattern of asset price more comprehensively.

Therefore, in the part of investment strategy, we will make investment decisions based on the predict results of the ARIMA-LSTM model.

## 5 Dynamic Programming Based Trading Strategy Determination

### 5.1 Dynamic Programming to Determine the Trading Strategy Process

To obtain the optimal trading strategy, we adopt a dynamic programming algorithm based on the predicted values in the previous section. By splitting the 5-year trading process and defining the relationship between trading states and states, we can solve the problem to in a recursive (or partitioned) manner.

#### 5.1.1 Splitting The Problem

For these 5 years of trading forecasts, we divide the five years in terms of days to get each daily return, for each trading day, we have a triplet like  $[C, G, B]$ , treat each day as a stage, and a daily result can be obtained, so that the problem can be solved by recursion or recursion.

#### 5.1.2 Determining The Problem State

We choose to represent the problem split in the previous section by quantifying it.

Since we know the price up to that day (when we make the decision) and the predicted price in 2-3 days, we can make a daily judgment based on the return. With this setup, we end up transforming the big problem into a small problem.

#### 5.1.3 Dynamic Planning Flow Chart

Dynamic planning flow chart is as below:

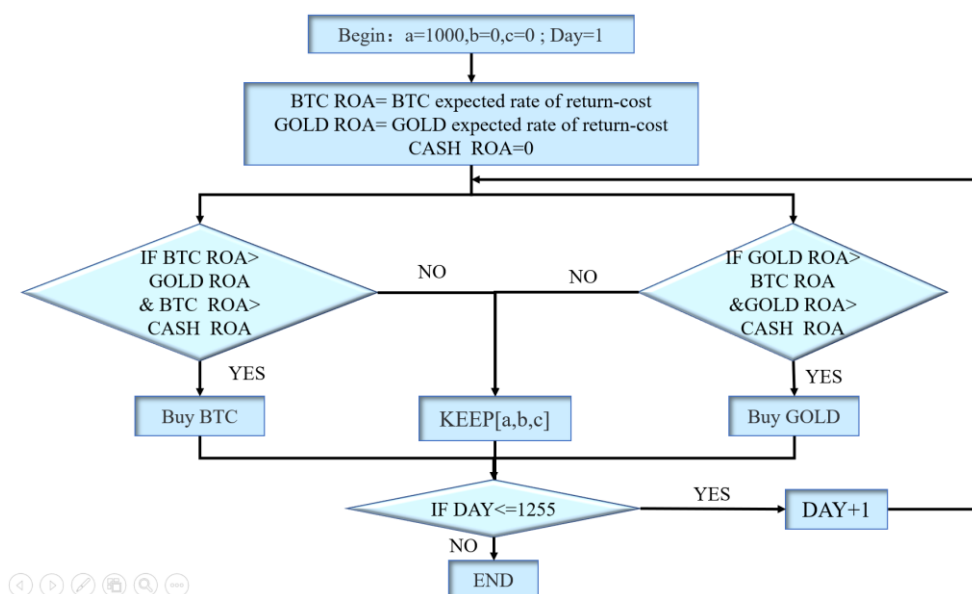


Figure 11: Dynamic Programming Flow Chart

## 5.2 Dynamic Programming Results

We obtained Figure 12 of the change in asset value of \$1,000 from 9/10/16 to 9/10/21 through dynamic programming and obtained a final estimated asset value of \$374,563.25 on 9/10/21.

But since money has time value, we need to discount our terminal value. Since it is common in academia to treat the U.S. Treasury bond yield as the risk-free yield, we consider the risk-free yield as the yield on the U.S. five-year Treasury bond on 9/10/16. The discounting formula is shown in equation 3.

$$CF_{real_d} = \frac{CF_{nominal_d}}{(1 + r_f)^n} \tag{3}$$

Where  $CF_{real_d}$  is the real value by dynamic programming,  $CF_{nominal_d}$  is the nominal value by dynamic programming,  $r_f$  is the risk-free yield.

After calculation, we get that \$374563.25 on 9/10/21 is worth \$360,595.32 on 9/10/16. The formula for calculating the rate of return is as Equation 4.

$$Yeild = \frac{(selling\ price - buying\ price)}{buying\ price} \tag{4}$$

Finally, we find that our total return is 3745.63%, which is an excellent result.

