Problem Chosen	2023	Team Control Number
C	MCM/ICM Summary Sheet	2223196

Hybrid Trading Model Based on MA and LSTM

Summary

Market trading is full of movement, and with the emergence and insane rise of cryptocurrencies, people are gradually starting to think of them as an investment tool just like gold.

In this paper we try to explore a quantitative trading procedure that is able to identify potential investment opportunities through historical price data. Based on the limitations of the data set, we follow several assumptions.

- The price of gold and bitcoin is only determined by the historical price.
- The market price does not change during the day.

Based on our research, we found that quantitative trading programs are generally divided into traditional programs that use only MA models [?] and new programs that use machine learning tools such as LSTM/SVM [?]. We first built an automated investment program using the basic MA strategy, from which we realized that such a model has the following drawbacks.

- When the price does not fluctuate much, the difference between the long average and the short average is small and the two curves will intersect frequently, resulting in frequent trading situations. Since we have a high commission, this can result in large losses.
- This strategy has a certain lag, so in the event of a spike or a crash, investment opportunities will be missed.

To optimize such a situation, we would predict the future price trend with the Long Short-Term Memory network. However, predicting price trends over time by historical prices is not reliable enough. Our new trading model chooses to combine MA strategies and LSTM neural networks, which gives us the following advantages.

- The LSTM network can detect the presence of a trend before the MA.
- The MA strategy provides reliable warning of trend changes to the LSTM network and allows the LSTM to predict the trend magnitude.

Eventually, we developed a new trading program that combined the advantages of both models and ran simulations on historical data to obtain a final asset of $17,834.09 with an average annualized interest rate of 336.68%. And the Sharpe ratio had a result of 1.32. It outperformed the benchmark MA model in terms of results across the board.

Keywords: LSTM; Moving Average; Quantitative Trading

Hybrid Trading Model Based on MA and LSTM

February 17, 2023

Summary

Market trading is full of movement, and with the emergence and insane rise of cryptocurrencies, people are gradually starting to think of them as an investment tool just like gold.

In this paper we try to explore a quantitative trading procedure that is able to identify potential investment opportunities through historical price data. Based on the limitations of the data set, we follow several assumptions.

- The price of gold and bitcoin is only determined by the historical price.
- The market price does not change during the day.

Based on our research, we found that quantitative trading programs are generally divided into traditional programs that use only MA models [**?**] and new programs that use machine learning tools such as LSTM/SVM [**?**]. We first built an automated investment program using the basic MA strategy, from which we realized that such a model has the following drawbacks.

- When the price does not fluctuate much, the difference between the long average and the short average is small and the two curves will intersect frequently, resulting in frequent trading situations. Since we have a high commission, this can result in large losses.
- This strategy has a certain lag, so in the event of a spike or a crash, investment opportunities will be missed.

To optimize such a situation, we would predict the future price trend with the Long Short-Term Memory network. However, predicting price trends over time by historical prices is not reliable enough. Our new trading model chooses to combine MA strategies and LSTM neural networks, which gives us the following advantages.

- The LSTM network can detect the presence of a trend before the MA.
- The MA strategy provides reliable warning of trend changes to the LSTM network and allows the LSTM to predict the trend magnitude.

Eventually, we developed a new trading program that combined the advantages of both models and ran simulations on historical data to obtain a final asset of \$17,834.09 with an average annualized interest rate of 336.68%. And the Sharpe ratio had a result of 1.32. It outperformed the benchmark MA model in terms of results across the board.

Keywords: LSTM; Moving Average; Quantitative Trading

Contents

1 Introduction

With the wild performance of bitcoin's sharp rise in early 2021, more and more investors and investment institutions are starting to buy bitcoin. In a research note to clients, Goldman Sachs analyst Zach Pandl said that bitcoin would take market share away from gold in 2022 as safe-haven assets. [**?**] Developing trading models by mining historical price data can be crucial in instructing investments in safe-haven assets.

1.1 Problem Summary

A trader wants to develop a trading model using gold and bitcoin market prices from 9/11/2016 to 9/10/2021. The daily trading strategy should base on price data before that date. And notice that gold can't be traded on weekends.

To achieve the goal, we need to solve the following problem.

