Problem Chosen	2022	Team Control Number
C	**MCM/ICM Summary Sheet**	**2214872**

How to earn from \$1000 to \$300000: LSTMIS-Based Quantitative Portfolio Investment Model

Summary

With the development of financial market, portfolio investment has become a new hotspot in the field of quantitative investment. Our team is asked to develop a model to propose the best strategy of gold and bitcoin portfolio investment.

As for question 1, we first use LSTMIS to predict the value of gold and bitcoin with input sequence consisting of last 30 days data. Then we use the predicted data and the mean value of last 5 days as long- and short-term moving average input of AMTM model respectively, hence judging whether to buy or sell. Afterwards we treat the portfolio of gold and bitcoin as venture capital, cash as risk-free investment. Considering both risk and gains, betterment of trading strategy can be addressed as an optimization problem based on Markowitz Mean Variance Model(MOP). Optimal combination of venture capital as well as daily business transaction can be obtained through optimization. We incorporate MOP with AMTM to get the iterative formula of daily shares $[D^{(T)}, G^{(T)}, B^{(T)}]$. Note that we should obey one rule: the total amount of gold remains unchanged on days when gold market closes. Ultimately, we can calculate that our total assets are **310445.96 dollors** after 5 years of trading based on our strategy with the start cash of 1000 dollars. We also make a table to record main trading processes.

As for question 2, we analyze the effect of our model used in question 1 and do **backtest** based on Python Lib backtrader, in order to evaluate our model. First we evaluate the prediction model LSTMIS by analyzing the effectiveness of updating the model daily using all historical data. We conclude that its prediction is more accurate and valid compared with other prediction models since it considers nonlinear factors and updates daily to learn up-to-date information, and sensitive to the length of input sequence. Next we use backtrader to backtest our trading strategy by only trading gold or bitcoin. Summary can be drawn that the profit rate of gold is **remarkably less** than the result we acquire in question 1, however the profit rate of bitcoin is **significantly greater** than that of previous result. In consequence, we can safely conclude that our model realizes fine trade-off between risk and gain.

As for question 3, we first make adjustments to the commission of gold and bitcoin in $[0, 3]$, and find out that the total assets after implementing our strategy for 5 years **decline monotonically** by and large with increasing commission, and **local maximum value** exists. It follows that the existence of commission, on the one hand, **reduces the profit rate** of total assets and, on the other hand, **increases trading thresholds**, therefore to some degree lowering the trading frequency.

Lastly, we test LSTMIS's sensitivity by changing **the length of input sequence** and discuss the strengths and weaknesses of our model.

Keywords: **LSTMIS Average Moving Markowitz Mean Variance Model Iterative Formular Gold Volume Conservation Backtrader Backtest**

Contents

1 Introduction

1.1 Background

Gold and digital currencies, both of which play an inflation-resistant role, have become a hedge against the backdrop of uncertainty and looming inflation in the global situation. The total market capitalization of the gold market is estimated to be over $11 trillion, reflecting the fact that gold is still a mainstream investment option as well as a safe-haven asset. in 2020, the average daily trading volume of gold was $125 billion, about 30 times the average daily spot trading volume of Bitcoin at $4 billion. Nonetheless, both assets have extremely liquid markets, which means there is plenty of room for both digital currencies and gold to grow, while making gold and bitcoin investments becomes an emerging portfolio. As the financial markets evolve, how investors construct the right portfolio has become an important topic of study in economics, and the rise of quantitative investing has gone some way to solving that problem. Quantitative investment is an investment method based on data, programmed trading, strategy models and the pursuit of absolute returns. Compared with traditional investment, it relies entirely on massive data and models and can fully realize the risk diversification of the strategy. It has a wide range of information sources, a short investment cycle, and issues buy and sell orders through a computer program, which excludes human cognitive bias and subjective judgment, and therefore has more stable returns. Therefore we need to use quantitative investment as the core to provide investors with the best portfolio strategy.

1.2 Restatement of the Problem

To facilitate more accurate decision-making, traders asked us to develop a model that uses only past daily price flows to date to determine whether a trader should buy, hold, or sell assets in his or her portfolio each day. Using the five-year period from November 9, 2016 to October 9, 2021 as the trading period, with November 9, 2016 as the starting time and $\$1,000$ as the initial capital, for each trading day thereafter, we are required to set up a portfolio consisting of cash, gold, and bitcoin $[C, G, B]$ in U.S. dollars, troy ounces, and bitcoin, respectively. The initial state is $[1000, 0, 0]$. In addition to being going to ration the shares of gold, cash and bitcoin, we need to note that there is a commission cost for each transaction, which is $\%$ of the transaction amount for each transaction (buy and sell). Let's assume that gold= is 1% and bitcoin= is 2%. There is no cost to hold the asset. It is worth noting that bitcoin and gold do not have different trading rules. Bitcoin can be traded daily, but gold is only traded on the day the market is open, as reflected in the data files **LBMA-GOLD.csv** and **BCHAIN-MKPRU.csv**. Your model should take this trading schedule into account. However, when building the model, we will only be able to use the**LBMA-GOLD.csv** and **BCHAIN-MKPRU.csv** data tables as the only reference data for the model.

1.3 Analysis of the Problem

- **Analysis of Problem1 :** In response to question 1, the first question asks us to develop a model and give the best daily trading strategy based only on the price data of the day. The essence is to use predictive models and portfolio principles to give reasonable investment patterns and position ratios, in addition to considering whether it is a trading day, when to buy and sell, and other important time points.

- **Analysis of Problem2 :** For question 2, we are ostensibly asked to provide evidence that the model proposes the best strategy, however, its The essence is the accuracy test and analysis of the model, that is, giving indicators to argue the rationality of the decision. The larger the error in the constructed model, the worse the decision will be. We can use theoretical backtesting methods in finance to apply the model to a period of known prices to determine the reasonableness of the model. Since there is a huge difference between the actual and simulated earnings, we can generate some indices: Sharpe ratio, annualized rate of return to think about the analysis and discussion from an economic point of view.

- **Analysis of Problem3 :** For question three, here the question is brought into the model again by increasing or decreasing the percentage of commission (transaction cost) to test the magnitude of sensitivity at different transaction costs. Here, in addition to analyzing the results for different sensitivity levels, the financial necessity of the sensitivity analysis needs to be analyzed in relation to the relevant knowledge in finance. In addition, the robustness of the model, i.e. whether the profitability of the constructed model is more stable under different commission costs, is still to be argued here.

- **Analysis of Problem4 :** In response to the fourth question, the fourth question asked us to write a two-page memo to summarize the decision metrics we used in the process, the robustness of the model, the decision payoffs, and several other aspects of the analysis, in addition to the analysis of the model itself, but also the analysis of the gold and bitcoin market shocks.

2 Assumptions and Justification

To simplify the problem and make it convenient for us to simulate real-life conditions, we make the following basic assumptions, each of which is properly justified.

- It is believed that investors are risk averse, i.e., more inclined to forgo fair bets and worse investments and more willing to accept investments with positive risk premiums, and their risk aversion is measured by A.

- The \$1,000 remaining assets are considered to be used for risk-free investments, i.e., where there is a stable low return r_f and no risk is taken into account. (without considering leveraged investments, futures investments and any borrowing, and without any fees)

- The magnitude of the return on each asset can be described by a normal distribution curve, i.e., the mean and variance (or standard deviation) can be calculated.

- Each asset is infinitely divisible, meaning that the investor can buy a portion of a share if he wishes.

- When making investment choices, the investment manager is interested in only two parameters of the probability distribution of the investment return: the expected return and the variance, respectively. The former reflects the investor's level of expectation of future investment returns, while the latter reflects the investor's risk trade-off

3 Notations

The primary notations used in this paper are listed in Table 1.