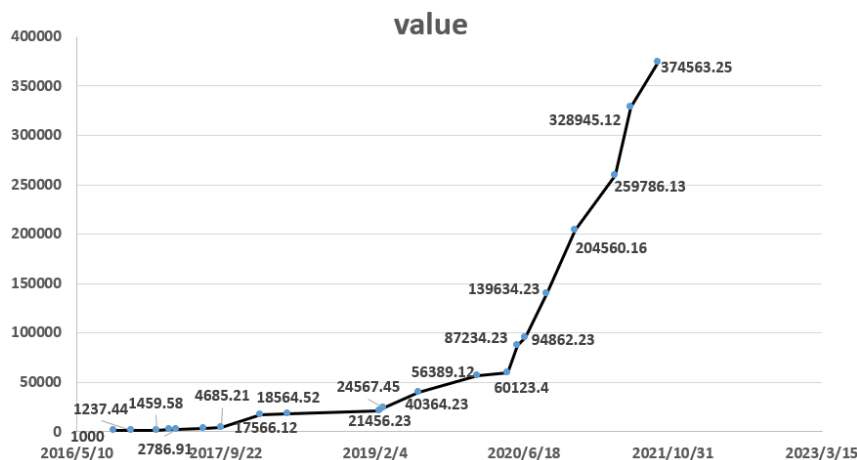


Figure 12: Asset Value by Dynamic Programming (2016/9/10-2021/9/10)

## 6 Model II: Strategy Improvement Model

### 6.1 Introduction to Genetic Algorithm

Genetic algorithm is a randomized search algorithm that draws on the natural selection and natural inheritance mechanisms in biology. It simulates the reproduction, crossover and genetic mutation phenomena occurring in the process of selection and inheritance. In each iteration, a set of better solutions is retained, and better individuals are selected from the group of better solutions according to some index; then use genetic operators to combine these indi-

viduals, so that it is capable to produce a better group of solutions; finally, the process will repeat until results converge to the optimal solution.

The operation flow of the genetic algorithm is shown in the figure 13.

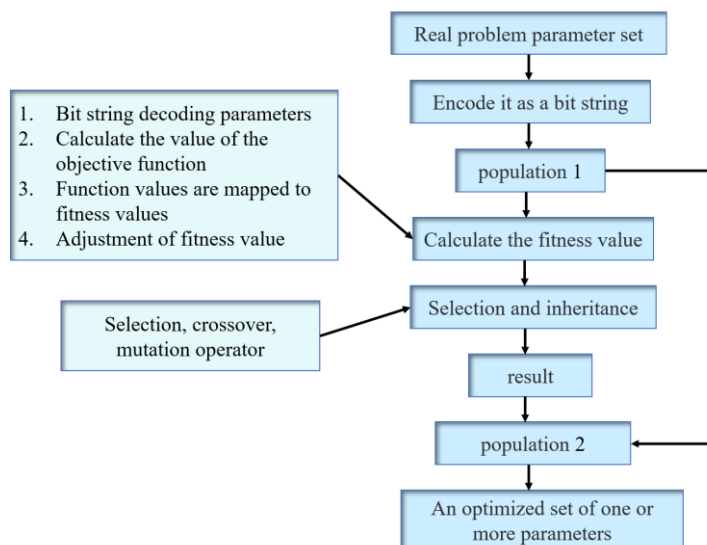


Figure 13: The Operation Flow of The Genetic Algorithm

## 6.2 Process of Genetic Algorithm Optimization Trading Strategy

### 6.2.1 Encoding

- **Gene1:** Trading Portfolio Information

According to the question, we know that the trader has an account consisting of cash, gold and bitcoin as  $[C, G, B]$ , but we cannot determine what trading strategy we need to use. So, we set the trading strategy as a non-deterministic variable in our trading system, i.e., Gene1. This variable contains information about which asset will enter our strategy. In this paper, we choose the method shown in the table below to encode it.

Table 2: Gene1

C	G	B	Gene1
0	0	0	000
0	0	1	001
0	1	0	010
1	0	0	100
0	1	1	011
1	1	0	110
1	0	1	101
1	1	1	111

In the table 2, the number "0" means that the strategy will not trade on this asset; the number "1" means that the strategy will trade on this asset. For example, if Gene1=101, it means that we will trade on cash and bitcoin. It is easy to see that this table covers all possible combinations.

- **Gene2:** Trading Frequency Information

In this article, we create another gene (Gene2) to determine the transaction frequency. The variable Gene2 is used to store our transaction frequency information. Like Gene1, we

use the following rules for coding.

**Table 3: Gene2**

Frequency	1 day	3 days	5 days	7 days	15 days	30 days	60 days
Gene2	000	001	010	100	011	110	111

- **Gene3: Buying Strategy**

There are several strategies for entering the market based on price trends: breakouts of large patterns, breakouts of large averages, breakouts of trends and so on.

