Problem Chosen	2022	Team Control Number
C	MCM/ICM Summary Sheet	2200237

Make a bundle in the Wall Street: AI assistant helps everyone become Buffett in investment.

Summary

During our childhood, investing in gold was seen as an extremely profitable investment driving people to spend every penny in their pocket. As the economy develops, new assets like bitcoin managed to make a great coup in the stock market and become the hottest star. The huge price fluctuation attracted many people holding dreams of getting rich overnight. In this paper, we established a model maximizing 5-year investment profits in both gold and Bitcoin. We hope to provide dear traders with strategies which will enable them to foresee the potential transaction points and make profits in the ups and downs. Following the logic sequence, we divided our model into two parts, the **Unit Price Prediction Model** and **Investment Decision Model**, denoted model 1 and model 2 respectively.

For Model 1, we construct a **Price Prediction Model based on Long Short Term Memory (LSTM) and Random Walk**. Inspired by reinforcement learning, we utilized **decay factor** to build our training set which allow the network to pay more attention to the recent prices. To make our model more robust and precise, we use the **Random Walk Model** on the basis of the principles o f **statistics analysis**, which performs excellent on the provided dataset. By a comprehensive consideration on the terms of strengths and weaknesses of two models, **Ensemble Learning** is put into use to get a more accurate prediction.

For Model 2, due to the limited time length of price prediction, the **Greedy Algorithm** is determined to be the main algorithm to obtain maximum profit, and the corresponding equation is established to **amplify profit and reduce loss**. Then, based on the forecast data of unit price in the next few days generated by Random Walk and LSTM methods in Model 1, today's buying and selling decision would be settled. After 5 years, the total value of the final assets held was **around 7 million**, as shown in Table 6.

In addition, we proved the effectiveness of our model by comparing it with **two existing statistic indexes as our baseline**. By evaluating the performance of the models, we find that our model doesn't only have a better profit income after a certain time period, but also maintain a high consistency with the strategies given by baselines. This indicates that our model could make most of correct decisions while making improvements for an even better profit gain.

Finally, we analyzed the sensitivity properties of our model and discussed the influence of transaction cost changes and the effective coping strategies. Regardless of other changes, traders should reduce the number of transactions, wait and see, and rely more on gold for profits when bitcoin transaction costs rise. Finally, with this strategy, we were able to increase our principal more than **30-fold** even when the transaction cost of Bitcoin doubled. As for gold, changes in transaction costs have little discernible impact on traders' decisions.

Keywords: Long Short Term Memory; Random Walk; decay factor; Greedy Algorithm

Contents

1 Introduction

1.1 Problem Background

As the international financial market develops, the scale of global stock market has grown to a massive scale and become the most important part in social economic activity. However, the complex market can be influenced by a dramatic number of factors and fluctuate erratically which make it a challenging work to predict. In recent years, a variety of methods have been proposed and implemented for forecasting stock market trends, but it remains unclear whether these models could predict new assets with new properties such as Bitcoin. Under the practical consideration of the limited spendable resources, how to effectively allocate the resources in our hands to enter the commodity trading market in order to obtain maximum benefits in the future has become a problem that we urgently need to consider.

1.2 Restatement of the Problem

In this problem we are given 2 market price datasets on gold and bitcoin in the form of time series. We were asked to develop a model, come up with strategies to gain maximum profit and conduct relative analysis on our model. We should only use the data up to a certain date to solve the following problems:

- Develop a model using price data from beginning till the day to determine present action.
- Proof that our model provides the best strategy which can grantee a largest profit with zero knowledge of future price.
- Discuss in what degree does transaction cost effect our model decision.
- Write a letter to communicate with traders on our model, advise them with decision procedures for their trading actions.

1.3 Our Work

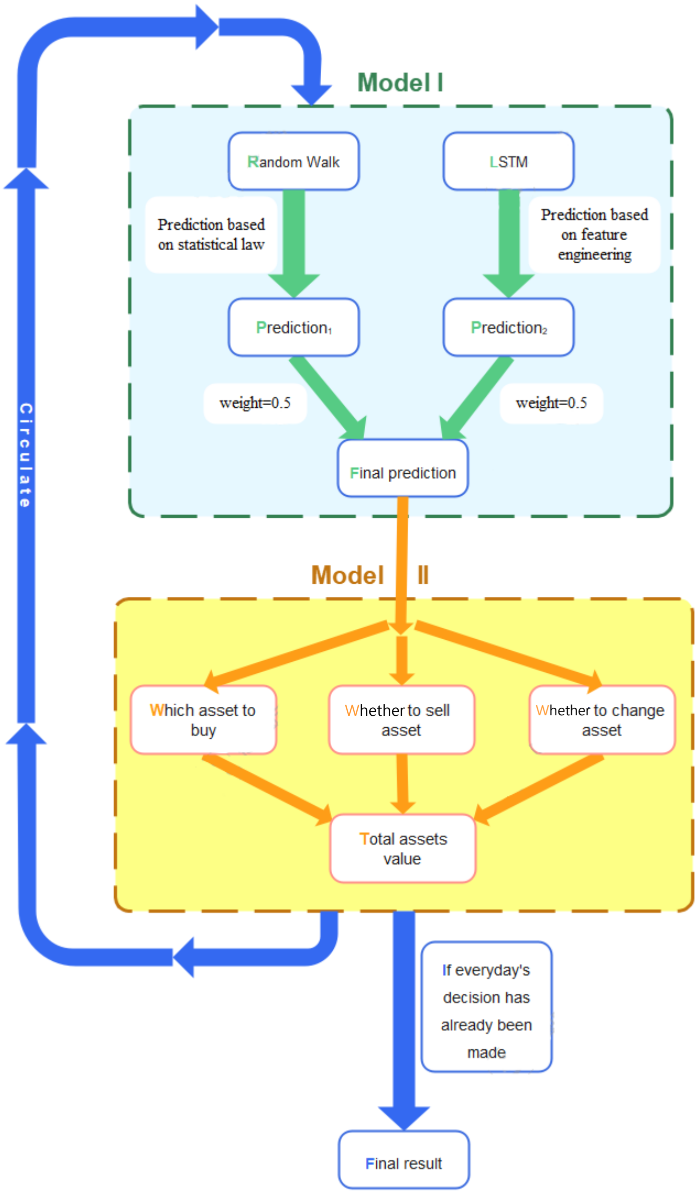


Figure 1: Model Overview

2 Assumptions and Justifications

To simplify our model and eliminate the complexity, we make the following main assumptions and justify their usage.