Table 1: Notations

Symbols	Definition	Type
$DTSM$	Daily Trading Strategy Model	Mapping
$LSTMIS$	LSTM-based Investment Strategy	
$MATS$	Moving Average Trading Strategy	Mapping
MOP	Markowitz Optimal Portfolio	Mapping
h	hidden state	Mapping
c	cell state	Mapping
$h_i^(j)$	i^{th} value of input-layer through value of output-layer	Scalar
W_f, W_i, W_0	Transformation matrix of forget , input, ouput gate respectively	Mapping
T	the index of current date	Scalar
$[D^{(T)}, G^{(T)}, B^{(T)}]$	Portfolio of USD,gold and bitcoin,superscript is its associated date index	Vector
$[w_g^{(T)}, w_b^{(T)}, y^{(T)}]$	Ratio of gold,bitcoin and USD respectively,subscripts g or b mean gold or bitcoin	Vector
r_f	Risk-free yield of USD,where we define it to be 0.01	Scalar
A	Risk aversion coefficient, where we define it to be 2	Scalar
$GM(T+1, m)$	Short-time forecast by average of slice T-m : T	Scalar
$GM(T+1, n)$	Lond-time forecast by LSTM of slice T-m : T	Scalar
$Cap^{(T+1)}$	Total capital associated with T+1	Scalar
β	Maximum asset shrinkage index	Scalar
α_{gold}	Transaction cost(purchase or buy) of gold	Scalar
$\alpha_{bitcoin}$	Transaction cost(purchase or buy) of gold	Scalar
S	Sharp ratio	Scalar

where we define the main parameters while specific value of those parameters will be given later.

4 Model Overview

In our basic model, we aim to achieve three objectives: modeling the trading strategy on a daily basis, evaluating the modeled trading strategy, and analyzing the impact of changes in trading costs in the model on the model's profitability results. We therefore divide the model into three main parts.

- The first part forecasts and fits a Markowitz portfolio investment model combining the quantitative investment part of economics through AI algorithms such as neural networks.
- The second part mainly evaluates the return and risk of the model by introducing some economic concepts such as annualized rate of return, Sharpe ratio, etc. In addition, we will use the concept of backtesting, which is commonly used in economics, to verify the effectiveness and feasibility of the model.
- The third part is a sensitivity analysis of the model by changing the amount of gold and bitcoin commissions in different ways to verify the robustness of the model.

The overall modeling and analysis process is shown in the figure below.

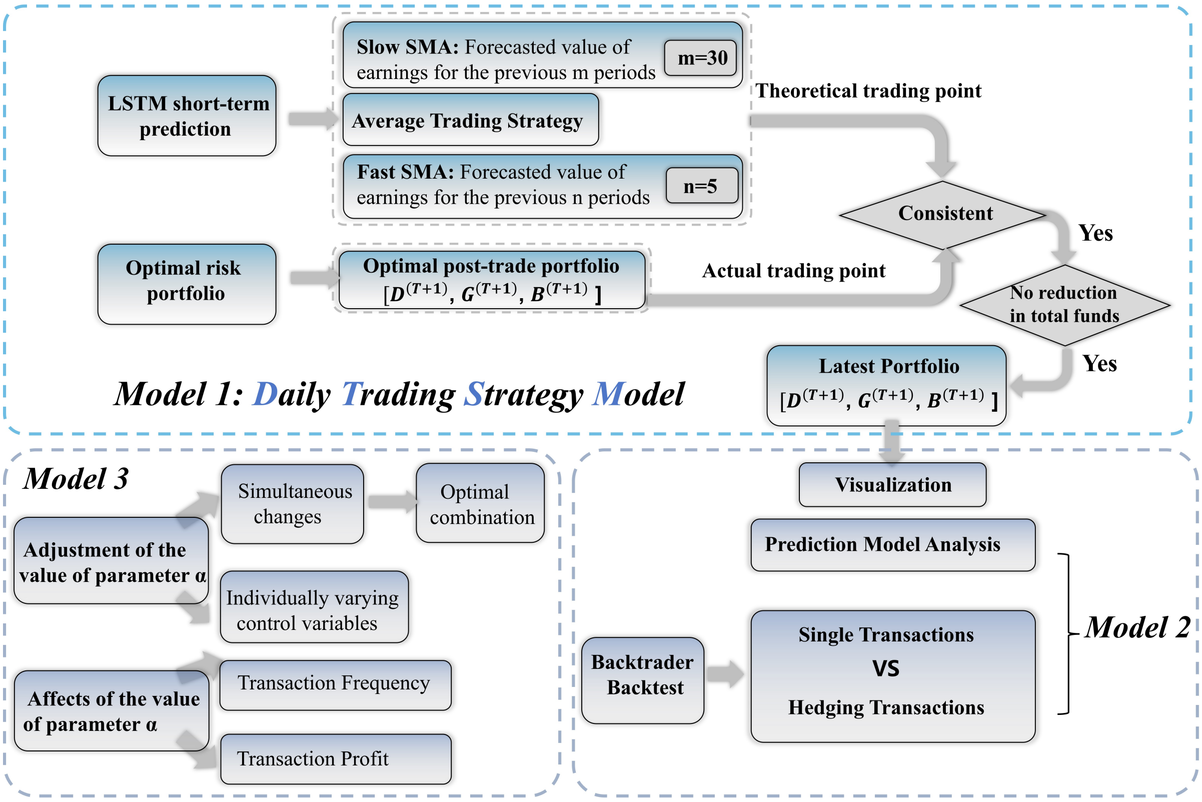


Figure 1: Modeling process

5 Daily Trading Strategy Model

Traditional Markowitz Mean Variance Model provides the optimal portofio, given two types of risk venture and one risk-free investment, which achieving the minimum of risk and the maximum of profit simultaneously.But can only give strategy of one day. Traditional Average Moving Trading Strategy judges the gold cross point and death cross point, through long-time and short-time average line.We creatively integrates the above two model, using AMTS to give trade signals and MOP to determine accurate volume of trade each day.Moreover, we substitute LSTM into the AMTS as long-time prediction, and develop strategies for the commission α and close day of gold respectively. The specific algorithmic process is shown below.

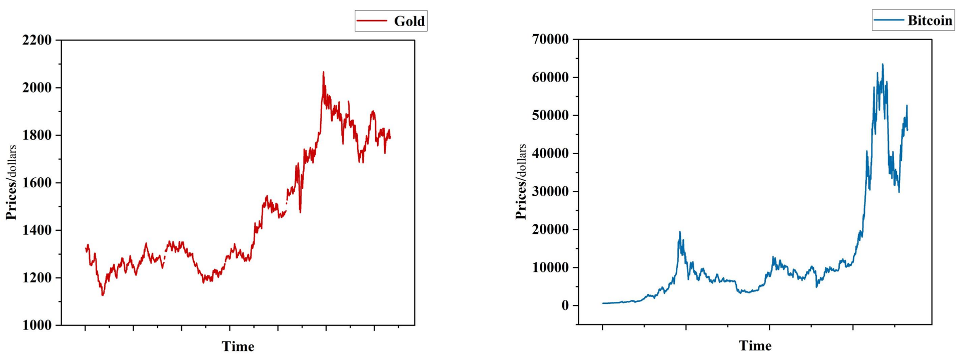
Algorithm 1: Daily Trading Strategy Model

input : Daily price of gold and bitcoin from 9/11/2016 to 9/10/2021
output : Final portfolio $[C, G, B]$

Initialize : $[D^{(0)}, G^{(0)}, B^{(0)}] = [1000, 0, 0], Cap^{(0)} = 1000, T = 0;$
Compute $GM(T+1, n)\, and\, GM(T+1, m)$;
`// modify long-time prediction of MATS with` **`LSTM`**
for $T \leftarrow 30$ **to** l **do**
  if *Both gold and bitcoin cross **Ideal Profit Point and Loss Point*** **then**
  Compute w_g^*, w_b^*, y^* with **Markowitz Optimal Porfolio** ;
  Update porfolio $[C, G, B]$ with **Updating Eq12** ;
  `// Eq12 obeys the conservation of total capital:`$Cap^{(T)} = Cap_0$
  if *Gold can be traded* **then**
    Calculate as Eq8
  `//` **`Eq8`** `obeys the conservation of gold holdings`
  if *Constraints are established successfully* **then**
    `/* Eq13 means ideal trade analysis(calculated by MATS) is consistent with actual trade strategy(calculated by MPO). */`
    Trading successfully according to Eq8 and Eq12 ;
  ;
  else Skip trading;

5.1 Data Pre-processing and Visualization

First, we need to process the data in the two official price data tables LBMA-GOLD.csv and BCHAIN-MKPRU.csv. provided by Metsä about gold and bitcoin between 2016 and 2021, and after unifying the formats we can get the graphs of gold and bitcoin prices over time respectively as shown in Figure 2.