For example, there are two types of new highs or new lows: one is a new high or new low reached by crossing a historical resistance or support level; the other is a new high or a new historical low that has been reached by crossing a historical high or low in the price run. Once the above breakout is reached, the capital under the corresponding strategy will choose to buy.

Due to the space limitations, we will not describe all the strategies in detail one by one. Depending on the buying strategy, we can code according strategies in order. Rules are displayed as table 4.

**Table 4: Gene3**

Strategies for breakouts	Large Patterns	Large averages	Trends	Consolidation	New Highs or New Lows	Small Cycle Signal
Gene3	000	001	010	011	100	101

- **Gene4: Selling Strategy**

For the selling strategy, we use fixed percentages of loss for the sell operation. We set These stops: 1%, 2%, 3%, 5%, 8%, 10%, as soon as the stop loss is hit, we will sell the asset.

We have the following table for the Gene4 code:

**Table 5: Gene4**

Stop Loss Percentage	1%	2%	3%	5%	8%	10%
Gene4	000	001	010	100	011	110

These 4 genes can form a chromosome that contains all the choices of our trading strategy, which contains choices for the dimensions of asset, trade frequency, SMA type, calculation period, trailing stop percentage, etc. For our design, the algorithm will first randomly assign "0" and "1" to each position in the chromosome and then repeat the iterations to create the first generation of chromosomes, in this particular case, each chromosome is a strategy.

### 6.2.2 Adaptation Function

The fitness function value is a metric used by genetic algorithms to evaluate the quality of the results. The larger the value of the fitness function, the better the quality of the result. Although there are many methods to evaluate the goodness of trading strategy selection, in this paper only the yield is used as the fitness evaluation function since we only know the asset price price in this problem.

### 6.2.3 Genetic Operators

**Step 1: Selection.** The genetic algorithm uses a selection algorithm to choose individuals with high fitness to be inherited into the next generation population. The roulette wheel selection method is used in the simple genetic algorithm (SGA). Also, backtesting algorithm

is used to perform selection with reading strings, which are then translated into the corresponding strategies and tested. After testing, the algorithm will give the return of each strategy.

**Step 2: Mating.** Its main function is to pair the winners selected by the selection algorithm, which becomes the basis for the next step of reproduction.

**Step 3: Crossover.** It is the process of exchanging some of the genes of two paired chromosomes with each other in some way based on the crossover probability, so that we can form two new individuals.

**Step 4: Mutation.** There are three possible forms of variation in this problem: genetic variation, strategy swapping and strategy variation.

### 6.3 Genetic algorithm to optimize trading strategy results

After continuous genetic inheritance, selection, crossover and mutation, we get a fitness value (rate of return), which is the optimal solution after genetic algorithm optimization. In this problem, this result is shown as follows.

#### 6.3.1 Rate of Return

We plot the returns of each transaction conducted in the five years according to the asset type in Tables 2 and 3, respectively. After observing the tables, we found that Bitcoin has a return of 259.406% in January 2018 and a return of 371.84% in April 2021; on the contrary, the return of gold is more stable and its return does not fluctuate much.

**Table 6:** Rate of Return on BTC & Gold

Date	Rate of Return on BTC	Date	Rate of Return on Gold
2017-1-10	44.77%		
2017-3-17	29.38%		
2017-9-07	120.17%		
2017-9-13	70.51%		
2018-1-15	259.05%	2018-4-19	0.002919
2018-7-31	0.95%	2019-2-27	0.106942
2019-7-16	174.81%		
2019-8-14	5.67%	2020-1-09	0.050922
2020-2-26	0.032783	2020-4-24	0.13822
2020-6-24	0.076138		
2020-9-03	0.193441		
2021-4-21	3.718426	2021-6-11	0.050733
2021-9-10	0.156693		

In the second and third columns of the table, there are some cells with no data in them because there were no transactions made for gold during the period.

#### 6.3.2 Buy-Sell Time and Price

Based on the results of the strategy implementation after the genetic algorithm optimization, the buy and sell positions of bitcoin and gold in 5 years are plotted in the figure. The red points are the buy points and the green points are the sell points.

As can be seen from the Figure, our strategy makes a corresponding buy after the price is about to rise or has risen for some time, while the sell points tend to be a small distance

away from the highest point of volatility. This shows that our strategy is well adapted to the market environment.