- **Assumption 1: Money cannot increase or decrease if they are not invested in Gold /Bitcoin market.**

  Justification: *By setting up this assumption, we ignored the complex real-world situation in which keeping money in bank alone can gain us a certain percentage of deposit interest and value of our money can decrease due to the condition like inflation.*
- **Assumption 2: We clarify that the profit of certain assets is calculated from the**

next day they are bought till the day they are sold.
Justification: *In another word, assets do not create profit during the day of purchasing.*

- **Assumption 3: During our modeling and calculation, we assume that there is no minimum quantity limitation hance the exchange quantity can be however small.**
  Justification: The minimum quantity in a real-world exchange is 10^{12} and the most common precision limitation is 0.01. In our calculation, we consider the minimum exchange quantity is small enough for us not to consider it. In fact, the exchange quantity precision goes along with our computer's calculation precision.
- **Assumption 4: The data provided in the topic is sufficient and reliable.**
  Justification: *Because we can only use the data provided in the topic without referring to other data, only the credibility of the data information can ensure the accuracy of the model.*

3 Notations

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

Table 1: Notations used in this paper

Symbol	Description	Unit
P_c	Current daily price of the holding asset	\$
TAV	Total assets value	\$
T_n	Threshold value used to enhance robustness	\$
Q	The quantity of assets held	—
$\alpha\%$	Transaction cost	—
I	The potential interest of today's Decision	\$
R_{atio}	Threshold ratio used to enhance robustness	—

4 Data Preprocessing

4.1 Data Cleaning

Examining the data files provided, including LBMA-GOLD.csv and BCHAIN-MKPRU.csv, we find that the dataset concerning gold contains ten missing values and the trade of gold stops at weekends. Therefore, we drop the rows corresponding to the missing values. For further visualization, the format of the original time data has been modified.

4.2 Min-Max Normalization

We perform a linear transformation on the original data so that the result maps to $[0,1]$, as it is essential to eliminate the dimensional effect between prices. The conversion function is as follows:

$$y_i = \frac{x_i - \min\limits_{1 \leqslant j \leqslant n} \{x_j\}}{\max\limits_{1 \leqslant j \leqslant n} \{x_j\} - \min\limits_{1 \leqslant j \leqslant n} \{x_j\}} \tag{1}$$

where $x_1, x_2......x_n$ represents the value of gold or bitcoin in the data files.

However, the prediction curves of the following models are shown after denormalization

4.3 Date Down-sampling

To reduce losses from commissions for each transaction, which probably exert great impact on the result, we prefer to trade less frequently. In that case, several local maxima and minima in the data curve have little influence on our decision-making so we choose to do data down-sampling. It can smooth the curve that is more conductive to global analysis.

5 Price Prediction Model Based on Long Short Term Memory and Random Walk

5.1 Single-Layer Long Short Term Memory Prediction Model

5.1.1 Model Development

A time series is a sequence of data points that occur in successive order over some period of time, conducting detailed research on the trend of the ongoing time series and forecast possible future movements has long been a field where massive attention is drawn. Because of the wide range of time series existence and possible profit gained from prediction, numerous methods have been invented and implemented which gained a satisfying outcome. Categorized by using machine learning algorithms or not, the current means of forecasting can be divided into statistical method and machine learning (ML) method. Before the existence of ML, this work is mainly done by statistical analysis. However, conventional techniques for time series prediction have their limitations dealing with big data and could have a hard time dealing with series patterns of new asserts like Bitcoin price series. The representative machine learning models include support vector machine (SVM)[1], Shallow Neural Networks like simple Neural Network (NN), Multi-Layer Perceptron (MLP) and error Back Propagation (BP). There are also shallow NNs having recurrent structures like Recurrent Neural Networks (RNNs) and ones having deep convolutional structures like Convolutional Neural Network (CNN). All of these models have shown their shiny advantages in some applications, but in this specific task predicting market prices, studies have shown the superiority of LSTM in fully utilize the data in a long time period.

Long Short-Term Memory (LSTM) model is a novel variant of Recurrent Neural Networks(RNN). Unlike traditional RNN, LSTM is suited for classifying and predicting long term time series data by introducing different control gates. A typical LSTM cell contains input gate, forget gate and output gate, these gates allow LSTM cells to processes data as it propagates forward[2]. For a recurrent network, it suffers a lot from short-term memory. Specifically speaking on this model, having RNN as a prediction model could let us easily waste the information further from the present while LSTM could have more control of what to forget[3]. In the structural design of our network, we surprisingly found that one layer of LSTM network is more than enough for our prediction to work, blindly adding more layers have no contribution to the precision of our task but a waste of computing power.

5.1.2 Training set generation

In this section, we applied sliding window method to create our training set.

In single step prediction, we demand our model to predict trading price one day in the future at a time, so each training sample is composed of a certain length k series and one number label.

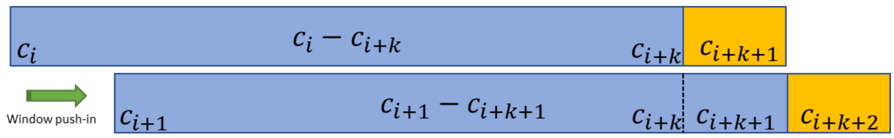
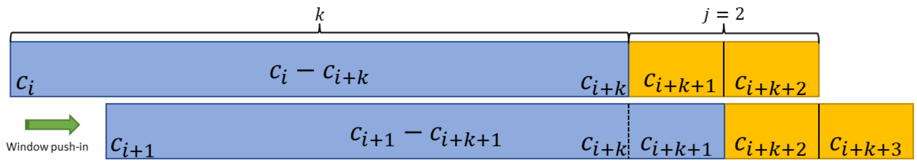


Figure 2: Single-step training set generation

each training sample is shown in blue and its label is shown in yellow

Similarly, in dual step prediction, the label length is changed to 2 points.



For example, in the scenario that on the 547^{th} day (30% of total days), we will have a training set which shape is $(530, 15, 1)$, shown in the form of (*samples, window-length, features*).

5.1.3 Symbol definition

The definitions of symbol in our model are shown as follow:

Table 2: Symbol definitions

Symbol	Definition
k	Length of input window
$x \in R^k$	Input vector for LSTM network
$h \in R^{128}$	Hidden state, carrying information flowing between the cells
$C \in R^{128}$	Cell state, carrying information of last cell and being the input of
$i \in R^{128}$	Information left, a part of input gate
$f \in R^{128}$	Product of forget gate, indicating information preserved in
$o \in R^{128}$	Product of output gate
W	Weight matrix the gates
b	Bias of certain gate
sig	The Sigmoid function
$tanh$	The hyperbolic tangent function

5.1.4 The Construction of LSTM

An ordinary LSTM cell computes data following the order of input gate, forget gate and output gate.

- First, the value of input gate i_t can be calculated:

$$i_t = \sigma \cdot (W_i \cdot [h_{t-1},\ x_t] + b_i) \tag{2}$$

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1},\ x_t] + b_c) \tag{3}$$

After multiplying i_t and C_t, we now have the output of input gate.

