PADRRI: Prediction And Decision Models For Best Return And Risk With High Interpretability

Gold and Bitcoin are two volatile assets that market traders often need to buy and sell frequently to maximize their total return. Due to market factors such as the volatility of prices, the policies of quantitative investment are of great significance to resist the risk and get profit. To provide the best daily trading strategy that achieves the maximum return, we establish the *PADRRI* model (**P**rediction **A**nd **D**ecision models for best **R**eturn and **R**isk with high **I**nterpretability).

The PADRRI model is divided into two parts. The first part is the forecasting model, whose accuracy has a direct impact on the strategy of our model, and in turn affects the final total return. We compared some common time series forecasting algorithms, including deep learning-based methods such as LSTM , Autoregressive Integrated Moving Average model (ARIMA) and so on. With comparison and experiments, *Prophet* algorithm is finally adopted which accurately and efficiently predicts the prices of gold and bitcoin with higher interpretability. Additionally, to reduce overfitting and improve the prediction accuracy, an improved Prophet algorithm called X-Prophet is used by introducing a strong classifier model, XGBoost.

Based on the high instability of the bitcoin price, *Sliding Window Mechanism* is proposed to predict the future bitcoin price using only data within a sliding window (e.g., the previous season). Interestingly, this design is highly consistent with the way economics experts predict bitcoin prices in real-world scenarios, which can be regarded as a combination of machine learning and human intuition. The proposed approach significantly improves the prediction accuracy and provides strong support for decision models. With the improved forecasting model X-Prophet cooperated with the sliding window mechanism, we achieve highly accurate predictions of gold and bitcoin prices and provides a solid foundation for the trading decision phase.

The second part is the Decision model that uses the prediction results of the forecasting model to generate the best daily trading policies to maximize the return. However, relying solely on forecasts of future prices may lead to higher investment risks. To balance this risk, a combination of forecasting results and the *5-day SMA Principle* are adopted. In this way, expert knowledge in finance can be combined with machine predictions to provide more robust decisions in a human-machine collaborative fashion. Our model has proven to be quite effective in making decisions and has resulted in a significant return.

For PADRRI, we also provide evidence that our algorithm can maximize the total return for market traders. The effectiveness of the prediction model X-Prophet is first demonstrated, followed by a presentation of the decision model's ability to provide the best trading strategy under the assumption of highly accurate prediction results. The experimental results show that our prediction model and decision model can achieve a pretty good result.

We also conduct sensitivity analysis, which shows that our model is robust to the key parameters like transaction costs. And then the strengths, weaknesses and further improvements of our model are summarized. A brief conclusion is presented in the paper. Finally, a memo which communicates our strategy, model, and results to the trader and provides suggestions on the best daily trading strategy is written in the end of paper.

Keywords: Forecasting, Prophet, Sliding Window Mechanism, 5-day SMA

Contents

1 Introduction

1.1 Background

Gold's price volatility has been moderate, and liquidity has been adequate, allowing investors to avoid investment risks. Bitcoin, when compared to gold, has a higher yield, greater volatility, and is unregulated and taxed, giving it a lot of room for growth in the financial market. It can also be used to hedge against gold. As a result, a gold-and-bitcoin portfolio is extremely important. Market traders frequently purchase and sell highly volatile assets in order to maximize their total income.

According to The Street's survey [6], bitcoin has gradually entered the lives of the general public and is having a greater impact on economic development. However, the need to pay a commission for each purchase and sale, as well as the volatility of the financial markets, makes it very telling how to allocate funds wisely for purchases and when to buy and sell gold and bitcoin. Typically, investment traders will forecast future gold and bitcoin prices based on recent gold and bitcoin price movements, and then decide whether to buy, sell, or hold based on their forecasts, maximizing their profits and mitigating risk.

Figure 1: Results of The Street's bitcoin poll, which interviewed 1,005 U.S. citizens. [6]

1.2 Restatement of the Problem

We need to develop a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio. Trader's portfolio only contains gold and bitcoin, and we need to note that bitcoin can be traded every day, but gold can be only traded on days the market is open.

After through in-depth analysis and research on the background of the problem, we can specify that our article should cover the following aspects:

1. Develop a model that gives the best daily trading strategy based only on price data up to that day. And we should use our model to figure out how much is the initial $1000 investment worth on 9/10/2021.