Figure 13: Bitcoin Buy-sell Point Chart

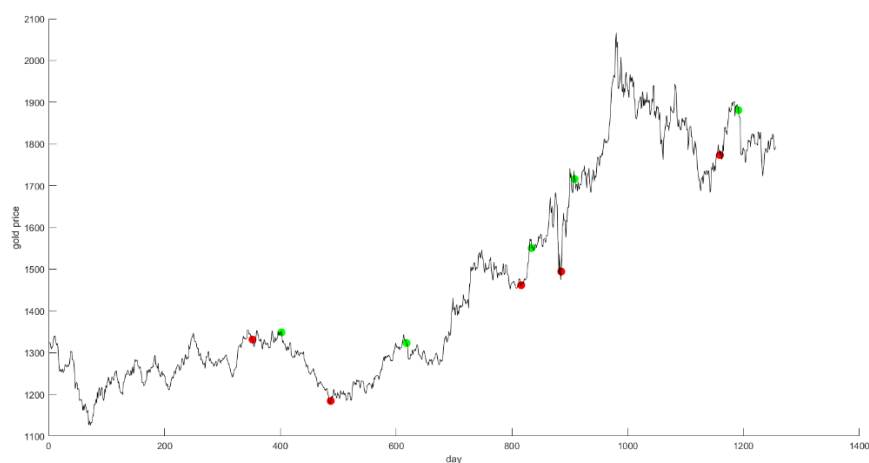


Figure 14: Gold Buy-sell Point Chart

### 6.3.3 Asset Value Estimation

After the above discussion, we obtained a graph of the change in asset value of the original \$1,000 from September 10, 2016 to September 10, 2021. It was eventually learned that the final asset value on September 10, 2021 was \$442,697.25, with a five-year total return of 44,169.70%, an extremely good result compared to bitcoins 7458.97% and golds 135.48% over the same period, with an annualized return of 338.08%

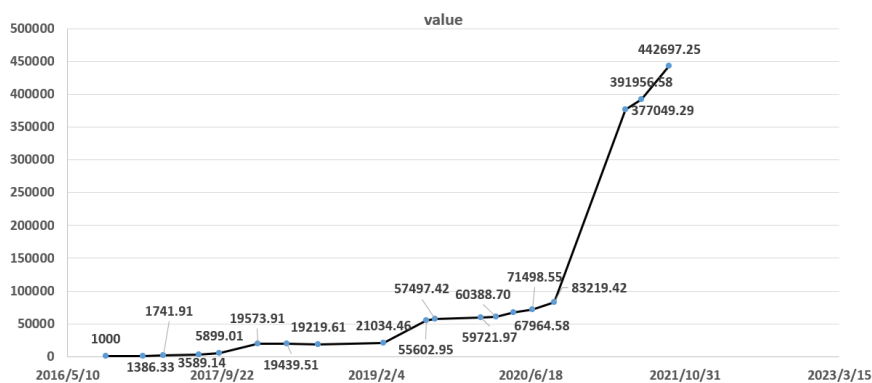


Figure 15: Asset Value (2016/9/10-2021/9/10)

After discounting, we get that \$442,697.25 on 9/10/21 is worth \$426188.16 on 9/10/16.

## 7 Effectiveness Evaluation of The Trading Strategy

In the previous section, we have calculated the value that we have on 9/10/21 after implementing our strategy. To prove that our strategy is optimal in both aspects, we try to change the strategy itself and change the market environment respectively.

In the previous section, we have calculated the value we have on 10 September 2021 after implementing our strategy. We have attempted to demonstrate that our strategy is optimal in both respects, in terms of changing the strategy itself and in terms of changing the market environment, respectively.

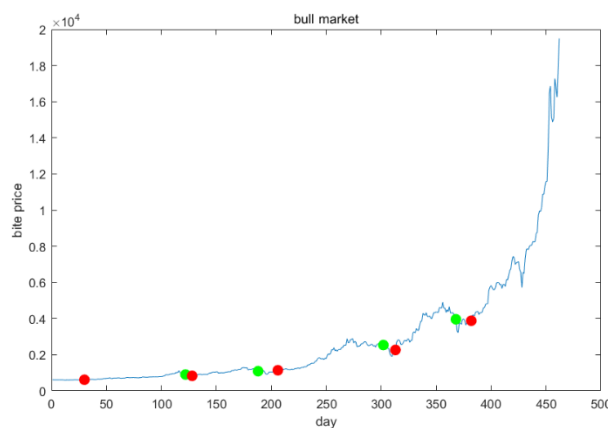
For changing the market environment, we zoomed in on the market at a particular time, extracting bull, bear and, shock markets, and proved that the strategy outperformed the market over the same period in terms of profitability in either market; for changing the strategy itself, we set it as a change in the stop-loss strategy, which was used as a proportional stop when we used the genetic algorithm for strategy optimization, in order to verify the In order to verify the validity of the model, we changed it to a capital stop loss, proving that our proposed strategy is better regardless of the stop loss strategy.