- The calculation of the forget gate involves h_{t-1} and x_t where:

$$f_t = \sigma(W_f \cdot [h_{t-1},\ x_t] + b_f) \tag{4}$$

After calculation above, the current state of the cell is determined in subsequent:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{5}$$

- The output gate value o_t and its output h_t are:

$$o_t = \sigma(W_o \cdot [h_{t-1},\ x_t] + b_o) \tag{6}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{7}$$

A better illustration of LSTM cell structure is shown in Figure 3:

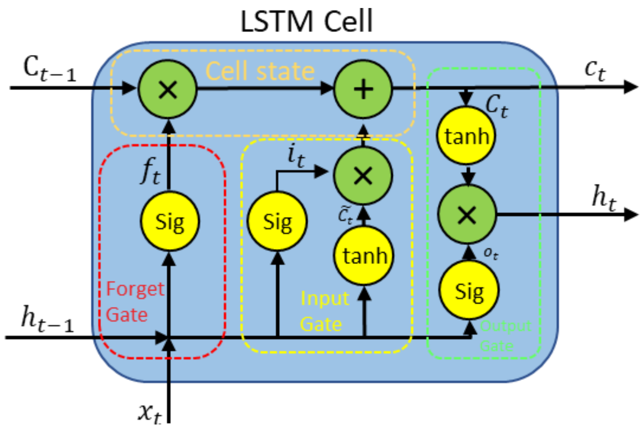


Figure 3: LSTM cell structure

For better understanding, we show the network structure of the model:

Table 3: The network structure of the model

Layer	Output Shape	Param number
Lstm_1	$(None, 200)$	161600
Droupout_1	$(None, 200)$	0
Dense_1	$(None, 1)$	201

Total params:161801 (100% trainable)

In the end, a Relu activation function is implemented to create its output.

5.1.5 Add Decay Factor

Inspired by the discount factor in reinforcement learning, which essentially determines how much the reinforcement learning agents cares about rewards in the distant future relative to those in the immediate future, we put forward the new concept of decay factor in this model.

We believe that the data closer to the day of decision-making is more significant, while the data information from a long time ago has less influence on the future. Thus, when generating the training set, we multiply $RealPrice_{n-1}, RealPrice_{n-2}, ...RealPrice_1$ by the coefficient of $1, \gamma, \gamma^2, ..., \gamma^{n-2}$ respectively, where $RealPrice_i$ represents the actual price of bitcoin or gold on Day *i*.

The values of parameter we test are shown below:

$$\gamma = [0.8, 0.9, 0.99, 0.999, 0.999] \tag{8}$$

During experiments, we discover that when $\gamma = 0.9999$, the predictions are of great accuracy.

5.1.6 Relevant Process Details

- **Normalization**:

  In order to preserve consistency between different datasets and guarantee a fast convergence, the data is pre-normalized just as the data preprocessing mentioned.

- **Training**:

  The training process contains 20 epochs and 10 batches in each epoch. The model is equipped with Adam optimizer.[4]

- **Price forecast**:

  To predict future price especially a rather long time in future, we introduced two methods, they are one-dimension-muti-step prediction and circular prediction.

  In multi-step prediction, we increased the label length to expected prediction length in training phrase.

  After the model training, we take one closest sample and make **one prediction** of *j* days:

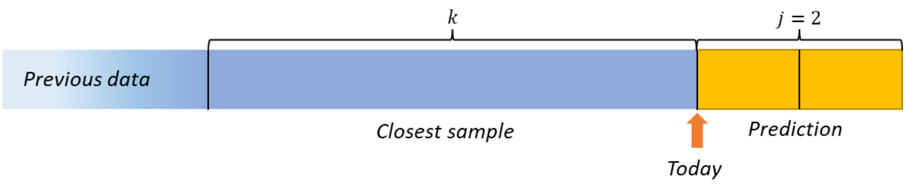


Figure 4: Multi-step Prediction

  This method provides a precise prediction but can only make short predictions ($j \leq 5$). Also the window length k needs to increase along with the prediction length j, practically the let $k \geq 5j$.

  In circular predictions, we initially use closest sample and make **multiple predic-**

tions. In each prediction, we predict only one data and push data in the sample afterwards. In this method can we theoretically make unlimited predictions but the error of each row accumulates and eventually lead the predictions to an unusable failure.

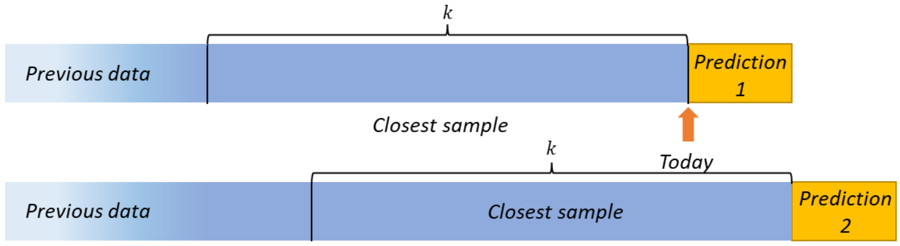


Figure 5: Circular Prediction

After a thorough consideration of the pros and cons in each way, we eventually utilize.

5.1.7 Result Presentation

The following experimental curve seems to perform well. The specific deviation will be compared and analyzed in the model evaluation.

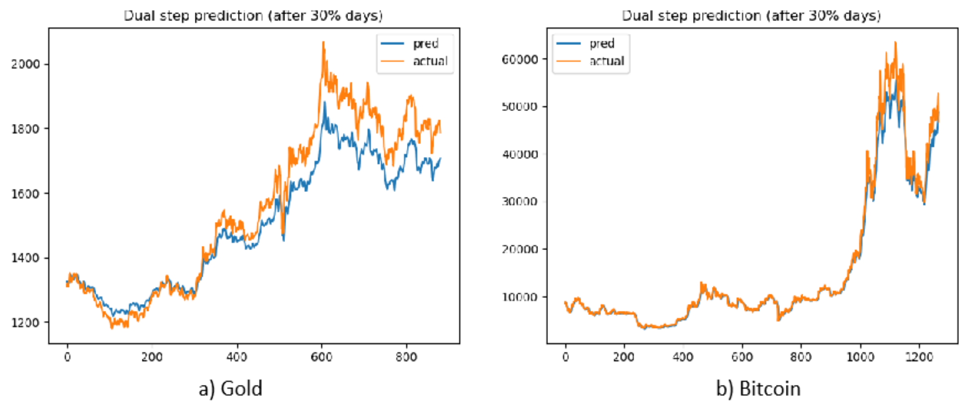


Figure 6: LSTM Prediction

5.2 Random walk in market prices

5.2.1 random walk theory introduction

Random walk theory, refers to the randomness and unpredictability of the price fluctuations of bitcoin and gold. The random walk theory points out that most people in the trading market are acquainted with analysis strategies, and the relevant information flowing into the market is public, so the current trading market prices have already represented the views of a great many shrewd people and are reasonable to a certain extent. The prices also fluctuate around the intrinsic value, which indicates the normal functioning and effectiveness of the market.

The theory suggests a number of reasons that could account for the fluctuation. For instance, news involving politics and economics stochastically enter into the markets, resulting in the reappraisal of Bitcoin and Gold valued by experts, who put forward transaction strategies based on alterations. However, this kind of stochastic information is unlikely to be foreseen. In another words, asset prices do not possess the memory system. It has no direction at all, wandering randomly, rising and falling. Another key point brought up is that market prices follow the law of normal distribution. Most asset prices rise and fall within a narrow range, while only very few can be sent into the stratosphere or plummet less than 100%. As the random walk is well known to investors, many economic workers have verified it and received positive feedbacks just as we expected.