2. Present evidence that our model provides the best strategy.

3. Determine how sensitive the strategy is to transaction costs. How do transaction costs affect the strategy and result.

4. Communicate the strategy, model, and results to the trader in a memorandum of at most two pages.
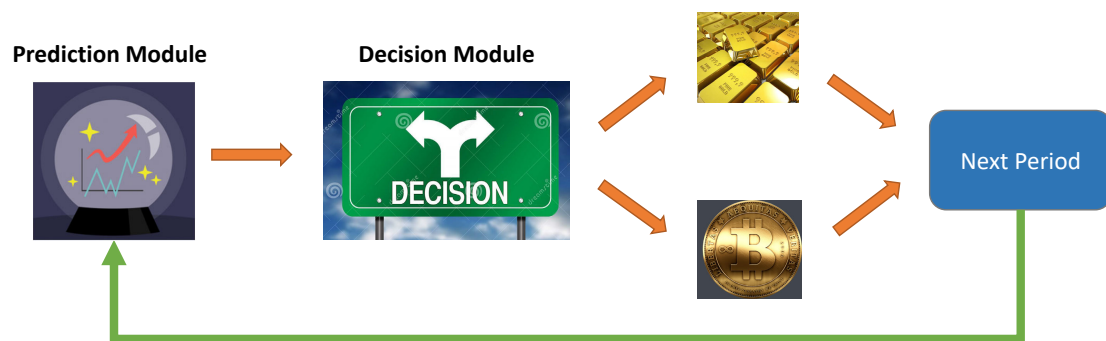
1.3   Our work

Figure 2: Overall structure of our model

To provide the best daily trading strategy that achieves the maximum return, the ***PADRRI*** model is established. It is divided into two parts, namely, forecasting model and decision model. The X-Prophet algorithm is adopted with sliding window mechanism, while the 5-day SMA Principle is combined with prediction result to generate the trading policies. Our work mainly includes the following:

1. The X-Prophet algorithm is adopted to improve the accuracy of the prediction and prevent overfitting with high interpretability based on statistic model.

2. The decision model is established to give the best daily trading strategy based on the prediction result and 5-day SMA Principle to balance the return and risk.

3. Present evidence that our model provides the best strategy.

4. Do sensitive analysis to show how transaction cost affect the strategy and result.

2   Assumptions

To simplify the analysis of our problem, we make the following assumptions, each of which is properly justified.

1. We assume that the prices of gold and bitcoin are not significantly correlated, since gold and bitcoin are two investment assets with different characteristics in financial markets.

2. Assume that the current prices of gold and bitcoin have a strong correlation with previous prices, so that we can use previous daily price to predict the tend of price in the future.

3. Since the pricing data files **LBMA-GOLD.csv** and **BCHAIN-MKPRU.csv** can be used, we assume that only the previous daily prices can affect the future price of gold and bitcoin.

3   Abbreviation and Definitions

For convenience, we introduce some important notations below.

Notations	Explanations
a(t)	the trend of the time series over the non-period
b(t)	the period term of Prophet
c(t)	the holiday term of Prophet
ϵ	the error term
H(t)	the model capacity
k	the growth rate
θ	the amount of change in the growth rate
m	the offset parameter
β	the seasonality prior scale
t	holidays prior scale, and the default value is 10

4   Model

4.1   Prediction Model

Gold and bitcoin price forecasting is a typical time series forecasting problem. The purpose of analyzing time series is to predict future trends. Traditional time series forecasting methods, such as moving average, are used to smooth out the time series. Increasing the number of periods in the moving average method will make the smoothing of fluctuations better, but it will also make the forecast values more insensitive to the actual changes in the data. With the development of machine learning and artificial intelligence, a large number of machine learning algorithms and deep learning algorithms are used to solve time series forecasting problems. For example, the long and short term memory network (LSTM) can solve the defect of recurrent neural network (RNN) which is difficult to remember the data features with long time interval, but it is still not a good solution for long sequences. The random forest algorithm can provide an effective method to balance the data set error, but it does not solve the regression problem well and is not able to make predictions beyond the range of the training set data.