### 7.1 Bull and Bear Market Strategies

When evaluating a fund manager, one looks beyond the returns of the fund manager's funds to judge the ability of the fund manager's funds to weather cycles. In layman's terms, this means judging whether the fund manager's portfolio has equally good returns in markets with different environments such as bull, bear and shock markets.

Given that both gold and bitcoin have experienced a full bull, bear and shock market over the 5 years of data given, in order to evaluate our strategy's ability to ride out the cycle, we have extracted the performance of our strategy in bull, bear and shock markets separately and compared it to the market conditions over the same period to draw our conclusions.

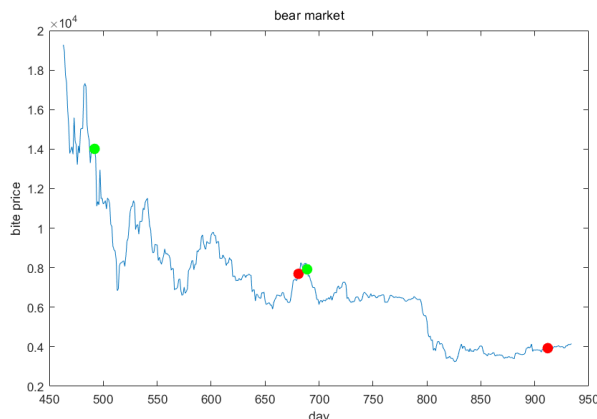
Firstly, the performance of our strategy in a bull market is shown in the chart below. As you can see from the chart, we have been able to maintain a holding position during large market rallies and sell assets in time to hedge our risk during some minor fluctuations, which shows that our strategy has also performed well in a bull market.



**Figure 16:** Trading Spots During Bull Market

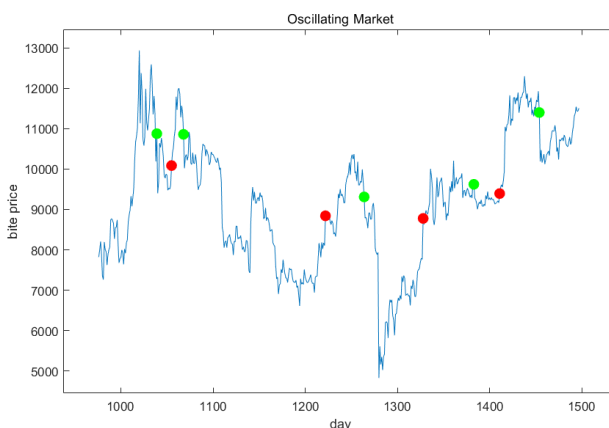


In the midst of bear markets, it can be seen that we have been in a short position during bad market conditions, while at the lowest point of the bear market, we chose to buy, and our strategy has performed well above the market level over the same period.



**Figure 16:** Trading Spots During Bear Market

In the midst of an oscillating market, we still managed to buy and sell in a timely manner and the strategy continued to perform well.



**Figure 16:** Trading Spots During Oscillating Market

## 7.2 Stop Loss on Capital

In gold trading, it is important to set a stop loss. Its main purpose is to help investors manage their investment risk. If an account is losing money, it can be sold to lock in the loss and prevent further declines.

Stop loss levels are used to avoid large price fluctuations by setting a profit and loss point in advance. A capital stop is usually set at a maximum loss that the investor can afford and once the strategy has been implemented, once the loss on a single trade reaches that value, a sell will be made, regardless of expectations.

In operational terms, we simply change Gene4 to Gene4' in the genetic algorithm optimization of the original strategy. As shown in the table.

**Table 6:** Stop Loss Comparison Table

code	000	001	010	100	011	110
Gene4' (Stop Loss Amount)	100	200	500	700	1000	1500
Gene4(Stop Loss Percentage)	1%	2%	3%	5%	8%	10%

Ultimately, we found that the yield after the change was reduced from 442% is reduced

to 273%, the return of the original strategy is much higher than the return of the current strategy, so it can be justified that the article is now a more reasonable strategy.