5.2.2 Explanations for model usage

There are some reasons for introducing the concept of random walk:

- This theory has a strong economic basis and its validity and authenticity has been proved by many experts.
- This theory has a strong economic basis and its validity and authenticity has been proved by many experts.
- In the early stage, the Long Short Term Memory model cannot train a credible model due to the small data samples obtained, but the Random Walk method can have an outstanding performance, which is able to predict the price in the short term based on very little data.
- The combined application of the Long Short Term Memory method and the Random Walk method increases the reliability of the prediction curve.

5.2.3 The Construction of The Model

When the Random Walk is represented by a mathematical formula, it can be written as:

$$PredPrice_n = RealPrice_{n-1} \cdot \epsilon, \ \epsilon \sim N(\mu, \sigma) \tag{9}$$

Where $PredPrice_n$ refers to the predicted price of bitcoin or gold on Day_n, $RealPrice_{n-1}$ means the actual price of bitcoin or gold on Day_{n-1}, the coefficient ϵ obeys Gaussian distribution that μ serves as the mean value and σ serves as the standard deviation.

μ and σ are calculated from historical data. Suppose we already obtain the data information from Day_1 to Day_{n-1}, denoted as $RealPrice_1$, $RealPrice_2$,...... ,$RealPrice_{n-1}$
We define that:

$$Daydiff_K = (RealPrice_{k+1} - RealPrice_k) \ / \ RealPrice_k, \ k = 1, 2, ..., \ n-2 \tag{10}$$

which is an evaluation indicator to measure the daily fluctuation. From computation mentioned above, we can conclude that:

$$\mu = E(Daydiff_i), \ i = 1, \ 2, ..., \ n-2 \tag{10}$$

$$\sigma = std(Daydiff_i), \ i = 1, \ 2, ..., \ n-2 \tag{11}$$

where $E(x_i)$ represents the mathematical expectation of x_i and $std(x_i)$ represents the standard deviation of x_i.

Objectively speaking, the longer we predict, the broader our horizon, the better for our decision-making. But meanwhile, applying the information of data before a certain day to prognosticate the future gold or bitcoin values of too many days, will inevitably lead to inaccuracy in the later period. This is confirmed in our experiments as well.

Results and Analysis

We set the parameter $precdict_{day}$ to describe how many days we predict at a time. The values of parameter we test are shown below:

$$predict_{day} = [1, 2, 3, 5, 10, 20, 60] \tag{12}$$

which is the same as the Long Short Term Memory model.

We also obtain the same result as the Long Short Term Memory model. When $precdict_{day}$ is greater than 2, the result of Random Walk model estimation becomes unreliable. It depends on the principle of random walk to great extent. The selection of ϵ is inherently stochastic and the fluctuation as well as deviation of prediction will be aggravated after several iterations.

Combined with our decision model for comprehensive consideration, two-day prediction is recommended. Following are the results:

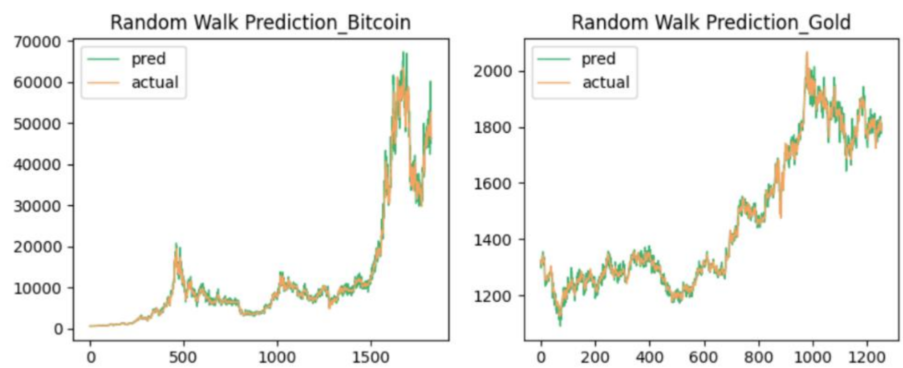


Figure 7: Random Walk Prediction

The experimental effect seems to be good. The specific deviation will be compared and analyzed in the model evaluation.

5.3 Ensemble Learning of Two Models

5.3.1 Reasons For Using Ensemble Learning

When we analyze the various metrics of the two models, which are shown in the table below, it is easy to judge that the indicators of random walk perform better, especially in the prediction of Gold.

However, the randomness of the random walk method seriously affects the stability of subsequent decisions, which probably result in predictions with large variance. Its theoretical basis is far less than Long Short Term Memory model. Therefore, Random walk is unreliable to some extent, as samples are not sufficient.

Table 4: Evaluation Indicators of Bitcoin

Model	R^2	MSE	MAE	RMSE	MAPE
LSTM	0.9918795	0.0004233	0.0112885	0.0205736	0.0439869
Random Walk	0.9903471	0.0004868	0.0112508	0.0220626	0.0538036

Table 5: Evaluation Indicators of Gold

Model	R^2	MSE	MAE	RMSE	MAPE
LSTM	0.7082408	0.0042472	0.0525663	0.0651703	0.0651212
Random Walk	0.9921163	0.0000001	0.0002642	0.0003494	0.0113165

What's more, at the beginning, we did some model comparisons, and the results proved that among a great many algorithms for extracting features, LSTM performed perfect.