In this paper, we use the Prophet model, which has a simple, easy-to-explain periodic structure that is well compatible with and explains event nodes such as holidays, and experiments show that Prophet performs relatively well in predicting future prices of gold and bitcoin. Since the original Prophet is prone to over-fitting problems at some specific time points, which affects the prediction results, we use an improved version of Prophet to predict the future prices of gold and bitcoin.

4.1.1 Model Selection

ARIMA is the most commonly used time series forecasting model. However, ARIMA requires the timing data to be stationary, or stable after being differencing. And, ARIMA can only capture linear relationships by nature, but not nonlinear relationships. Therefore, to predict the time series data using ARIMA model, it must be stable, and if the data is unstable, it cannot capture the pattern. For example, the reason why Bitcoin data cannot be predicted with ARIMA is that it is non-stationary and often fluctuates due to policies and news.

LSTM is also commonly used in time series prediction because of its powerful performance. However, on the one hand, deep learning-based methods such as LSTM rely on the quantity of data and are prone to overfitting. In other words, even if LSTM fits the current data with high quality, it may lose generalization due to overfitting, which is not very meaningful for practical scenarios. On the other hand, LSTM requires high demand on computational resources and may take longer training time in the current environment. Finally, LSTM lacks interpretability and is difficult to convince the market.

Therefore, a more interpretable and efficient time series prediction model is required for this task.

4.1.2 Prophet

Prophet [1] is a data prediction tool based on python and R language open source by Facebook. As shown in Fig. 3. The algorithm is designed based on time series decomposition and machine learning fitting, and can handle not only the case of some outliers in the time series, but also the case of some missing values. Prophet can predict the future trend of the time series almost fully automatically, due to the use of pyStan, an open source tool, in the fitting model, also makes prophet can get the prediction results in a very short time.

Prophet is an additive model that consists of four components, namely the trend term, the period term, the holiday term and the error term.

$$p(t) = a(t) + b(t) + c(t) + \epsilon \tag{1}$$

where a(t) denotes the trend term, which is the trend of the time series over the non-period. b(t) denotes the period term, c(t) denotes the holiday term, which indicates the presence of a holiday on that day, and ϵ denotes the error term.

There are two important functions for the a(t) trend term, one based on the logistic regression function and the other based on the segmented linear function.

The logistic regression-based growth model is as follows:

$$a(t) = \frac{H(t)}{1 + exp(-(k + z(t)'\theta) * (t - (m + z(t)^T \gamma)))} \tag{2}$$

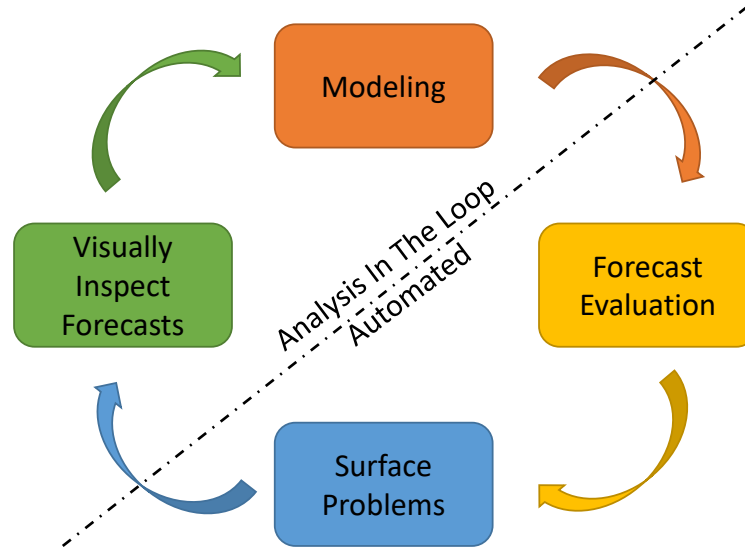$$z(t) = (z_1(t), ..., z_s(t))^T \tag{3}$$

Figure 3: Schematic view of the analyst-in-the-loop approach to forecasting at scale, which best makes use of human and automated tasks. [1]

$$\theta(t) = (\theta_1, ..., \theta_s)^T \tag{4}$$

$$\gamma(t) = (\gamma_1, ..., \gamma_s)^T \tag{5}$$

where H(t) denotes the model capacity, k denotes the growth rate, and, θ denotes the amount of change in the growth rate. As t increases, a(t) converges to H(t). When using Prophet's growth = 'logistic', the value of H(t) needs to be set in advance.