- Develop a model that offers the best trading strategy based on historical prices.
- Prove that our model gives the best scheme.
- Determine the influence of transaction costs on our strategy and results.

1.2 Previous Work

According to Adam Zaremba's research [**?**], futures management strategies can be divided into the following categories.

Figure 1: Futures management strategies

Trend following strategy like Moving averages (MA) is statistically well-defined and easy to implement programmatically. It smoothes the time series, which will dampen the noise of short-term price fluctuations. MA produce signals with a delay while non-trend following strategies try to use pattern recognition to forecast future price movement.

Other researchers have tried to use neural networks to predict the price movements of the market. [**?**] Long short-term memory (LSTM) networks are state-of-the-art sequence learning for time serise forecasting [**?**], researchers hope to predict the future direction of market prices by mining the intrinsic linkage of time-series data.

1.3 Our Method

We first traded the market relying exclusively on the MA strategy and found that the model has a lag for trend detection and suffers from frequent buying and selling during periods of price stabilization. Then based on a study by Yunpeng, Liu et al. [**?**] we use LSTM networks for multi-step ahead time series to predict the future prices of our portfolio.

We want to combine the advantages of the MA strategy and the LSTM network. Enabling us to use past data to identify potential future growth possibilities and to sell our positions at the highest point possible.

Figure 2: Idea of LSTM + MA hybrid model

2 Preparation

In this section, we will discuss our preparation for solving the problem. Consists of notation and data preprocessing and our basic assumptions and rationale for the problem.

2.1 Notations

Symbol	Description
i	Date, $i = 0, 1, 2...1825$
j	$j = 0$ for USD, $j = 1$ for gold, $j = 2$ for bitcoin
$C_{i,j}$	Cell state, one of the inputs to the output gate.
$S_{i,j}$	Output information of the network
L	Number of days that LSTM networks make the decision based on
$X_{i,j}$	Current day's price data and previous L -1 days' prices data
$P_{i,j}$	Real historical price on day i
$P^*_{i,j}$	LSTM networks' prediction of prices
MA_N	N days moving average
$Portfolio_{i,j}$	USD cash/Gold/Bitcoin held on day i
$Slope_{i,j}$	Slope of the price for the next seven days
$ESP_{i,j}$	Buy/Sell Inhibition Threshold
$Risk_{i,j}$	Price increase indication
$Rb_{i,j}$	Percentage of $Portfolio_0$ that can be used to inveset

Table 1: Notion list

2.2 Data Preprocessing

All the data for our model comes from the two spreadsheets provided by COMAP. From a preliminary analysis of the data, we found that gold has missing prices on specific dates and there is no price data for gold on holidays. The data validation in Fig.3 shows that the average of the errors using the two-day average price as a replacement for the real price is only 0.44%. To align the transaction dates of bitcoin and gold on the timeline, we use the same approach to fill in the missing values.

We do the following approaches to clean the data-set:

- Use the average price of the two adjacent days to fill in the missing values of the gold price.
- For those days when gold is not tradable, we also fill the average price of the two adjacent days.

Figure 3: Error of replacing true value with two-day average

2.3 Assumptions and Rationale

- **Assumption 1:** The price of gold and bitcoin is only determined by the historical price of gold and bitcoin and is not influenced by any other factors. This is the most fundamental assumption of this paper because we only explore the impact of historical prices on the future prices of gold and bitcoin.
- **Assumption 2:** The same investment theories apply to gold and bitcoin. According to a study by Jamal Bouoiyour et al. [**?**] gold and bitcoin benefit from the same economic conditions. Therefore we use the same investment strategy for gold and bitcoin.
- **Assumption 3:** On the first day of the trading day, there will be an equal propensity to buy bitcoin and gold. Because without the support of historical trading data, we believe that the returns obtained from investing in gold and bitcoin are equivalent. This makes it fair for the investment model to adjust its trading propensity on its own after several trades.
- **Assumption 4:** We assume that the market price does not change during the day, because the dataset provides only one price for a day.

3 Model I: Price Prediction Model

We want to anticipate the future price of the market for our portfolio so that we can buy at the lowest possible price and sell before the price plummets as much as possible. Therefore, it is extremely important to build a price prediction model, and in this section we will discuss how we can build our price prediction model and validate its predictive effect.