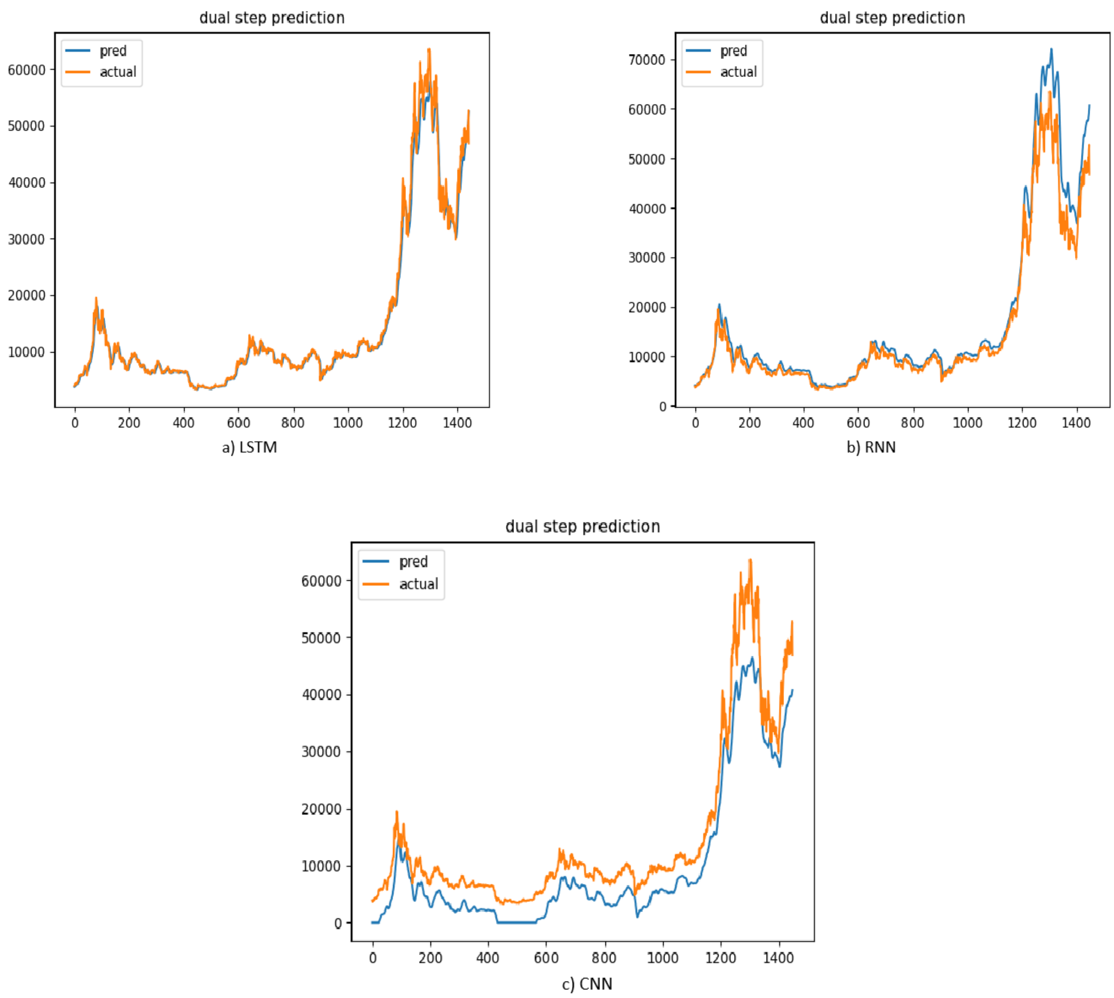


Figure 8: The prediction result of three different networks

Experiment conducted using bitcoin data and training set : test set = 2:8

Taking all above factors into consideration, we propose to use ensemble learning.

5.3.2 The Combination Strategy of Ensemble Learning

For numerical prediction problems, the method of average has been proved to be an excellent method to improve model performance. Additionally, as we are not able to get the real label of whether a transaction is made on a certain day, we fail to train a classifier.

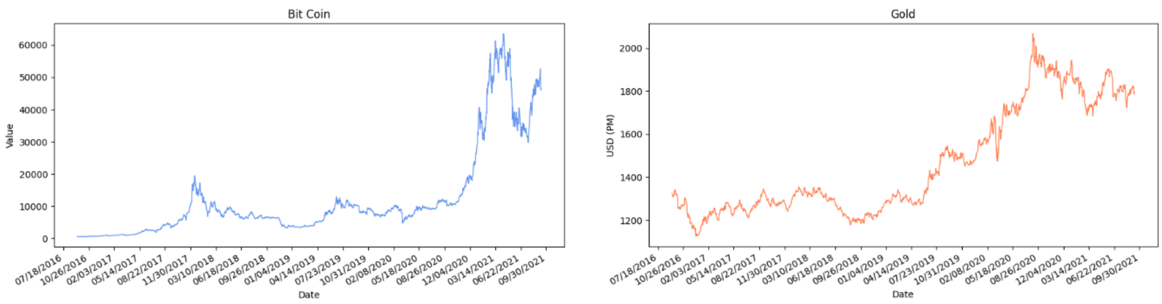
Ultimately, we make a balanced decision as follow:

$$Pred_{final} = 0.5 \cdot Pred_{LSTM} + 0.5 \cdot Pred_{rw} \tag{13}$$

6 Model II: Gold-BTC Portfolio Investment

6.1 Data Description

Since the amount of data is large and not very intuitive, we directly visualize the data for display, which creates great convenience when evaluating this model.



(a) Bit Coin Daily Price *(b) Gold Daily Price*

Figure 9: Value trend

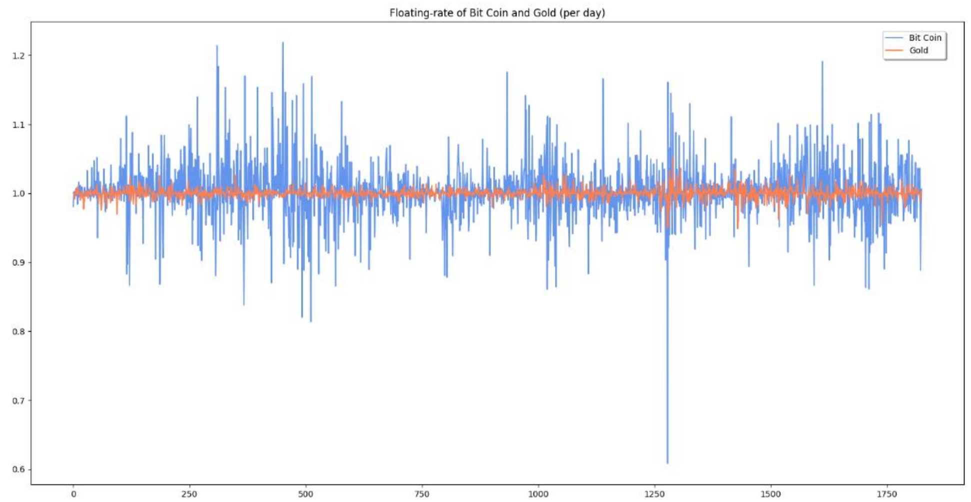


Figure 10: Floating-rate of Bitcoin and Gold (per day)

Bitcoin's fluctuations are obviously much more dramatic than Gold's, which could mean that most of our purchasing and selling will be on Bitcoin, with Gold playing a supporting role.

6.2 The Establishment of Model II

We will start with investing in one asset, for it's difficult to directly model investing in two. The assets in hand for each day will be recorded as:

$$state\,[Cash\ ,\ Bitcoin\ ,\ Gold] \tag{14}$$

6.2.1 Consider only one asset

At the beginning of day one, we are faced with the situation of choosing the right purchasing point. In order to maximize the profit, we always want to buy at the minimum price. So, in the case that the predicted value of Model I may not be accurate, in order to find the right min-price, a threshold is set to make sure the decision-making progress doesn't miss it, if the following formula is satisfied, today is considered to be the minimum point:

$$|\ P_c - P_{min}\ | < \ T \tag{15}$$

P_{min} is the minimum of the predicted value for two days and three days history including today

After confirming that today is the minimum, we should not blindly buy, for the profits brought later may not be able to cover the transaction cost, so we can only buy when the potential profits exceed a certain ratio of the cost in the foreseeable future:

$$(P_{maxf} - P_c) \cdot Q \cdot R_{atio} > (P_{\max f} + P_c) \cdot Q \cdot \alpha\% \tag{16}$$

P_{maxf} is the maximum of the predicted future days

Algorithm 1: The process of deciding whether to buy or not (consider only Bitcoin)

Input: $state\,[C,B,G]$, P_c, P_{maxf}, T,α
Output: $state\,[C,B,G]$
if *B=0* **and** $|\,P_c - P_{min}\,| < T$ **:**
  if $(P_{maxf} - P_C) \cdot Q \cdot R_{atio} > (P_{maxf} + P_c) \cdot Q \cdot \alpha\%$ **:**
    Convert all *Cash* into *Bitcoin*

When it comes to selling assets, we need to find the right max-price point to maximize profit, the method is identical with finding the minimum point:

$$|\ P_c - P_{\max}\ | < \ T \tag{17}$$

After confirming that today is the maximum, we should decision whether to sell it, there are several conditions:

- $P_{maxf} < P_c$:

  Decision: If you hold it, it will depreciate. So, sell it today.
- $P_{maxf} > P_c$ **and** the date of P_{maxf} is earlier than the date of $P_{\min f}$:

  Decision: There is still room to rise, continue to hold.
- $P_{maxf} > P_c$ **and** the date of P_{maxf} is later than the date of $P_{\min f}$:

Decision: In order to prevent the cost loss caused by frequent blind buying and selling, we consider two buying and selling ways below in the foreseeable span, calculate and compare the total asset value (TAV) after different decisions at the P_{maxf} point, then finally choose the operation method with higher asset value.