A model based on segmented linear functions:

$$a(t) = (k + z(t)\theta) \cdot t + (m + z(t)^T \gamma) \tag{6}$$

where k denotes the growth rate, θ denotes the change in growth rate, and m denotes the offset parameter.

The periodic function of the seasonal periodicity b(t) can be represented by a sine or cosine function. So the periodicity of the time series can be modeled using a Fourier series, for example:

$$b(t) = \sum_{i=1}^{I} (a_i cos(\frac{2\pi it}{p}) + b_i sin(\frac{2\pi it}{p})) \tag{7}$$

we can define the following equation:

$$\delta = (a_1, b_1, ..., a_I, b_I)^T \tag{8}$$

$$x(t) = [cos(\frac{2\pi 1 t}{p}), ..., sin(\frac{2\pi I t}{p})] \tag{9}$$

Therefore, we can express b(t) into vector form:

$$b(t) = x(t)\delta \tag{10}$$

For a sequence with a yearly period, p = 365.25 and I = 10, then:

$$x(t) = [cos(\frac{2\pi 1t}{365.25}), ..., sin(\frac{2\pi 10t}{365.25})] \tag{11}$$

For a sequence with a weakly period, p = 7 and I = 3, then:

$$x(t) = [cos(\frac{2\pi 1t}{7}), ..., sin(\frac{2\pi 3t}{7})] \tag{12}$$

Thus, the seasonal term of the time series is: $b(t) = x(t)\delta$, and the initialization of δ is δ = normal(0, β^2). β = seasonality prior scale, and a larger value of β indicates a more pronounced effect of seasonality.

Holiday items c(t) model can be written as:

$$c(t) = m(t)k = \sum_{i=1}^{L} *I_{l \in D_i} \tag{13}$$

$$m(t) = (l \in D_i, ..., l \in D_L) \tag{14}$$

$$k = (k_1, ..., k_L)^T \tag{15}$$

The initialization of k is k = Normal(0, t^2), t = holidays prior scale, and the default value is 10, when a larger value indicates that holidays have a greater impact on the model.

For the Change-point Selection in Prophet model, before introducing the changepoint, it is important to introduce the Laplace distribution, which has a probability density function as follow:

$$f(x|\theta, t) = exp(-|x - \theta|/t)/2t \tag{16}$$

where θ denotes the location parameter and t > 0 denotes the scale parameter. The Laplace distribution is somewhat different from the normal distribution. In the Prophet algorithm, it is necessary to give the location of the variable points, the number of them, and the rate of change of the growth. Therefore, there are three more important indicators, namely:

1. changepoint-range

2. n-changepoint

3. changepoint-prior-scale

Changepoint-range refers to the percentage of variable points that need to be set in a time series as long as the previous changepoint-range, in the default function changepoint-range = 0.8, n-changepoint refers to the number of variable points, in the default function is n-changepoint = 25. changepoint-prior-scale indicates the distribution of the growth rate of changepoints. [4]

4.1.3 Improvement Prophet - X-Prophet

The original Prophet model often tends to fall into overfitting at some special time points, which makes the error of some prediction behaviors large and makes the prediction's performance not very good. Therefore, we use XGBoost model to optimize the Prophet model, so that the optimized model can have the advantages of both models and thus reduce the prediction error.

The structure of X-Prophet is showed in4. The X-Prophet improved model is based on the original Prophet model with an additional confidence value optimization module. The function of this module is to combine the Prophet and XGBoost sub-models, and use the XGBoost model to optimize the Prophet model and reduce the model prediction error. [2]
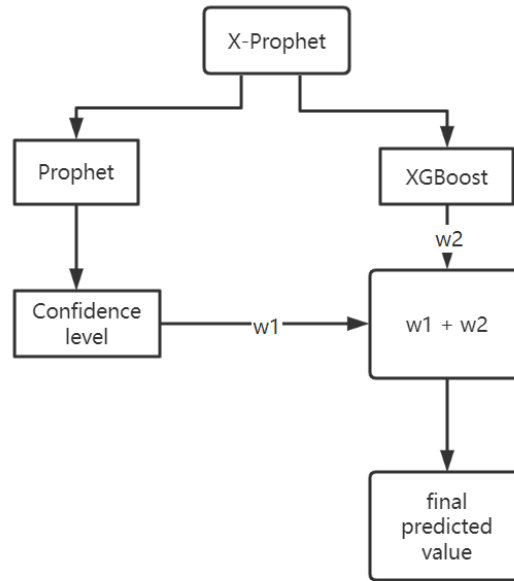