### 8 Transaction Cost Sensitivity Analysis

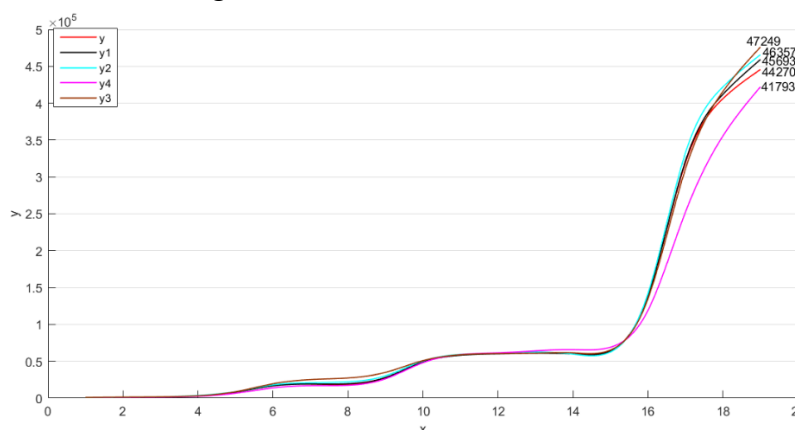
Transaction costs are one of the key indicators that affect a strategy's yield. Many strategies have good rates of return under ideal conditions, but may have lower rates of return once transaction costs change. If trading outcomes are more sensitive to transaction costs, it is an indication that the strategy may not be realistic in the face of changing market conditions.

To see the impact of transaction costs on trading outcomes, we replaced the transaction costs with the following four scenarios:

**Table 7:** Table of Changes in Commission

reducing commissions for gold only	$\alpha_{GOLD} = 0.5\%, \alpha_{BTC} = 2\%$
reducing commissions for bitcoin only	$\alpha_{GOLD} = 1\%, \alpha_{BTC} = 1\%$
reducing commissions for both assets	$\alpha_{GOLD} = 0.5\%, \alpha_{BTC} = 1\%$
increasing commissions for both assets	$\alpha_{GOLD} = 2\%, \alpha_{BTC} = 3\%$

The results are shown in figure 17:



**Figure 17:** Asset Values Under the Influence of 5 Costs

We find that when the commission is reduced on only one asset, the transaction frequency of the asset with reduced transaction cost will increase. But bitcoin has increased more frequently than gold. The reason lies in the large fluctuation and high frequency of bitcoin, which makes it easier for the price to touch the upper and lower limits set by us when making decisions, so as to initiate buy or sell orders. However, for the final return on assets, the reduction of commission on the two assets will increase the return on assets accordingly.

When the commission of the two assets is reduced simultaneously, we find that the return of the asset after the change is greater than the return of the commission of only one asset. Lower commissions on two assets are traded more frequently than if they were reduced on just one asset.

When the commission of two assets is increased at the same time, the final return and trading frequency of the assets are the smallest in the four cases. The reason is obvious. The increase in commissions will reduce our potential profit opportunities, and our capital will be significantly reduced as we trade.

## 9 Memorandum to Investor

Dear Investor:

I am the developer of this investment strategy. There are a few things we would like to tell you about our models, strategies and results. We have used the following models to develop and improve our investment strategy, which has resulted in good returns.

**Firstly**, we used an ARIMA-LSTM model to forecast the future movement of asset prices. The reason for forecasting is that our investment decisions are based on our expectations, or judgements, of the future movement of asset prices. This forecasting model is the cornerstone of our strategy formulation. The forecasting model is a combination of the ARIMA model and the LSTM model, as both models have strengths and weaknesses that complement each other relatively well; the ARIMA model has a good grasp of the general trend of assets, but lacks the volatility of asset prices, while the LSTM model has the opposite characteristics of the ARIMA model. It was therefore reasonable to assume that the combination of the two would make our forecasts more accurate, and this was indeed the case.

**Secondly**, based on the forecasting results, we use a dynamic programming approach to develop a preliminary strategy. We stand at each point in time and forecast the return for a number of days ahead, buying if the projected return is greater than the commission rate and selling if the loss is greater than the commission. However, such an approach proved to be delayed in capturing price trends and was only sensitive to short, large changes in price movements, missing many profit opportunities for nothing.

**Finally**, based on the shortcomings of the original strategy, we proposed to optimise the strategy using a genetic algorithm to improve indicators such as trading time, trading frequency, and stop-loss point settings, ultimately achieving a 440x return, much higher than the returns of bitcoin and gold over the same period.

Our strategy incorporates two assets, one in bitcoin and one in gold. The strategy is, in general, biased towards bitcoin rather than gold. This is because bitcoin tends to be more volatile in price than gold and its gains tend to be greater than gold, according to forecasts. Therefore, when capturing a profitable opportunity in bitcoin over the next few days, we will choose to sell gold at the right place in order to get the bigger gains inherent in the bitcoin trade.