3.1 Long short-term memory

A hidden structure proposed by Hochreiter et al. in 1997 [**?**], called LSTM. it was designed to solve the problem of gradient disappearance in long-term sequences. The LSTM structure contains three modules: forget gates, input gates and output gates.The forget gate is responsible for removing information from the network, the input gate controls the information saved by the network and finally the correct information is output through the output gate. the LSTM also introduces a "Cell State" for the network that allows information to be saved for a long time. The structure of the network is shown in the Fig.4 below.

Figure 4: Structure diagram of LSTM single layer network

Based on our research on previous papers [**?**] [**?**], we decided to use long and short term neural networks for multi-step ahead time series to predict the future trend. As shown in the Fig.5, forecasting multi-step ahead time series means that we add the predicted price of day i to the data set, then remove the data of day $i - L$, use this data set as input to forecast the price of day $i + 1$, add this data, remove the data of day $i + 1 - L$ and so on. Until all the data in the data set were replaced by predicted data, then stop further forecasting.

Figure 5: Multi-step time series forecast

We demonstrate through extensive experiments that using a four-layer network with a setting L of 30 and training the last 180 days of price data provides the best prediction results.

3.2 Results and Validation

By verifying the positive and negative nature of the product of $P_i + 7 - P_i$ and $P_i^* + 7 - P_i^*$, we arrive at a 7-days prediction accuracy of 56% for bitcoin and 50% for gold. Some of the predicted results are shown in Fig.6.

Figure 6: Bitcoin and Gold 7-days prediction

By analyzing the real price of gold we conclude that the low accuracy of the seven-day forecast for gold may be influenced by the low price of gold itself, and that small price forecast errors are magnified by the multi-step ahead time series forecast.

4 Model II: Trading Model

4.1 Basic Rules

- **Rule 1:** According to Assumption 2, our model assigns equal Rb of 0.5 to bitcoin and gold, and at the daily settlement, the gains obtained by gold and bitcoin are counted, and the party with the higher gain receives 0.05 Rb of the other party. The highest value of Rb is 0.9.
- **Rule 2:** When we hit a date when gold cannot be traded, we automatically block gold from trading, whereas bitcoin is not subject to this restriction.
- **Rule 3:** All of the following discussions are based on individual investment targets, and the amount of USD they can use each time has been determined through Rule 1.

4.2 Baseline Model

We first use pure doule moving mverage trading as a benchmark for our trading model. Equation.1 demonstrates definition for moving average (MA). [**?**] k represents the date of the day we traded and N represents the average of the price of the days used by MA.

$$MA_N = \frac{1}{N} \sum_{i=k-N+1}^{k} P_i \tag{1}$$

We use the Equation.2 to determine whether to buy or sell. 1 means buy 0 means sell.

$$f(MA_{30,i}, MA_{30,i-1}, MA_{5,i}, MA_{5,i-1}) = \begin{cases} 1 & MA_{5,i} - MA_{30,i} \geq 0 \& MA_{5,i-1} - MA_{30,i-1} < 0 \\ 0 & MA_{5,i} - MA_{30,i} \leq 0 \& MA_{5,i-1} - MA_{30,i-1} > 0 \end{cases} \tag{2}$$

4.3 Model Analysis

We run the most basic MA trading model, as shown in Fig.7 to see how our asset allocation looks on the time series chart. The buy and sell scenarios forbitcoin and gold are shown in Fig.9 and Fig.8 along with the price and moving average movements, respectively. At the end our total assets reached $6899.23.

Figure 7: Asset allocation time series chart for MA strategy

Figure 8: MA strategy for buying and selling gold

We got a 6.89x improvement in assets, but the MA strategy still has the following problems. As shown in Fig.9, the MA strategy is unable to sell off assets before a real decline in prices occurs.

Figure 9: MA strategy for buying and selling bitcoin

We can also find in Fig.9 that our baseline model is influenced by the crossover line to buy at the highest point and sell at the lowest point. Also in Fig.9 we can see that the MA strategy makes frequent purchases and sales during periods of small price fluctuations, losing a significant amount of commission.

The above experimental results all prove that there is a delay in generating signals from MA crosses, so we would like to introduce a new model to make the trading program aware of uptrends or downtrends earlier.