(a) Gold price fluctuations between 2016 and 2021 (b) Bitcoin price fluctuations between 2016 and 2021

Figure 2: Gold and Bitcoin price fluctuations between 2016 and 2021

As we can see from Figure 1, the price curve for gold is not continuous, which is in line with the market rule that gold is not traded every day. In contrast, the curve for Bitcoin is continuous, indicating that Bitcoin is not affected by trading hours.Based on the data and the images obtained, it is easy to see that the two datasets are extremely nonlinear, so we can use AI prediction algorithms such as neural network algorithms to model and analyze

5.2 LSTMIS: LSTM-based Investment Strategy

LSTM, short for long-short term memory recurrent neural network, is a powerful and effective tool to process sequence data. Based on LSTM, we create a prediction model named LSTMIS, which uses the most recent 30 days to predict tomorrows bitcoin value or gold value, so that we can wisely invest our money.

To predict the value on the t^{th} day v_t, we input v_{t-30},v_{t-29},...,v_{t-1}, together with hidden states of yesterday $[h_0^1, h_0^2, ..., h_{256}]_{256*1}$ and cell states $[c_0^1, c_0^2, ..., c_{256}]_{256*1}$, which convey historical information and help to form gates.Gates determine what information should be remembered, what information should be forgot and what information should be output. They consist the inner structure of LSTM, as is shown in ... The gates, states and outputs are defined as follows:

$$\begin{cases} c^t = z^f \odot c^{t-1} + z^i \odot z \\ h^t = z^0 \odot \tanh(c^t) \\ y^t = \sigma(W'^{h^t}) \\ z^f = \sigma(W_f^T[h_{t-1}, x_t]) \\ z^i = \sigma(W_i^T[h_{t-1}, x_t]) \\ z^0 = \sigma(W_0^T[h_{t-1}, x_t]) \\ z = \sigma(W^T[h_{t-1}, x_t]) \end{cases} \tag{1}$$

where $\odot$ means element-wise multiplication and sigma means sigmoid function $y = 1/(1 + exp(-x))$

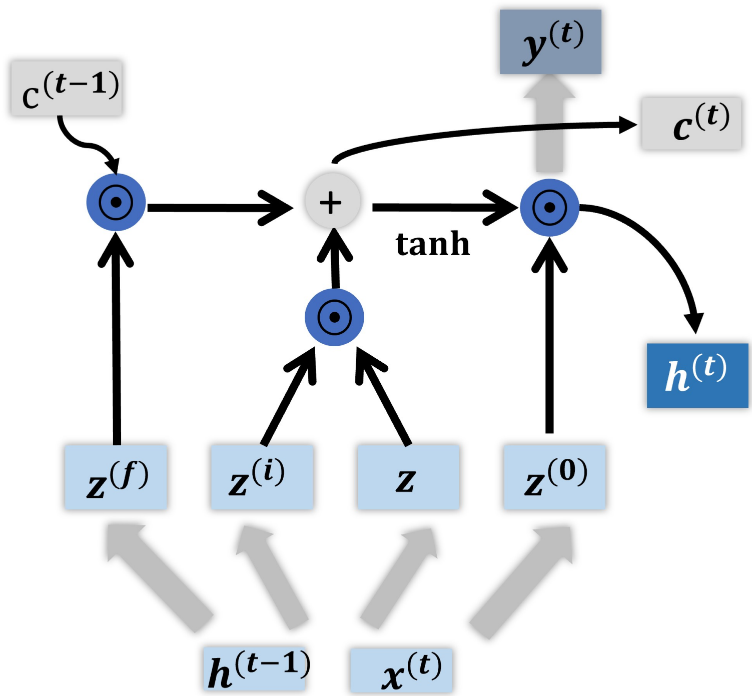


Figure 3: LSTM inner structure

When v_{t-30} together with $h_0^{(1)}, c_0^{(1)}$ are input into the first LSTM, $h_{t-30}^{(1)}, c_{t-30}^{(1)}$ are output. Then $v_{t-29}, h_{t-30}^{(1)}, c_{t-30}^{(1)}$ are input into the first LSTM, $h_{t-29}^{(1)}, c_{t-29}^{(1)}$ are output. Similar cycle occurs until $v_{t-30}, v_{t-29}, ..., v_{t-1}$ have all been input into the first LSTM and corresponding hidden states $h_i^{(1)}, i = t-30, ..., t-1$ and cell states $c_i^{(1)}, i = t-30, ..., t-1$ are output. Then these hidden states are input into the second LSTM network and the same process occurs, with $h_i^{(2)}, i = 0, t-30, t-29, ..., t-1, h_i^{(2)}, i = 0, t-30, t-29, ..., t-1$input and output in the same way. Note that all cell states and hidden states have the size of 2561. The whole two-layer LSTM network generates 256 outputs, and they are input into a fully-connected neural network (FCN), which consists of a hidden layer (64 neurons) and an output layer (1 neuron). Finally, the output of FCN is tomorrows value. The whole structure is portrayed in figure 4.

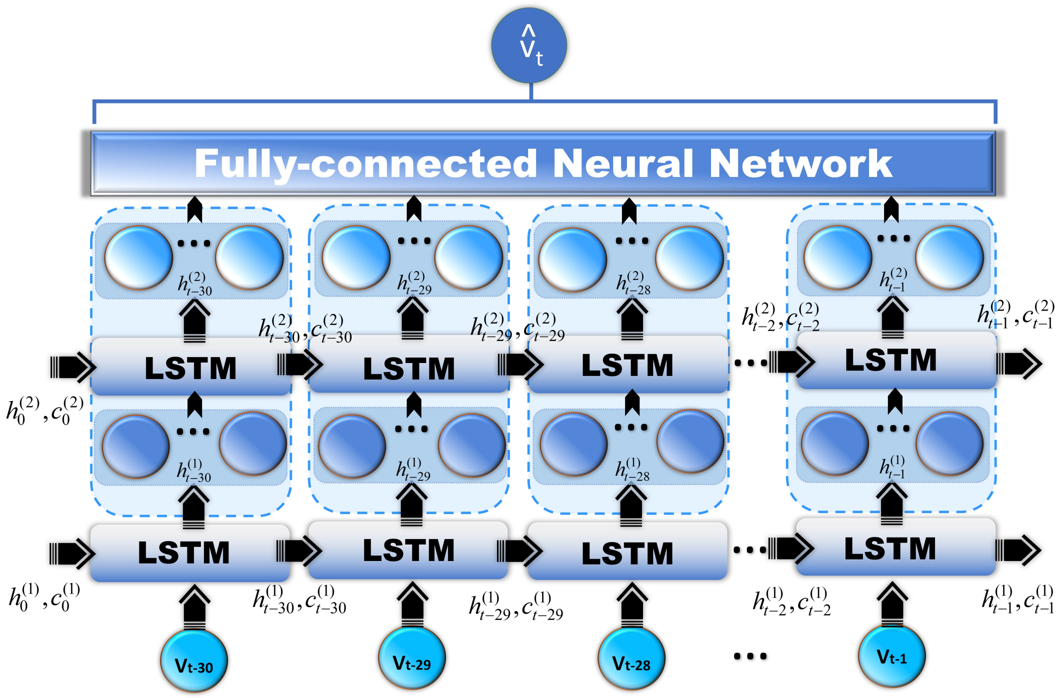
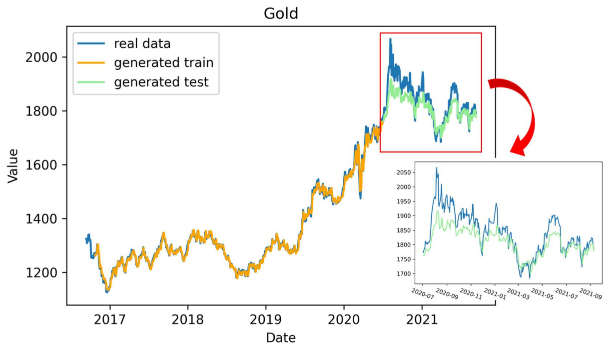
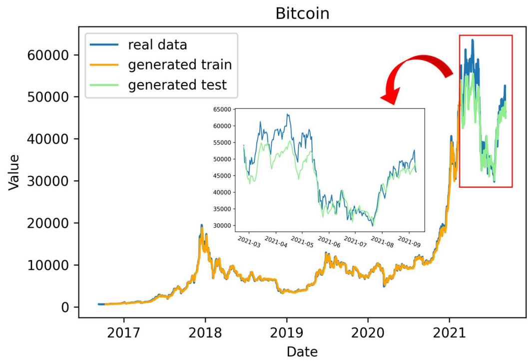


Figure 4: fully-connected neural network (FCN)

Based on mean square error loss and Adam optimizer, the model is trained using back propagation (100 epochs, batch size=10) and tested. We use data of the last 200 days as the test set for bitcoin value and data of the last 300 days as the test set for gold value, while the rest data as their training set respectively. The results are shown in figure above. We can see that our model makes predictions that are very accurate in the beginning. Even if the gap between generated data and real data widens gradually as time passes by, it still accurately predicts the trend value changes, with acceptable minor delays.