Case one: sell today and buy at the P_{minf} point:

$$TAV_1 = \frac{P_c \cdot Q \cdot (1-\alpha\%)}{P_{minf} \cdot (1+\alpha\%)} \cdot P_{maxf} \tag{18}$$

Case two: hold today until the P_{maxf} point:

$$TAV_2 = P_{maxf} \cdot Q \tag{19}$$

Algorithm 2: The process of migration of fish with consider of randomness

---

Input: $state[C,B,G]$, P_c, P_{maxf}, T_1, T_2, α
Output: $state[C,B,G]$
if $B \neq 0$ **and** $|P_c - P_{\max}| < T_1$**:**
  if $P_{maxf} < P_c$**:**
    Convert all *Bitcoin* into *Cash*
  else:
    if the date of P_{maxf} is **earlier than** the date of $P_{\min f}$:
      Continue to hold
    else:
      Calculate TAV_1 according to equation (5)
      Calculate TAV_2 according to equation (6)
      if $TAV_1 - TAV_2 > T_2$
        Convert all *Bitcoin* into *Cash*
      else:
        Continue to hold
else:
  No selling
end

6.2.2 Consider two assets

In the traditional investment model, investing by combining multiple assets is a way to avoid risks on the premise of ensuring a higher rate of return. Here, we believe that investors believe enough in their predictions to avoid risk through portfolio investment, so in order to maximize the expected return, investors will take an aggressive approach—holding a single asset for investment. Therefore, we only need to consider the following two cases:

- There is cash, one asset shows that there's a good chance of making a profit if you buy it today, but another asset may also be lucrative if you buy it a few days later. Should you trade today or keep the cash for future?

Decision: Based on the prediction of Model I, we calculate the potential profit of the two assets:

$$I = (P_{maxf} - P_c) \cdot \frac{C_{ash}}{P_c \cdot (1 + \alpha\%)} \tag{20}$$

$$I' = (P'_{maxf} - P'_{minf}) \cdot \frac{C_{ash}}{P'_{minf} \cdot (1 + \alpha'\%)} \tag{21}$$

Again, for possible forecasting errors, a threshold is set to prevent the miss of profit opportunities by holding on to cash:

$$I' - I > T_2 \tag{22}$$

- Holding gold or bitcoin without cash, now, the other asset shows that there's a good chance of making a profit if you buy it today, should you replace the asset that you own or hold on to it?

  Decision: In consideration of transaction costs and forecasting accuracy, we will replace the holding assets when the difference between the potential profits of the two assets exceeds a certain threshold:

$$(P'_{maxf} - P'_c) \cdot Q' - (P'_{maxf} + P'_c) \cdot Q' \cdot \alpha'\% - P_{maxf} \cdot Q \cdot \alpha\% - (P_{maxf} - P_{In}) \cdot Q \geq T_3 \tag{23}$$

6.3 Calculation Results

We can get the predicted daily price of both Gold and Bitcoin in the future 2 days according to Model I. However, it is noted that due to the mechanism of prediction, the prediction results are different each time. Therefore, in order to obtain more accurate results and better measure the model, we randomly select 20 running results and get their average value as the final calculation result of the model, the standard deviation is used to measure the fluctuation degree of the results.

Table 6: Numeral characteristics

Numeral characteristic	Value	Unit
Max	17670192	$
Min	4024800	$
$\mathbb{E}$(x)	7030220.000000	$
σ	3790473.313191	$

The daily total assets value of a relatively excellent investment choice can be obtained in Figure 11.

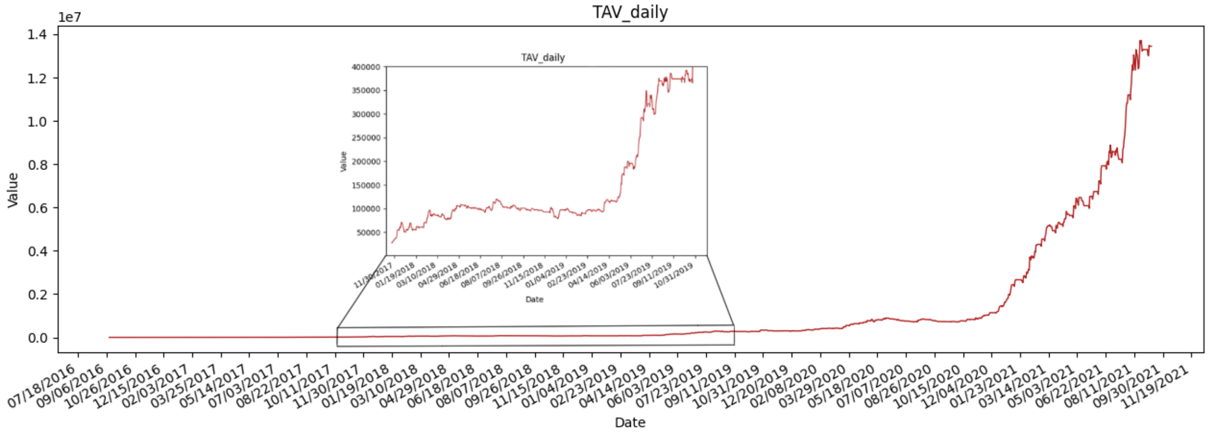


Figure 11: Daily total assets value

7 Proof of the best trade

In this section, we will prove the effectiveness of our model. Giving that calculating the mathematical ***max profit*** is a challenging task especially on large scale data, we introduced two present statistical indexes to compare with our model. Going through years of practice in stock market, the statistical indexes are proven to be efficient enough to guide traders of investment strategies and win a maximum profit. Our goal is to show the superiority of our model winning profits while having the similar decisions with indexes' strategies.

The indexes we would like to introduce are Relative strength index (RSI) and Exponential Moving Average (EMA):

7.1 Relative Strength Index

- **Introduction**:

  The relative strength index (RSI) is an indicator which measures magnitude of recent n-days price changes to indicate whether the stock market is in overbought or oversold condition. This index was originally invented by J. Welles Wilder Jr.

- **Calculation:**

$$RSI = 100 - \frac{100}{1 + RS}$$

$$RS = {}^{n\_days\ Average\ Gain}\!/_{n\_days\ Average\ Loss}$$

Where n is set to 5 here.

- **Metrics For Buying And Selling:**

  If RSI reaches over 70, the market is considered as overbought condition. If RSI drops below 30, the market is considered as oversell condition.