4.4 Model Improvement

To improve the shortcomings of the previous model, we introduced LSTM to forecast future trends with a view to trading at earlier points in time. The new improved model is divided into two parts, MA crossover occurs and MA non-crossover occurs.

We first define the variable $Slope_{i,j}$, which obtains the slope from a linear regression of the set $\{P_k^*|k = i, i+1, \ldots, i+6\}$. Then we derive Buy/Sell Inhibition Threshold $ESP_{i,j}$ from Equation.3 . The $ESP_{i,j}$ is compared with $Slope_{i,j}$ for the purpose of eliminating the effect of the size of the price value on $Slope_{i,j}$. We also define the variable $Risk_{i,j}$ to represent the price increase indication through Equation.4 .

$$ESP_{i,j} = 0.4 * \frac{1}{20}\sum_{i}^{20} Slope_{i,j} \tag{3}$$

$$Risk_{i,j} = \frac{MA_{5,i,j} - MA_{30,i,j}}{MA_{30,i,j}} \tag{4}$$

For MA with crossover occurrence, We do not trade for cases where $Slope_{i,j}$ is less than $ESP_{i,j}$, while when $Slope_{i,j} \geq ESP_{i,j}$, there are two cases.

- When the result of Equation.2 equals 1, We invest $Portfolio_{i,0} * Rb_{i,j}$;
- For the result of Equation.2 equals 0, We sell all our $Portfolio_{i,j}$.

For MA without crossover occurrence, We first determine whether $Slope_{i,j}$ is greater than 0.05*$P_{i,j}$, if not then we end, otherwise we then divide into two cases.

- $Slope_{i,j} > 0$, We invest $Portfolio_{i,0} * Rb_{i,j} * MAX(1, -risk)$;
- $Slope_{i,j} < 0$, We sell $Portfolio_{i,j} * MAX(1, risk)$

The complete model diagram is shown in Fig.10

Figure 10: Hybrid model structure

4.5 Model Validation

Results of the improved model

We run the Improved model with the asset allocation timing chart shown in Fig.11, and at the end of the trade we have total assets of $17,834.09. Its trading strategy for bitcoin and gold is shown in Fig.12 and Fig.13.

Figure 11: Asset allocation time series chart for hybrid model of MA and LSTM

Figure 12: Hybrid model for buying and selling bitcoin

Figure 13: Hybrid model for buying and selling gold

We can see that our hybrid model has made great strides in forecasting ahead of the trend (as shown in Fig.12), and we have also used the improved model to successfully avoid the problem of frequent trading in the baseline model leading to loss of fees (as shown in Fig.12).

We can also note that no transactions occurred in the assets between 1500 and 1600 days, which represents an effective suppression of transactions during periods of small price fluctuations by our program, avoiding fee losses.

Comparing the improved model with the baseline model

The baseline model has an asset of $6,899.24, while our hybrid model ends up with an asset of $17,834.09, which has a 2.58x return improvement. The annualized interest rate for the hybrid model is 356.68% and the baseline model for the MA strategy is only 137.98%.

The comparison between Fig.9 and Fig.12 shows that we achieve better market trend prediction and win more profit compared to the base model; thanks to the addition of the LSTM model, we are able to predict the possibility of small price fluctuations and avoid loss of fees due to MA crossovers that misjudge this.

To better evaluate our model, we introduced the following three parameters, the annualized interest rate (defined in Equation.5), the Sharpe ratio (defined in Equation.6) [**?**] and the maximum retracement rate (defined in Equation.7).