(a) Forecast of gold price trend



(b) Forecast of bitcoin price trend

Figure 5: Gold and Bitcoin price fluctuations between 2016 and 2021

In practice, when making a daily trading strategy, we can only get data up to that day. It means we cannot use the all the data in this 5-year period to make a perfect regression and use it to make a daily strategy. Given this fact, we use data of all days prior to the day we want to predict as the training set, and train the LSTMIS model for 100 epochs our model daily. When it comes to the next day, we add that days data to our training set and update the model, in a bid to better predict the value of gold and bitcoin in the future.

When conducting our strategy, we can not only use next days predicted value, but also use predicted value on days after tomorrow, which can be calculated using real data and predicted data, to obtain a likely change trend, therefore helping us better our strategy. And since the model is updated daily, the trend it predicts is more valid and grounded since it learns up-to-date information.

5.3 Moving Average Trading Strategy

The Moving Average Trading Strategy (MATS) forecasts the future trend of the price or price index by selecting the long and short time periods (abbreviated as L, S) respectively, and then generates trading signals and implements the trading strategy based on the results of L and S forecasts.

The short-term forecast is used to track the short-term price trend of the market, and the average of the above $m = 5$ days is used in this study; the long-term forecast is used to track the long-term price trend and possible price reversal of the market, and the forecast result of the LSTM of the above n days is used in this study. In the moving average trading strategy, when the fast average breaks upwards through the slow average as a **golden cross**, it is considered as an ideal buying point to open a position as the price may rise significantly at a future date. When the fast SMA breaks down the slow SMA, a **death cross** is formed and the price is considered to have a tendency to fall in the future, which is considered to be the ideal point to sell short. The timing of both buying and selling is the opening price on the second day of the signal. The rules for calculating the moving average are:

- **Fast averages:** $MA(t,m) = \frac{\sum_{i=1}^{m} r_i}{m}, m = 5$

- **Slow averages:** $MA(t,n) = LSTM(n = 30)$

Considering the existence of handling fees for actual transactions:$\alpha_{gold} = 1\%, \alpha_{bitcoin} = 2\%$,
When $MA(t,m) > (1+\alpha) * MA(t,n)$, buy to open a position at the opening price of T+1
When $MA(t,m) < (1+\alpha) * MA(t,n)$, sell the short position at the opening price of T+1

We use MATS to get the golden cross point and death cross point, and judge whether to trade or not.Then we will introduce the MOP to determine the accurate volume of trade.

5.4 Markowitz Optimal Portfolio

5.4.1 Optimal Hedge Portfolio

Portfolio theory refers to a portfolio of several securities whose return is the weighted average of the returns of these securities, but whose risk is not the weighted average risk of the risks of these securities. The portfolio reduces the unsystematic risk, and we improve and modify the existing Markowitz portfolio model by treating gold and bitcoin as two different stocks and completing the matching of the two through the model, so as to make quantitative investments and reduce the risk while satisfying a higher rate of return.

- **The Optimal Risk Combination of Gold and Bitcoin**

Category	Earnings Value	Risk	Investment ratio
Gold	$E(r_g)$	σ_g	ω_g
Bitcoin	$E(r_b)$	σ_b	$\omega_b = 1 - \omega_g$

When the investment ratio ω_g is reasonably chosen, better portfolio returns and lower portfolio risk can be obtained. We use the maximum Sharpe ratio and the minimum value at risk as the optimization objectives, respectively, and thus jointly build an optimization model with two sets of equations as the core.

- **Optimization model building and constraints** We developed an optimization model with the objectives of maximum Sharpe ratio and minimum value-at-risk

$$\begin{cases} \underset{\omega_g}{max} \frac{E(r_p)-r_f}{\sigma_p} \\ \\ min \sigma_g^2 \end{cases} \tag{2}$$

The restrictions are

$$S.t.\left\{\begin{array}{c}\omega_g+\omega_b=1 \\ 0 \leq \omega_g \leq 1 \\ E(r_p)=\omega_g E(r_g)+\omega_b E(r_b) \\ \sigma_p^2=\omega_g^2 \sigma_g^2+\omega_b^2 \sigma_b^2+2 \omega_g \omega_b \cdot \\ E\left(r_g\right)=\operatorname{mean}\left[r_g^{(T-N)}: r_g^{(T-1)}\right] \\ E\left(r_b\right)=\operatorname{mean}\left[r_b^{(T-N)}: r_b^{(T-1)}\right] \\ \sigma_g=\operatorname{var}\left[r_g^{(T-N)}: r_g^{(T-1)}\right] \\ \sigma_b=\operatorname{var}\left[r_b^{(T-N)}: r_g^{(T-1)}\right]\end{array}\right. \tag{3}$$

Among the restrictions we have the following points to note

- $\operatorname{cov}(G, B)$ is the covariance of r_b and r_g
- When $\omega_g > 1 \text{or} \omega_g < 0$, we take the boundary value
- N is the data period, where N is taken as the period of 7 days

- **Maximum Sharpe Ratio** According to the above formula, we can solve for

$$\omega_g^*=\frac{E(R_g) \sigma_b^2-E(R_b) \cdot \operatorname{cov}(G, B)}{E(R_g) \cdot \sigma_b^2+E\left(R_b\right) \cdot \sigma_g^2-\left[E\left(R_g\right)+E\left(R_b\right)\right] \cdot \operatorname{cov}(G, B)} \tag{4}$$

And among these equations

$$E(R_g)=E(rg)-r_f ; E\left(R_b\right)=E(r_b)-r_f \tag{5}$$

$$\omega_b^*=1-w_g^x \tag{6}$$

- **Minimum total risk value** We can calculate that

$$\omega_g^*=\frac{\sigma_b^2-\operatorname{cov}(G, B)}{\sigma_g^2-2 \operatorname{cov}(G, B)+\sigma_b^2} \tag{7}$$

5.4.2 Markowitz Mean Variance Model

People make investments, essentially choosing between uncertain returns and risks. Portfolio theory uses mean-variance to portray these two key factors. By mean, we mean the expected return of a portfolio, which is a weighted average of the expected returns of individual securities, weighted by the corresponding percentage of the investment. Of course, the return of a stock includes both dividend payout and capital appreciation. By variance, we mean the variance of the portfolio's return. We refer to the standard deviation of returns as volatility, which portrays the risk of a portfolio. How should people choose the combination of return and risk in their portfolio investment decisions? This is the central question in the study of portfolio theory. Portfolio theory is the study of how "rational investors" choose the optimal portfolio. A rational investor is an investor who maximizes expected return for a given level of expected risk or minimizes expected risk for a given level of expected return.

Therefore, based on the above knowledge of finance, we created a flow analysis of the portfolio as shown in the figure 6 below.