It is worth noting that as an emerging asset, Bitcoin comes with a high level of risk along with a high level of reward. As our strategy is not position-controlled in every trade, but chooses to invest as much of the asset as possible, only the extremely risk-averse will be able to take advantage of our strategy. It is important to note that our strategy does not necessarily trade on a daily basis. Warren Buffett once said, "To be a good hitter, you have to have good balls to hit." This means that we need to be patient and wait for good opportunities rather than struggling to take every swing every chance, and indeed, this is almost impossible for traders to do. We have to give up something to get more. Secondly, our strategies are not foolproof. The vast majority of the time it is not possible to buy at exactly the lowest point and sell at the highest. Investors therefore need to be understanding of this phenomenon.

Finally, thank you for your interest in our strategies and we wish you all the best in life.

Yours sincerely,  
Lee.

February 21, 2022

## 10 Model Evaluation

### 10.1 Strengths

- By combining the ARIMA model with the LSTM model, both linear and non-linear situations can be taken into account, allowing attention to be paid to both trend changes and volatility.
- The ARIMA-LSTM model can be used to forecast the future trend of stock prices simply by using the evolution of the historical state of the stock itself, which is simple and reliable compared to traditional mathematical and statistical methods.
- The genetic algorithm is applied to the optimization of trading strategies, which not only optimizes existing strategy combinations, but also generates new trading strategies through the evolution of the system, and allows the optimal strategy to be selected for future trading in real time.
- In order to study the effectiveness of the model, we analyze both the two market environments and the strategies themselves to make the discussion of effectiveness more comprehensive and reasonable.

### 10.2 Weaknesses

- The financial markets are highly volatile, and the time series analysis method only uses historical price data in the hope of obtaining useful information to predict future movements, without considering the causes of stock price movements, so it is generally an intuitive analysis and only makes short term predictions.
- For the strategy combination approach of generating new stocks, because of the single measure taken and the small number of indicators, it does not take into account the different risk-return characteristics of different trading strategies, and the fitness function in this paper only measures the performance over time, without considering the cumulative performance.

## 11 Conclusion

As our team set out to come up with a strategy on what would be the most efficient way to solve the problem of investment strategy setting. First of all, we should not stand in the perspective of God to invest, and we need to predict the daily asset price. In recent years, the problem of asset price prediction has been realized by many scholars in different ways. We summarize the previous practice, and further propose the method of combining linear and nonlinear models, namely ARIMA-LSTM model. In order to further analyze the strategy setting, we use the mathematical model of dynamic programming to transform long-term investment problems into short-term ones, refine the investment process, and more importantly, we use genetic algorithm to optimize the trading strategy. We proved the effectiveness of the strategy by changing the strategy and changing the market environment. In short, we were

sure that the strategy we proposed was reasonable and effective.

In the following calculation and research, we reasonably extended our strategy and result, studied the impact of commission on strategy and result, and obtained the relationship between commission and them. However, as the title only requires the use of price to set strategies, there are many other factors that affect asset price changes, not only past prices, which can be improved in subsequent studies.

## References

- [1] Yao HX, Lai JW, Xia SHENGHAO, Chen SHUMIN. Fuzzy transaction decision making based on Apriori algorithm and neural network[J]. *Systems Science and Mathematics*,2021,41(10):2868-2891.
- [2] Hu Wenxiu, Su Zhenxing, Yang Li. A study of investors' conceptual attention on conceptual index return prediction and trading strategy based on random forest method[J]. *Forecasting*,2021,40(01):60-66.
- [3] Gong Yajian, Wei Xianhua, Meng Xiangying, Liu Chenhao. Can genetic programming strategies be applied to the Chinese stock market? -- A study of stock index trading strategies based on stochastic multi-objective genetic programming[J]. *Systems Science and Mathematics*,2020,40(12):2381-2400.
- [4] Sun Dachang,Bi Xiuchun. High-frequency trading strategies based on deep learning algorithms and their profitability[J]. *Journal of the University of Science and Technology of China*,2018,48(11):923-932.
- [5] Peng Songlin,Su Dongwei. Short-selling transactions, investment strategies and stock returns - an empirical test based on the Chinese A-share market[J]. *Industrial and Economic Review*,2017,8(04):135-153. doi:10.14007/j.cnki.cjpl.2017.04.012.