  Generally, when the RSI surpasses the horizontal 30 reference level, it is a bullish sign, and when it slides below the horizontal 70 reference level, it is a bearish sign. Put another way, one can interpret that RSI values of 70 or above indicate a security is becoming overbought or overvalued and may be primed for a trend reversal or corrective price pullback. An RSI reading of 30 or below indicates an oversold or undervalued condition.

7.2 Exponential Moving Average:

- **Introduction:**

  The exponential moving average is an average algorithm which give more weight on the most recent data points.

- **Calculation:**

$$EMA_t = \left(Price_t \cdot \frac{a}{1+n}\right) + EMA_{t-1} \cdot (1 - \frac{a}{1+n})$$

  Where a is a smoothing factor usually equals to 2 and n is the number of days.

- **Metrics For Buying And Selling:**

  When n equals 12, the 12-day EMA line is denoted as EMA12. EMA12 served as a warning line, whenever stock price reaches this line, there is likely a turning point ahead. In such, the buying and selling should happen whenever stock price reaches hits this line.

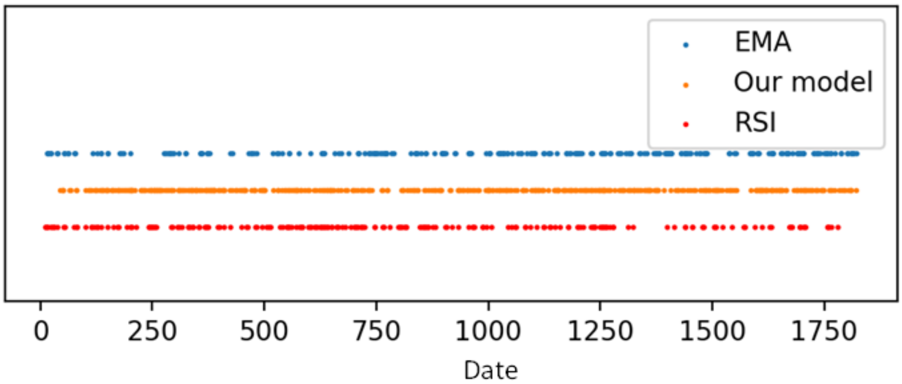


Figure 12: The buy/sell movements of our model and suggestions

The buy/sell movements of our model and suggestions by other indexes is shown above, each dot represents a buy/sell event. The Closeness index CI is defined as:

$$a_i = \begin{cases} 1, \min\left(abs\left(a_i - b_j\right)\right) < \alpha \\ 0, \min\left(abs\left(a_i - b_j\right)\right) \geq \alpha \end{cases}$$

where $i \in [0, n], j \in [0, m], n = model\_1\ transaction\ number, m = model\_2\ transaction\ number, \alpha\ denotes\ a\ threshold.$

$$A = \sum a_i$$

$$CI = \frac{A}{n} \times 100\%$$

The closeness index indicates the proportion of points in one sequence that are closely matched to the other sequence. One thing worth mentioning is that this algorithm only counts the proportion of the "close points", as for the rest points, their position does not influence the outcome because they are likely where the better performance originated.

7.3 Explanations

From the above formula, the indicators can be calculated directly.

Table 7: Metrics Results

Metrics	CI Result	Profit Ratio
EMA	0.9095238	1.083
RSI	0.9166667	1.100

In conclusion, the model we developed has shown its consistency with the strategies brought up by existing models. This indicates that our model makes right buy/sell decisions similar to what is known the "right ones" by practical experience. Besides that, our model also made some decision which other two do not have, this should be resulted from the prediction module in our model which give it abilities to "think" base on the past, present and future instead of just making summaries from the past. This advantage is also shown in the outcome, we are delighted to see that the ability of foreseeing future does contribute to the final profit. As such, we firmly believe that our model has its ability producing the best strategy.

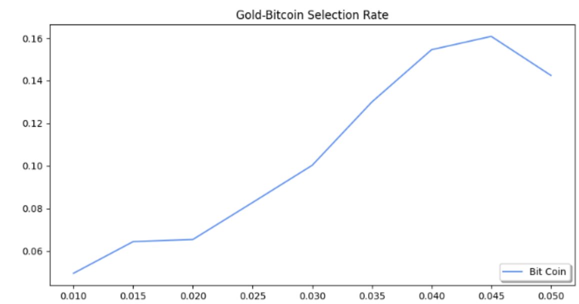
8 Sensitivity Analysis

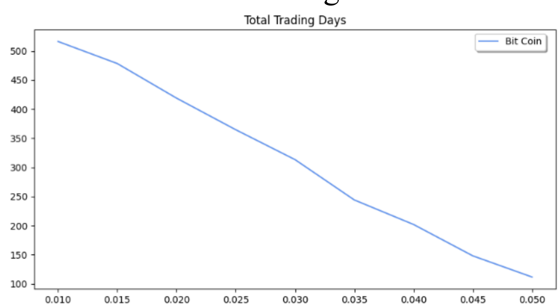
8.1 Transaction Costs

Transaction cost is the most important value of all parameters, because once the cost constraint is not considered, the global optimal solution can be achieved only by following the simplest greedy algorithm -- minimum point buy, maximum point sell. In order to obtain more accurate results and better measure the sensitivity, the approach used in section 6.3 is adopted again.

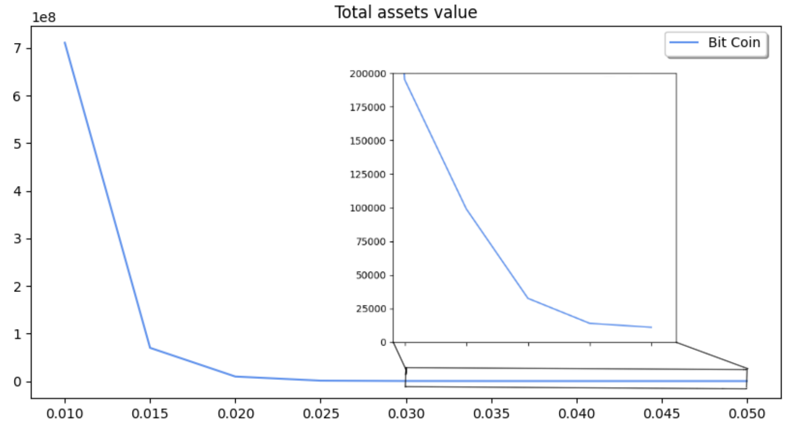
8.1.1 Transaction Cost of Bitcoin

Since most transactions will be conducted on Bitcoin, its transaction cost is more important than gold's, and we limit it to 0.01~0.05 to observe changes in total assets value, total trading days, and gold-bitcoin selection rate, which can be obtained in Figure 13.