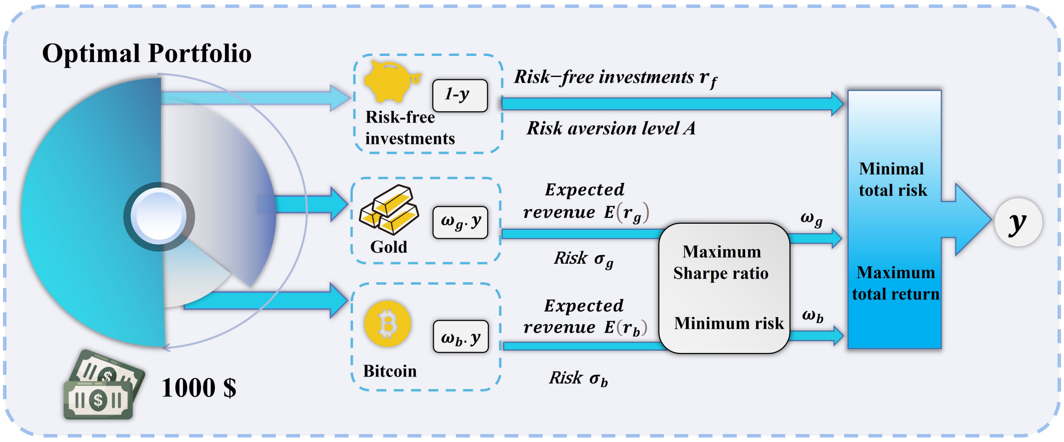


Figure 6: Process analysis of the optimal portfolio

- **Non-differential curve** Now introduce the risk aversion index A, where the value of A is taken to be 2. Considering the total benefit and total risk, the utility function u is constructed, where the contour of the utility function is the undifferentiated curve. The undifferentiated curve is a portfolio that has the same attractiveness to investors. Portfolios located on the undifferentiated curve have the same attractiveness to investors.

- **Capital Allocation Line CAL** CAL denotes all feasible risky portfolios of the investor (determined by ω_g), with a slope S (Sharpe ratio) equal to the expected return that rises for each unit increase in the standard deviation of the chosen asset portfolio. That is, the degree of additional return generated by each unit of additional risk.

- **Introduction of optimization models** The system of equations for the optimization model is as follows

$$\underset{y}{max}(u) = E(r_c) - \frac{1}{2}A\sigma_c^2 \tag{8}$$

$$S.t. \begin{cases} E(r_c) = r_f + y[E(r_p) - r_f] \\ \\ \sigma_c^2 = y^2\sigma_p^2 \\ \\ 0 \leqslant y \leqslant 1 \end{cases} \tag{9}$$

From equations above,we can easily find the solution .

$$y^* = \frac{E(r_p) - r_f}{A \cdot \sigma_p^2} \tag{10}$$

We consider that the total amount cannot be greater than \$1000, and if y is greater than or equal to 1, then we consider that y takes 1.

- **Visual Analysis of Optimal Portfolio** We visualize the above optimal portfolio in the feasible set by using the relevant functions in matplotlib.pyplot in order to illustrate more intuitively the value area and rationality of the portfolio. We use the relevant function in matplotlib.pyplot to visualize the position of the above optimal portfolio in the feasible set, and the obtained image is shown in figure 7.

Figure 7: Modeling process

In the chart above, the pentagon marks the portfolio with the highest Sharpe ratio (i.e. risk-return equilibrium point), and the positive hexagon marks the minimum variance portfolio. We use the minimum variance portfolio as the boundary and divide the feasible set into the upper and The edge of the upper half is the effective boundary, which we represent by the purple curve. The image illustrates the feasible region of the portfolio and gives options that can reduce the risk while guaranteeing a certain rate of return, proving that the model can achieve the unity of risk and return for the client.

5.4.3 Dummy Variables for Gold Close Days

Due to market trading rules, gold cannot be traded on Saturdays, Sundays and holidays, i.e. it cannot be bought and sold, while bitcoin does not have restrictions on trading hours, so we need to consider whether it is a gold opening day as a dummy variable when building our model

- **Gold Trading Day** On a gold trading day, we need to combine gold and bitcoin to get the risk and return.And we can get the following functions. $\omega_g^{(T+1)} y^{(T+1)}$

- **Gold non-trading days** On a non-trading day for gold, we then consider $\omega_g^{(T)}$,$y^{(T)}$ to remain constant and ration only bitcoin and risk-free investments, so we can obtain the following equation.

$$\left\{\begin{array}{l} E\left(r_p\right)=E\left(r_b\right) \\ \sigma_p=\sigma_b \end{array}\right. \tag{11}$$

From the above equation we can easily calculate the value of y^*

So we can get the following system of equations.

$$\left\{\begin{array}{c} \omega_b^{(T+1)}=\left[1-y^*\right] \cdot\left[1-\omega_g^{(T)} \cdot y^{(T)}\right] \\ y^{(T+1)}=y^* \cdot\left[1-\omega_g^{(T)} \cdot y^{(T)}\right] \\ \omega_g^{(T+1)}=1-\omega_b^{(T+1)} \end{array}\right. \tag{12}$$

5.5 Daily Trading Strategy Model

According to 1.2 it has been derived that for a given handling fee $\alpha_{gold}=1\%, \alpha_{bitcoin}=2\%$ when the ideal buy to open a position and sell short position. The optimal portfolio w_g, w_b, y that combines maximum return and minimum risk is also given in 1.3. We define that if the buy-sell relationship satisfied by the reallocation of resources according to the optimal portfolio for gold and bitcoin at time T+1 is consistent with the ideal buy-sell relationship, and the total value of

assets at time T+1 after the exchange does not significantly shrink compared to time T (defining the maximum allowable shrinkage index β), then the transaction is considered to be possible at time T+1, and the post-trade gold, bitcoin The allocation relationship $[D^{T+1}, G^{T+1}, B^{T+1}]$ with the remaining dollars is exactly in line with the optimal portfolio.
Known: The distribution of gold, bitcoin and remaining dollars on day T$[D^{(T)}, G^{(T)}, B^{(T)}]$ Total assets at time T are:

$$Cap^{(T)} = D^{(T)} + G^{(T)} * price_g^{(T)} + B^{(T)} * price_b^{(T)}$$

Optimal risk portfolio for T+1 day$[w_g^{(T+1)}, w_b^{T+1}, t^{(T+1)}]$,The short-term forecasts of prices are$price_g^{(T+1)}, price_b^{(T+1)}, D^{(T)} * (1 + r_f)$The total value of assets not to be traded is

$$Cap_0^{(T+1)} = D^{(T)}(1 + r_f) + G^{(T)} * price_g^{(T+1)} + B^{(T)} * price_b^{(T+1)}$$

Then, at the moment T+1, the total amount $Cap_0^{(T+1)}$ that can be allocated according to the optimal risk portfolio is distributed in the portfolio distribution $[D^{T+1}, G^{T+1}, B^{T+1}]$ as shown in Eq.

$$\begin{cases} G^{(T+1)} = Cap_0^{(T+1)} * y^{(T+1)} * w_g^{(T+1)} / [price_g^{(T+1)} * (1 + \alpha_{gold})] \\ B^{(T+1)} = Cap_0^{(T+1)} * y^{(T+1)} * w_b^{(T+1)} / [price_b^{(T+1)} * (1 + \alpha_{bitcoin})] \\ D^{(T+1)} = Cap_0^{(T+1)} * [1 - y^{(T+1)}] \end{cases} \tag{13}$$

The total value of assets at the T+1 moment after the transaction is updated as follows:

$$Cap^{(T+1)} = D^{(T+1)} + G^{(T+1)} * price_g^{(T+1)} + B^{(T+1)} * price_b^{(T+1)}$$

The actual gold and bitcoin transformations are compared to the theoretical buy-sell scenario, requiring that the following are satisfied:

$$\begin{cases} GA(T,n) > (1 + \alpha_{gold}) * GA(T,m) \&\& G^{(T+1)} > G^{(T)} \\ GA(T,n) < (1 + \alpha_{gold}) * GA(T,m) \&\& G^{(T+1)} < G^{(T)} \\ GA(T,n) > (1 + \alpha_{bitcoin}) * GA(T,m) \&\& B^{(T+1)} > B^{(T)} \\ GA(T,n) < (1 + \alpha_{bitcoin}) * GA(T,m) \&\& B^{(T+1)} < B^{(T)} \end{cases} \tag{14}$$

We further define Maximum asset shrinkage index β ,to avoid to much shrink of total capital.

$$Cap^{(T+1)} > (1 - \beta) * Cap^{(T)}, \beta = 0.01 \tag{15}$$

Obtain the distribution of the portfolio after trading according to the optimal portfolio$[D^{T+1}, G^{T+1}, B^{T+1}]$ and total assets$Cap^{(T+1)}$

5.6 Analysis of the Results

Using the model we constructed, we derived the specific transactions and transaction amounts for $\$1000$ for this 5-year period from 2016 to 2021 as shown in the following table.