## Appendices

### Appendix 1

#### Introduce: Prediction code

```

a= xlsread('biteb.xlsx');
data=a(:,1);
TTrain = floor(0.95*numel(data));
dataTrain = data(1:TTrain+1);
dataTest = data(TTrain+1:end);
sig = std(dataTrain);
Standardized = (dataTrain - mu) / sig;
XTrain = Standardized(1:end-1);
YTrain = Standardized(2:end);
numFeatures = 1;
numResponses = 1;
numHiddenUnits = 200;
layers = [ sequenceInputLayer(numFeatures) lstmLayer(numHiddenUnits)
fullyConnectedLayer(numResponses)  regressionLayer];
options = trainingOptions('adam', ...
    'MaxEpochs',250, ...
    'GradientThreshold',1, ...
    'InitialLearnRate',0.005, ...
    'LearnRateSchedule','piecewise', ...
    'LearnRateDropPeriod',125, ...
    'LearnRateDropFactor',0.2, ...
    'Verbose',0, ...
    'Plots','training-progress');
dataTestStandardized = (dataTest - mu) / sig;
XTest = dataTestStandardized(1:end-1);
net = predictAndUpdateState(net,XTrain);
[net,YPred] = predictAndUpdateState(net,YTrain(end));
numTimeStepsTest = numel(XTest);
for i = 2:numTimeStepsTest
    [net,YPred(:,i)] = predictAndUpdateState(net,YPred(:,i-1),'ExecutionEnviron
    ment','cpu');
End
figure
plot(YTest)
hold on
plot(YPred,'.-')
hold off
legend(["Observed" "Forecast"]); ylabel("Cases") title("Forecast")

```

**Appendix 2****Introduce: Strategic planning code**

```

function [nn,bn,cn,an,sv]=strategy(x,t)
a=1000;b=0;c=0;sx=1;sv=0;nn=[];bn=[];cn=[];an=[];sxf=0;
bite = xlsread('bitezhen.xlsx');
gold = xlsread('goldzhen.xlsx');
by = xlsread('biteyuce.xlsx');
hy = xlsread('goldyuce.xlsx');
[a,b]=bitemai(a,bite(1));
[bl,hl]=shoouyilv(x);
sx=sx+1;
n=1;nn=[nn,n];bn=[bn,b];cn=[cn,c];an=[an,a];
for sx=2:t
    if hl(sx)>=bl(sx)
        if hl(sx)>-0.01
            k=a+b*bite(sx)*0.98-gold(sx)*0.99;
            if k>=gold(sx)
                a=0;b=0;
                [a,c]=goldmai(a+b*bite(sx)*0.98,gold(sx));
                if a==0
                    n=2;
                else
                    if bl(sx)>0
                        [a,b]=bitemai(a,bite(sx));n=4;
                    else
                        n=5;
                    end
                end
            end
        else
            if c>0
                if bl(sx)>=0
                    [a,b]=bitemai(a,bite(sx));n=4;
                else
                    a=b*bite(sx)*0.98+a;b=0;n=5;
                end
            else
                if bl(sx)>=0
                    [a,b]=bitemai(a,bite(sx));n=1;
                else
                    a=b*bite(sx)*0.98+a;b=0;n=3;
                end
            end
        end
    end
end

```

```
                end
            end
        else
            a=a+b*bite(sx)*0.98+c*gold(sx)*0.99;b=0;c=0;n=3;
        end
    else
        if bl(sx)>-0.02
            [a,b1]=bitemai(a+c*gold(sx)*0.99,bite(sx));c=0;
            b=b+b1; n=1;
        else
            a=a+b*bite(sx)*0.98+c*gold(sx)*0.99;b=0;c=0;n=3;
        end
    end
    nn=[nn,n];
    bn=[bn,b];
    cn=[cn,c];
    an=[an,a];
    sv=a+b*bite(sx)+c*gold(sx);
    end
    sv=sv-1000;
    nn=nn';
    an=an';
    bn=bn';
    cn=cn';
    function [a,b]=bitemai(sv,x)
    b=floor(sv*10/x)/10;n=0;
    a=sv-x*b*1.02;
    if a<0
        n=0.1;
        while n*x<(-a)
            n=n+0.1;
        end
    end
    b=b-n;
    a=a+x*n;
    function [a,b]=goldmai(sv,x)
    b=floor(sv/x);n=0;
    a=sv-x*b*1.01;
    if a<0
        n=1;
        while n*x<(-a)
            n=n+1;
```



```
    end
end
b=b-n;
a=a+x*n;
function [blz,hlz]=shoouyilv(x)
blz=[0];hlz=[0];
by = xlsread('biteyuce.xlsx');
hy = xlsread('goldyuce.xlsx');
for sx=2:1255
    if sx-x>0
        bl=(by(sx)-by(sx-x))/by(sx-x);
        hl=(hy(sx)-hy(sx-x))/hy(sx-x);
    else
        bl=(by(sx)-by(1))/by(1);
        hl=(hy(sx)-hy(1))/hy(1);
    end
    blz=[blz,bl];hlz=[hlz,hl];
end
```