(a) Gold-Bitcoin Selection Rate *(b) Total Trading Days*



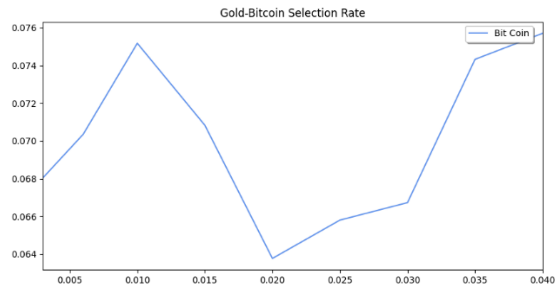
(c) Total assets value

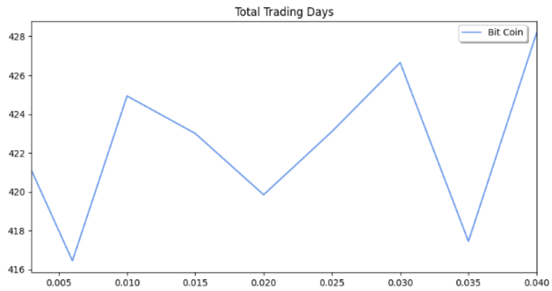
Figure 13. Gold-Bitcoin Selection Rate is the ratio of days of purchase for gold and bitcoin multiplied by the revised difference in the total number of days in the data; total trading days is the total number of days of trading.

It is clearly indicated that the total number of trading days has decreased because the cost of trading has increased, which has also led to a significant decline in the total value of assets because there are fewer days to buy and sell frequently for profit. At the same time, the investment model is more in favor of gold, which has a relatively much lower transaction cost than bitcoin. **In conclusion, the model is very sensitive to the change of Bitcoin's transaction cost.**

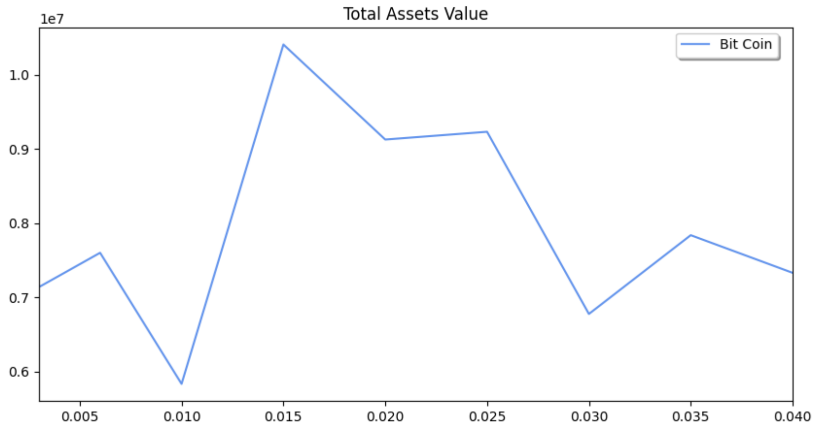
8.1.2 Transaction Cost of Gold

The variation of gold transaction cost is limited to 0.003~0.04





(a) Gold-Bitcoin Selection Rate *(b) Total Trading Days*



(c) Total assets value

Figure 14

Compared with bitcoin, the transaction cost of gold has almost no obvious impact on the value of assets, and the three indicators float randomly with the change of it. **In conclusion, the model is insensitive to the change of Gold's transaction cost.**

9 Model Evaluation

9.1 Strengths

- **Comprehensive consideration:** We comprehensively consider various intelligent algorithms and select the best basic model, combining with the Random Walk method based on economic statistics, and achieve accurate predictions of future value through Ensemble Learning.
- **Innovativeness:** We are not restricted to the basic algorithms, proposing the concept of decay factor on the basis of the principles of economics. Our mathematical decision model is designed by ourselves without relying on any existing algorithms. New metric has been come up with in the sensitivity analysis and plays an important role in the process.
- **Rigorous Logical Reasoning:** In the mathematical decision-making model, we analyze a variety of situations and make choices strictly based on constraints. Every step of reasoning is well-founded.

9.2 Weaknesses

- **Only short-term forecasts are possible.** Although our model can achieve excellent results in short-term prediction, it is knotty to realize long-term forecasting with historical data .
- **Subjective error when making decisions.** Due to strategies we make based on local prediction data, our algorithm is greedy to a certain extent even if the results are theoretically close to the optimal solution.

References

[1] Zhongyang Han, Member, IEEE, Jun Zhao , Member, IEEE, Henry Leung , Fellow, IEEE. A Review of Deep Learning Models for Time Series Prediction.

[2] Jerome T. Connor, R. Douglas Martin, Member, IEEE, and L. E. Atlas, Member IEEE. Recurrent Neural Networks and Robust Time Series Prediction.

[3] Reza Ghanbari and Keivan Borna. Multivariate Time-Series Prediction Using LSTM Neural Networks.

[4] Kiril Andreev Koparanov, Krasin Krasimirov Georgiev, and Vasil Aleksandrov Shterev. Lookback Period, Epochs and Hidden States Effect on Time Series Prediction Using a LSTM based Neural Network.

Memorandum

To: Traders
From: Team # 2200237
Subject: Make a bundle in the Wall Street: AI assistant helps everyone become Buffett in investment.
Date: February 22, 2022

Dear traders,

Having heard of your requirements finding a precise and robust model to assist you with your investments, we are more than delighted to make our effort. After a thorough analysis of data properties and the exchange market, we made a decision model with her eyes open to the future. Powered by powerful machine learning algorithm, the decisions can be made with precise forecast which eventually leads to a very promising result.

According to your requirements, we analyzed the historical price data of gold and bitcoin. During the investigation, we found that the fluctuations on bitcoin price is significantly larger than that on gold price, which could be a signal of greater profit potential. With that in mind we divided our model in to two modules which are prediction module and decision module.

To give her a pair of bright eyes, we implemented both machine learning and statistical method to let the advantage of the each to shine. The statistical method has an extremely high precision especially in the early days of your investment journey but its precision usually shows on a statistical scale or in another word, large scale. We do aware that every tiny prediction error could make your precious money suffering, an artificial neural network called LSTM is also implemented to ensure a reliable prediction. LSTM network could easily utilize the information hidden in the past to see the future. All parameters in this network are carefully tuned to preform the best in real-life scenario. Sufficient literature research has done by our team to ensure the fully optimization of the network structure. By doing so, I would say that the prediction performance is splendid and far beyond my expectations which gives a solid foundation to the decision module that coming subsequently.

Seeing far, thinking smart. I would say that having a great decision module is like being born with a smart brain. With reliable foreseeing data in hands, the decision module works in its confidence. This module works base on the ongoing trends and the predictions to make decisions. It can automatically recognize the potential minimum price point to buy your favored assets and sell them when the price reaches to the highest. Meanwhile, due to the transaction fees, the decision module can also balance the possible income and potential fees and pick the right time to act. During the modeling, we surprisingly found that it is the most beneficial holding one asset at a time, so please follow the recommendations of our model and make decisive calls.

With all that being said, we formulate reasonable strategies for you to consider: (1) Avoid making frequent transactions when the price increasement is pace and slow. (2) Try making more transactions on the bitcoin where violent up and down create great opportunities to realize the appreciation of your property. (3) We recommend you have fully trust in our model, it will tell you when the opportunity comes.

Thanks for your time out of your tight schedule. In the simulations, our model could gain around 18million at most. We sincerely hope that our model can lend a hand to your success.
Sincerely,
MCM Team # 2200237

Appendix:

Github Code link, creating convenience for you to check our results.

https://github.com/zhichunjing/2022MCM_C