From the above table we can easily see that $1000\$$ reached $310,445.96\$$ at 9.6.2021 with a return of $30944.6\$$, so we can see that the daily-based portfolio trading strategy model we have constructed can effectively quantify the investment in gold and bitcoin, increasing the return while keeping the risk within a manageable range.

We then selected a portion of the total gold and bitcoin transaction records to focus our analysis. We recorded the shares of gold, bitcoin and cash for each transaction and calculated

Table 2: Key Transaction Log Sheet

Date	Gold			Bitcoin			Cash
	Price	Amount	Holdings	Price	Amount	Holdings	
2016.10.26	1270.5	-122.43	0.095	0	0	0	577.57
2016.12.20	1125.7	0.314	0.31	0	0	0	628.54
2017.01.16				830.5	-573.46	0.6766897	55.02
2017.06.12	1266.4	49.76	0.35				5.26
2017.06.14				2706	807.52	0.37	812.78
2017.07.10	1211.9	376.41	0.661				436.37
2017.07.17				1931.214	-427.63	0.5892	8.74
2017.09.08	1280.2	730.8	0.0843				739.54
2017.11.11				6569.22	-739.54	0.6995	0
2017.12.17				19279.9	7994.7	0.2764	7994.7
2018.08.17	1178.4	-473.6	0.4821				7521.1
2018.12.14				3278.37	7521.1	2.52468	0
2019.06.26				11766.4	12131.195	1.4726	12131.2
2019.09.03	1537.85	717.4	0.011				12848.595
2020.03.15				5166.26	12792.35	3.899	56.25
2021.01.09				40670.25	57502.38	2.456	57558.63
2021.02.01				33136.46	57558.63	4.158	0
2021.04.16				63252.63	257794.12	0	257794.12
2021.07.20				30815.94	257609.6	8.1924	184.5
2021.9.6				51769.06	310261.46	2.0769	310445.96

the corresponding percentages and plotted the percentage stacking bar chart. In addition to plotting the ratio of the three investment elements, we also plot the change in the total amount of capital over this time, and we can see that in the first 3 years of this 5-year period the investments are overall on the steady side, with lower returns on investments, while as the price of bitcoin surges in the later years, the returns on investments produce huge fluctuations and eventually stabilize at a higher level.

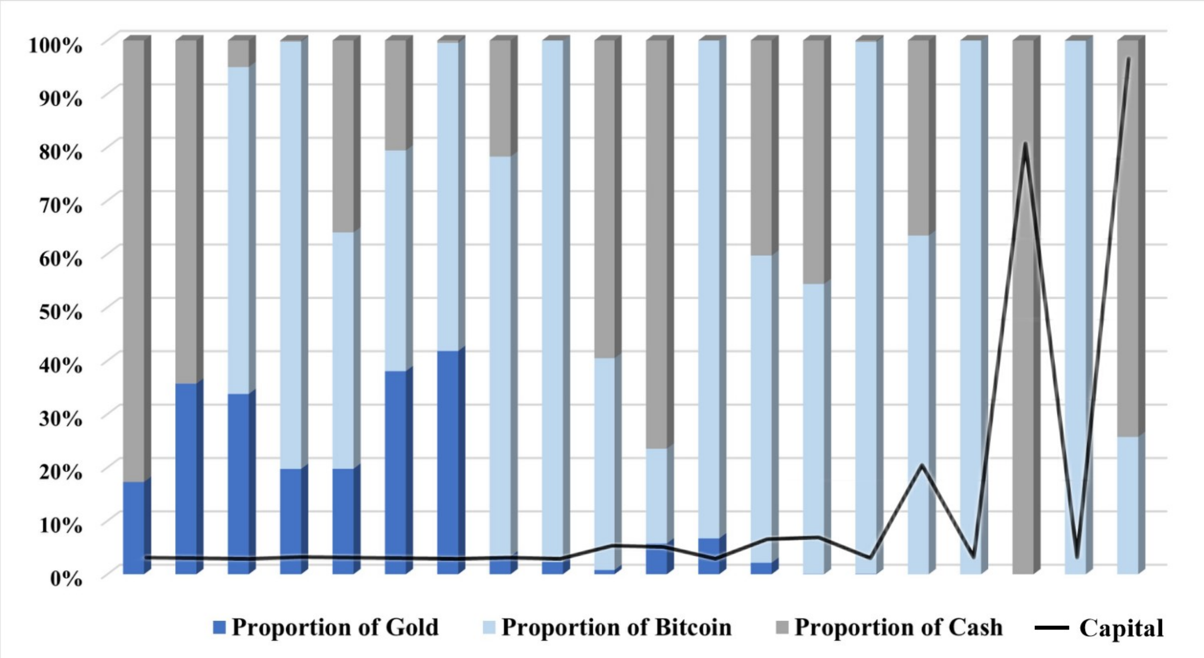


Figure 8: Cumulative histogram of trade profolio

6 Multidimensional Evaluation Model

The evaluation about the effect of the model is divided into two main parts as follows

- Considering the effect of the model itself, DTSM is consists of prediction model and optimization model. For lack of our knowledge, we mainly put our attention on the predicted performance.
- Backtesting of other common strategies in the market, evaluating and analyzing various economic indicators from expected return value, annualized profitability, annualized rate, Sharpe ratio, maximum retracement rate, etc.

6.1 Visualization of DSTM

Above all, to present the result of our prediction, we further visualize the data.

In addition to the line chart, in order to provide a clearer picture of the ratio of gold, bitcoin and cash at each trade, we have drawn pie charts at the trade nodes to show the size of the ratio of the three, and the size of the pie chart is a reflection of the total amount of capital traded, i.e. the larger the total amount of capital the larger the pie chart. Finally, to illustrate the details of each transaction, we use red stars to indicate buy operations and blue triangles to indicate sell operations. From the above chart it is easy to see that at each transaction, the ratio of gold, bitcoin and cash has a large difference, in the early stage the ratio of gold is higher and the ratio of bitcoin is lower, I guess because the value of gold is higher in the early stage, with higher stability and good investment prospects. And with the gradual rise and surge in value of bitcoin, the portfolio model is more inclined to invest with bitcoin, and you can see that in the later part of the portfolio almost no longer hold gold, but only bitcoin and cash.