In Equation.6, E represents expected value, R_a represents asset return, R_b represents the risk-free return (We use the interest on bank deposit as an estimate), and σ_a represents the volatility of the portfolio. In the actual calculation, we use the historical annualized return and its standard deviation instead of the ER_a and σ_a.

$$AnnualizedIntereestRate(i) = \sum_{j=0}^{2} Portfolio_{i,j} - \sum_{j=0}^{2} Portfolio_{i-365,j} \tag{5}$$

$$SharpeRatio = \frac{E[R_a - R_b]}{\sigma_a} \tag{6}$$

$$MaxRetracement = MAX(\{\frac{\sum_{j=0}^{2} Portfolio_{i,j} - \sum_{j=0}^{2} Portfolio_{k,j}}{\sum_{j=0}^{2} Portfolio_{i,j}} | i \in (0, N) \& k \in (i, N]\}) \tag{7}$$

According to the equations above, our baseline model has a Sharpe ratio of 1.323 and a maximum retracement of 60.61%, while our hybrid model has a Sharpe ratio of 0.821 and a maximum retracement of 45.93%. As a rule of thumb, a Sharpe ratio above 1 indicates that the return of an investment is higher than the risk, while a Sharpe ratio less than 1 indicates that the risk of an investment is higher than the return. The numerical comparison shows that our hybrid model has a higher Sharpe ratio compared to the baseline model, which represents a higher return-to-risk ratio for our improved model.

Figure 14: MA strategy maximum retracement

Figure 15: Hybrid model maximum retracement

The periods when the maximum rate of retracement occurs for the benchmark and hybrid models are shown in Fig.14 and Fig.15, respectively. A lower maximum retracement rate represents a smaller maximum loss that an asset can receive and, therefore, our hybrid model has better asset stability performance compared to the benchmark model.

Therefore, a comparison of the final data metrics for the hybrid and baseline models is shown in the Table.2 below. With all these data, we have demonstrated that our model provides the optimal strategy.

Indicators	Hybrid model	MA strategy
Final Total Assets	$17,834.09	$6,899.24
Annualized Interest Rate	356.68%	137.98%
Sharpe Ratio	1.323	0.821
Max. retracement rate	45.93%	60.61%

Table 2: Comparison of hybrid and baseline models

5 Conclusions

5.1 Problem Results

Develop a model that offers the best trading strategy based on historical prices.

Our hybrid model combines both LSTM neural network and MA strategies. When MA detects a trading trend, LSTM intervenes in the prediction of future volatility, preventing frequent buying and selling when small fluctuations ($Slope_{i,j}$ is less than the inhibitory value $ESP_{i,j}$); when MA does not detect a trading trend, LSTM model is responsible for predicting the future trend, when $Slope_{i,j}$ greater than our trading risk control (risk factor 0.05*$P_{i,j}$), we make buying and selling decisions in advance of the MA strategy. Our model compensates for the disadvantage that the baseline model uses the MA strategy that causes decisions to be delayed behind the market, and anticipates market movements in advance by using historical data.

Prove that our model gives the best scheme.

According to Section 4.5, Our hybrid model has an Annualized Interest Rate of 356.68%, a Sharpe ratio of 1.323, and a maximum retracement of 45.93%; all metrics outperform the baseline model across the board. Both ensure earlier market trend detection and also ensure that frequent trading will not be withheld due to small market fluctuations.

Our model is well optimized for high fees by reducing the frequency of trading; and the model is fully based on historical data and does not rely on a priori knowledge of the market, thus having good robustness. In summary, our model implements the optimal strategy.

Determine the influence of transaction costs on our strategy and results.

As shown in Table.3, the results of our two models for different transaction costs are as follows. Since our trading logic is not tied to fees, fees have no impact on our strategy. However, our strategy itself is to reduce the frequency of trading as much as possible to reduce the impact of fees on our asset value.

5.2 Strengths

- MA strategy will make frequent buy/sell operations when the price of investment is less volatile and be charged with frequent fees. The hybrid model can prevent such trades by predicting future trends, effectively reducing the commission.

- The hybrid model can predict the future price trend by LSTM, so that you can avoid more

Gold Fees	Bitcoin Fees	MA strategy Final Assets	Hybrid model Final Assets
1%	2%	6899.24$	17834.09$
0%	0%	15657.04$	35712.03$
2%	1%	7105.01$	23199.3$
2%	2%	5339.12$	16840.98$
1%	1%	9158.83$	24557.97$

Table 3: Final asset returns for different models with different fee ratios

losses and get more profits by trading before the MA strategy intersects the average line.

- The LSTM model has poor prediction ability for price peaks and troughs, and the hybrid model can make more reasonable decisions by averaging the averages to assist in judgment.