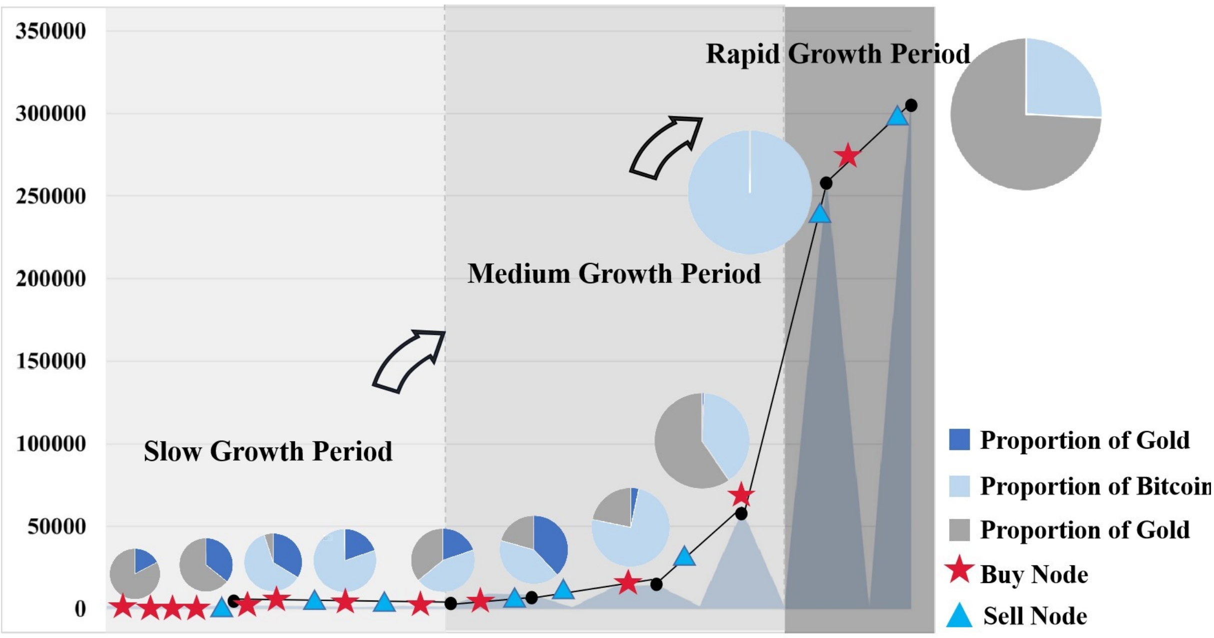


Figure 9: Detail portofio analysis of DTSM

In figure 9 above, we divided the 5 years into three main time periods based on the growth rate of the total capital, the first period has a low yield and a slow growth rate, but there are frequent buying and selling operations, the second part is the medium growth stage, with less frequent buying and selling operations compared to the first stage, while the total capital has a certain increase, but the yield is still at a low level. The third stage is the rapid growth stage, we can clearly see that the total amount of funds has surged, the number of transactions is very small, while the yield has increased significantly and maintained at an overall higher level.

6.2 Evaluation of LTSMIS

To better demonstrate that our model can predict more accurate value as well as change trend in the future, we use data prior to a certain date (training set end date) before 2/3/2021 to predict the value of bitcoin on 2/3/2021, in a bid to see how the mean absolute error between real data and predicted data changes with daily updates.

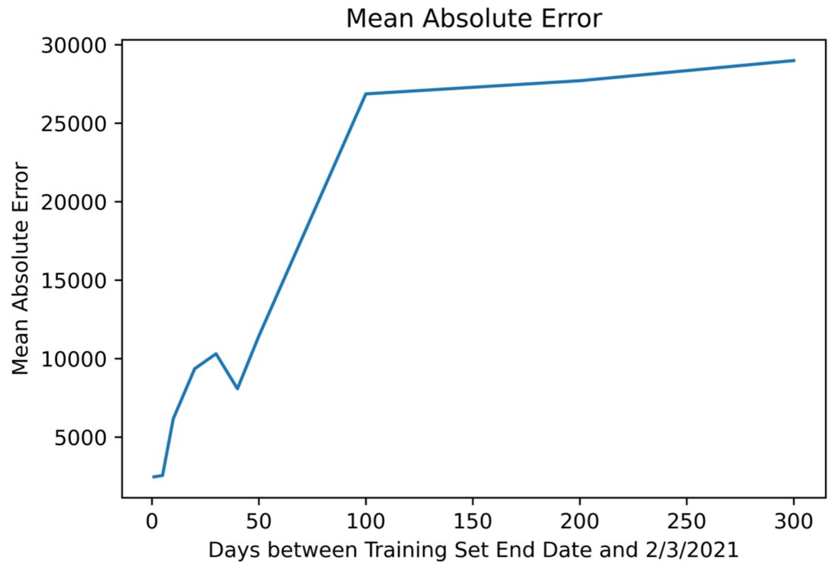


Figure 10: Mean absolute error between real data and predicted data

Shown in figure 10 above is the relationship between the number of days from training set end date to 2/3/2021 and mean absolute error of predicted value and real value of bitcoin. Since the result of neural network is highly related to parameter initialization, which are random variables following Gaussian distribution, we train our model for several times given different number of most recent days and use the average absolute error as the yardstick. The figure shows that with daily updates, our model is becoming more and more able to predict the change trend of future data more accurately.

6.3 Backtrader Backtest

6.3.1 The Core Idea of Backtrader

Backtrader is a library of python that consists of the following parts.

1. Data loading: loading the data of trading strategies into the backtesting framework
2. Trading Strategies: This module is the most complex part of the programming and requires designing trading decisions and deriving buy or sell signals
3. Backtesting framework settings: need to set (i) initial capital (ii) commission (iii) data feed (iv) trading strategy (v) trade position size
4. Run backtest: Run Cerebro backtest and print out all executed trades
5. Evaluate performance: Evaluate the backtesting results of trading strategies with graphs and indicators such as risk-return

6.3.2 Data Pre-processing and Backtest Strategies

Since the data given in the table is not comprehensive and needs to be further expanded and analyzed, we first make a reasonable expansion of the data related to gold and bitcoin.

- **Step1 :** We conduct basic processing first.Import the table data into DataFrame format in python and set the openinterest column and volume column to be 0 considering the incomplete data.
- **Step2 :** Extended data to the format that datetrader can process.
  - $Close^{(T)}$ Takes the latter day's open price as today's close price.
  - $Open^{(T)}$ Data given by problem.
  - $Low^{(T)}$ Suppose 0.99 times of open price to be the lowest price.
  - $High^{(T)}$ Suppose 1.01 times of open price to be the highest price.

  Here we listed part of gold data.

Table 3: The completion results(part of gold)

Close	Low	High	Open	Volume	Open interest	data time
1323.7	1311.4	1337.8	1324.6	0	0	9/12/16
1321.8	1310.4	1336.9	1323.7	0	0	9/13/16
1310.8	1308.5	1335	1321.8	0	0	9/14/16
...	...	...	...	...	...	...

- **Step3 :** ***Set parameters of backtesting strategy***
  1. Period of Average Trading Strategy
  2. Commission of trade(purchase or sell)
  3. Number of dataset
- **Step4 :** Visual analysis of backtesting.

6.3.3 DTSM Strategy vs. Single Trade Strategy

To better illustrate that the model's portfolio trading model has better results, we compare our portfolio trading strategy with trading gold alone and bitcoin alone, respectively. We exported the data of each of the three and compared them visually, and the resulting images are shown below.

Total capital	Gold amount
995.78	0.3329
949.805	0.8252
1050.769	0.4772
1033.076	0.8258
1129.168	0.4726
1074.328	0.7629
1042.639	0.8772
1175.64	0.6471
1303.87	0.4721
1260.54	0.8595
1353.612	0.6762
1421.95	0.5019
1336.71	0.8763
1780.739	0.19
1705.23	0.9892
1868.604	0.2745

Total capital	[D,G,B]		
577.57	0.065	0	0.31
628.54	0.198	0	0.35
55.02	0.296	0.537	0.04
5.26	0.197	0.800	0.01
812.78	0.213	0.479	0.38
436.37	0.486	0.528	0.26
8.74	0.656	0.905	0.01
739.54	0.054	1.219	0.35
0	0.084	3.576	0
7994.7	0.011	0.576	0.86
7521.1	0.065	0.200	0.86
12131.2	0.481	6.686	0
12848.59511	0.047	1.285	0.90
56.25	0.001	1.063	0.89
57558.63	0.010	13.01	0.03
0	0.000	1.681	0.96
257794.12	0.010	73.79	0
184.5	0.007	0	0.99
310445.96	0.009	125.2	0.09

Total capital	Bitcoin amount
988.49932	0.916
1293.288042	1.3517
3878.806346	0.6742
3175.47604	1.6365
8031.011108	1.0742
6056.859333	1.8831
36169.35286	0.4763
29672.62784	4.0714
41012.30147	2.1476
26973.90102	7.8132
92521.3127	4.1472
66576.17847	10.3853
106160.2493	10.9359
416305.0224	8.4715
348202.1552	11.4468
710316.4537	3.1265
602522.5767	18.8009
774309.1682	18.9401
930683.3365	4.3152

Final asset projections:

$ 1868.604	$ 310445.96	$ 930683.3365

Figure 11: Comparation among our DTSM, pure gold trade and pure bitcoin trade

As seen from the chart above, we compare main trade points of our DTSM with trade gold alone or trade bitcoin alone. We can find that gold is traded less frequently because gold has a trading day limit. In addition, we represent buying and selling in red and green respectively, which can clearly reflect the capital flow of each transaction. It can be seen that the trading strategy of pure gold is mainly based on selling, while in pure Bitcoin's trading strategy is mainly based on buying, and in the portfolio investment strategy we propose, buying and selling are more balanced. Finally, we compare the final returns of the three strategies. It can be seen that the minimum return of pure gold trading is only $1868.604$$, the return of all bitcoin transactions is $930683.3365$$, and the return of using the portfolio investment model is $310445.96$$. Although pure bitcoin investment has the highest return, the risk it bears is correspondingly high. The model we built can achieve a balance between buying and selling under the same high return, making good use of bitcoin and Gold's hedging properties reduce risk.

7 Transaction Costs Analysis

7.1 Sensitivity Analysis

In the construction and analysis of the previous model, we considered the commission for each transaction (purchase or sale) costs $\alpha\%$ of the amount traded. Assume $\alpha_{gold} = 1\%$ and $\alpha_{bitcoin} = 2\%$. There is no cost to hold an asset. However, in this question, we are going to test the stability and robustness of the model to investigate the effect of $\%$ on the final predicted return of the model when taking different values. We make adjustments to the commission of gold and bitcoin in $[0, 3]$, and find out that the total assets after implementing our strategy for

5 years decline monotonically by and large with increasing commission, and local maximum value exists.

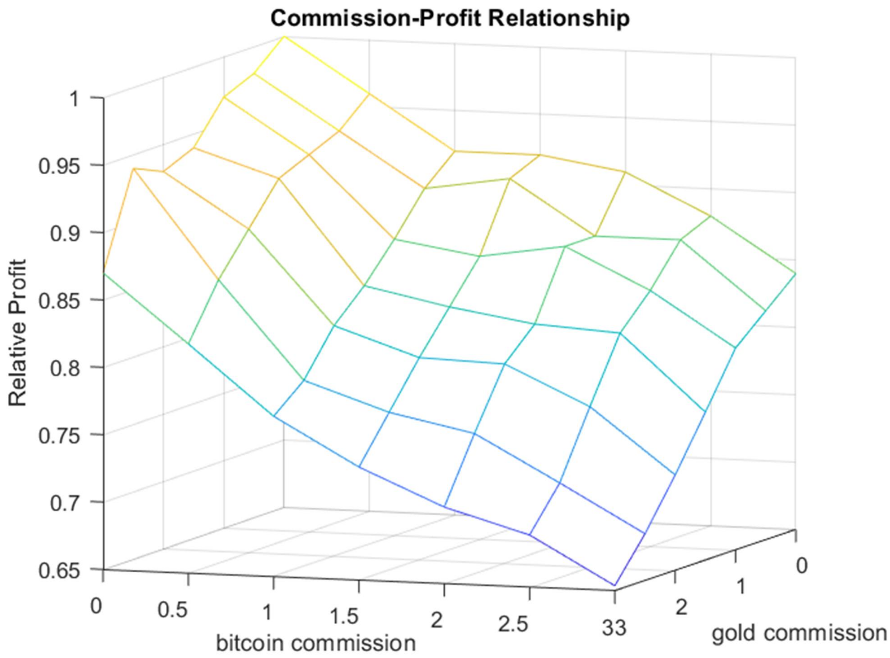


Figure 12: Commission-Profit relationship

7.2 Action Mechanism

We consider the affects of transaction costs on the strategy and results in the following two aspects.

- **Transaction Frequency :** In typical Average Moving Strategy, we calculate long-time average(L) and short-time(S) average $GM(T+1, n), GM(T+1, m)$, and then judge whether to purchase or sell products through whether they cross each other. We could say credibly that when the "L" line surpass "S" line, it means that we could profit in the future if we take shows now probably. It can also be analogized when "S" line surpass "L" line.
  When we take transaction costs into account, we must modify the principle above. In our model, we simply imple transaction costs by compute $1 + \alpha$ times of short-time average, since we consider it as current short-term costs.
  It can be seen obviously that transaction costs will significantly increase our transaction threshold and decrease our times of transaction, which consistents with the fact that the more the trade risk is, the less the trade frequency is.

- **Transaction Profit :** There is no doubt that transaction costs will have a strong and direct relationship with our total capital besides transaction frequency. We take some extreme situations as example. We all know that we must sell or purchase a lot of items when come across sharp growth or decline, sometimes even sell out use out cash. The transaction costs will noticeably impact our transaction profit, which could lead to approximate monotonic variation between total capital and transaction costs.

8 Memorandum

9 Sensitivity Analysis

To test the sensitivity of our model, we change the number of most recent days that are used to predict future value. We use data before 7/2/2019 to predict the value of bitcoin on 7/3/2019 and 4/27/2020 respectively, and records the absolute error between real and predicted value. Since the result of neural network is highly related to parameter initialization, which are random variables following Gaussian distribution, we train our model for several times given different number of most recent days and use the average absolute error as the yardstick. The relationship between absolute error and the number of most recent days are shown in figures above.

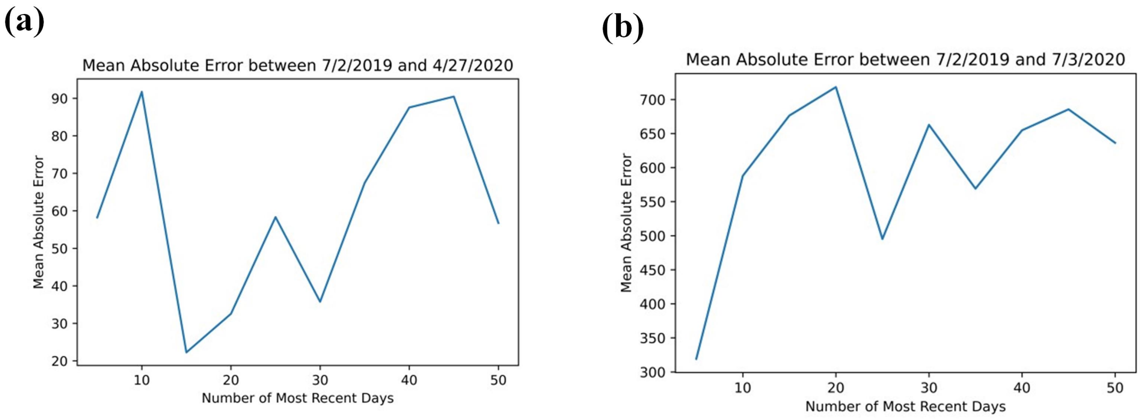


Figure 13: The relationship between absolute error and the number of most recent days

As can be seen, our model is highly sensitive to the number of most recent days.

10 Strength and Weakness

- **High Creativity :** Use LSTM as long-term forecast in MATS; deal with the condition that gold only be traded on some days by Conservation of Total Captal;introduce commission into MATS.
- **High Generalization :** Use backtrader Package to execute back test, which is easy to extend to other trading strategy.
- **Strong Robustness :** Our model is relatively robust to various parameters such as α_{gold}, $\alpha_{bitcoin}$ according to sensitive analysis.
- **High Forecast and Strategy Complexity :** Using MATS as Circular trading strategy and MOP as daily trade point add high time and programming complexity to our model.
- **High Forecast and Strategy Complexity :** Using MATS as Circular trading strategy and MOP as daily trade point add high time and programming complexity to our model.

References

[1] Zhou Zuwen. CNN-based quantitative trading model for financial markets [D]. Dalian University of Technology, 2020. DOI:10.26991/d.cnki.gdllu.2020.001173.

[2] Sun, Lipo. An empirical study of Makowitz's portfolio theory based on Python[J]. Times Finance,2020(25):46-47+50.

[3] Zhang Heqing. Research on Markowitz portfolio model with mean and variance changes[D]. Harbin Institute of Technology,2015.

[4] Yu Shule. Neural network modeling and analysis for stock price prediction[D]. Hangzhou University of Electronic Science and Technology,2016.

[5] Multi-factor quantitative stock selection strategy under machine learning algorithm[J]. Journal of Jilin College of Commerce and Industry,2021,37(06):90-97.DOI:10.19520/j.cnki.issn1674-3288.2021.06.014.

[6] Ye W.Y.,Sun L.P.,Miao B.Q.. A dynamic cointegration study of gold and bitcoin–based on a semiparametric MIDAS quantile point regression model[J]. Systems Science and Mathematics,2020,40(07):1270-1285.