5.3 Possible Improvements

We can still find imperfections in our model, for example, the trend judgment of gold is still not accurate enough (as in Fig.16), which we speculate is caused by the poor setting of the training parameters of the LSTM model. Also as shown in Fig.12, our model has some flaws in the judgment of the trend for the extreme upward price trend.

Figure 16: Hybrid model failed to predict gold price trend

6 Memorandum

To: The trader who asked for the model

From: MCM Team #2223196

Subject: Introducing you to an excellent trading model

Date: Feburary 21, 2022

Dear Trader,

Thank you for trusting us to help you develop your trading model.

Our model relies solely on gold and bitcoin price data from September 11, 2016 through the day the decision was made to come up with the decision. Since gold cannot be traded on double holidays and some holidays, when the model decides to trade on a non-tradable day, we record this trading decision and trade on the first trading day thereafter.

Strategy of the model

During the first 30 days, we need to collect prices to provide reliable data for subsequent decisions. Therefore, we do not make any trading decisions during the first 30 days to avoid unwanted losses.

After the 30th day, we will find the average price of the 5 days before the decision day and the previous 30 days, which are the long average and the short average, respectively. The long average is connected to form the long average and the short average as well. We find that.

- when the short average crosses the long average from the bottom up, the price will be at the beginning of an upward phase and we will invest a certain percentage of dollars in the investment.
- When the short average crosses the long average from the top, the price starts to fall and we sell all our investments.

As you may have noticed, when adopting the above strategy, there are two unavoidable problems. First, when the price does not fluctuate much, the difference between the long average and the short average is small and the two curves will intersect frequently, resulting in a frequent trading situation. In this case, the money we earn through investment may not even cover the fees brought by the transaction, resulting in a loss. Second, This strategy has a certain lag, so in the event of a sharp rise or fall, this strategy will miss investment opportunities.

To prevent this, we will optimize the above strategy by predicting future price trends with the Long Short-Term Memory neural network. However, since training the neural network requires at least 180 days of data, our trading decisions for the first 180 days will have to rely solely on the Long Average vs. Short Average decision.

After 180 days, we will use the historical prices since six months and train the LSTM neural network with this data so that the network can predict the price of the next day by the historical prices of the previous 7 days. Then, we use the predicted price with the previous 6 days' historical price to predict the price of another day, and so on, thus predicting the price of 7 days. Of course, due to the uncertainty of the market, the predicted prices will not be particularly accurate, but we can still use these predicted prices to see the price trend and the approximate increase or decrease. We established two thresholds, one large and one small.

- When the predicted increase or decrease is within a small threshold, we do not place any trades to reduce trading frequency and commission expenses.
- When the predicted rise or fall exceeds a large threshold, implying a spike or plunge, we perform a buy/sell operation even if the averages do not intersect. Of course, we do not buy/sell all the assets to reduce the impact of a failed neural network forecast. The amount of buy/sell here is related to the ratio of short averages above/below long averages. The larger this ratio is, the higher we think the probability of a crash/spike occurs, so we will sell/buy assets with a larger share.

The above section is a strategy for one investment. When this strategy is used for both bitcoin and gold, it constitutes our model. However, we put some restrictions on the buying operation because we can't use all of our money to buy the same investment.

We initially assume that both products have the same earning power. So for both investments, we invest half of the money when the buy decision is made. Later we compare the amount of money made by both, increase the proportion of the product with the higher earning power, and decrease the proportion of the other one. After all, there's a proverb that goes: Don't put your eggs in the same basket.

Performance of the model

We used the model to simulate trading the price of gold and bitcoin from September 11, 2016 to September 11, 2021, with very good results. In the simulated trades, our assets rose from $1,000 at the beginning to $17,834.09 at the end, with an average annualized rate of 336.68%. The Sharpe ratio, on the other hand, was 1.32, which shows that the payoff of the trade was higher than the volatility risk. The maximum retracement rate in the simulated trade, on the other hand, was 45.93%, which is a big improvement compared to the more popular MA model's maximum retracement rate of 60.61%.

Finally, we would like to thank you again for choosing us to develop your model and hope you can achieve more gains in your future trading with our model.

Sincerely yours,

MCM Team #2223196

Appendices

Here are simulation programmes we used in our model as follow.
Data Cleaning(main logic): Hybrid Model(main logic):