Figure 4: The structure of improvement prophet X-Prophet

The XGBoost model is a strong classifier model that can quickly and accurately mine the data for special time point features such as holidays, and it has a regularization term to avoid overfitting in the prediction behavior when analyzing time series data. Therefore, an improved model, X-Prophet, is proposed here with the addition of a predictor optimization module to the X-Prophet model, as shown in Fig. 4. The purpose is to introduce the XGBoost term in the optimization module and optimize Prophet by using the feature of XGBoost with regular terms, so that the two models can complement each other to obtain the improved X-Prophet model and finally achieve the purpose of reducing the prediction error and improving the prediction accuracy. For the X-Prophet model after the introduction of XGBoost optimization term, the model construction can be expressed as follows: [2]

$$\hat{y}_i^{(0)} = c \tag{17}$$

$$\hat{y}_i^{(1)} = \hat{y}_i^{(0)} + f_1(x_i) \tag{18}$$

$$\hat{y}_i(t) = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{19}$$

$$f_t(x) = w_{q(x)} \tag{20}$$

$$loss(f) = yT + \frac{1}{2}\gamma \sum_{j=1}^{T} w_j^2 \tag{21}$$

$$W(t) = w_1 p(t) + w_2 \hat{y}_t (w_1 + w_2 = 1) \tag{22}$$

where equation (16) - (18) indicates that the X-Prophet model analyzes the data to obtain the optimal solution at the corresponding time point. Equation (19) denotes the penalty term of the model, and the penalty term is used to control the complexity of the model. Equation (20) denotes the loss function of the model, and in each iteration, the loss function minimizes the error sum of the predicted value and the loss value, i.e., the optimal model is constructed. Equation (21) denotes the final predicted value of the model. [5]

4.2 Decision Model

We use the **Improved Adaptive Price Prediction Model for the N-Day Future Based on the 5-day SMA Operating Principle** to make decisions on buying and selling gold and bitcoin. The model is based on price forecasts for the next N days (e.g., 15) and is supplemented by the 5-day SMA method, which has been developed by expert research and is consistent with price movement patterns. We use this model on a daily basis to make decisions on whether to buy or sell gold and bitcoin on that day. The results of this model rely entirely on price data from the previous day: the price forecast for the next N days is based entirely on past data; the 5-day SMA uses price data from the 5 days prior to the current day. The calculation results show that our model can make quite effective decisions to maximize returns.

4.2.1 The 5-day SMA method of decision making

The 5-day SMA [3] is the average of the last 5 days' prices and is a good indicator of the recent price situation. Compare the price of the day with the 5-day SMA on the day, if the price of the day is higher than the 5-day SMA price, it means that the price of the day is higher than the average of the last 5 days, that is, the price is in an upward trend and will maintain the possibility of rising, the decision is optimistic. However, at this point, it is necessary to further combine the "5-day deviation" to finally determine whether to buy or not to buy. This is because the current price above the 5-day SMA may be much higher than the 5-day SMA and may appear to have risen significantly. According to the laws of the market, when the price rises too much, its possibility of falling also increases greatly. The "5-day deviation" is used to characterize the ratio of the current price above the 5-day SMA. If this indicator is higher than the threshold, it means that the current price has risen too much and has a high probability of falling, so it is not recommended to buy, but to sell. If this indicator is not high, it is recommended to buy. Conversely, if the current stock price is below the current day's 5-day SMA value, it indicates that the current day's price is lower than the recent price and is in and maintains a high probability of a downtrend, so buying is not recommended and selling is suggested. Using the past 5-day price trend combined with a forecast of future prices can also be a good way to make decisions. If the price has fallen in the last 5 days and will rise in the future forecast data, a buy is recommended.

4.2.2   Decision-making method based on adaptive price forecasts for the next N days

Previously, we used the X-Prophet model to make effective forecasts of future price changes based on previous data. We take the data for the next N days as the basis for the current decision. In our model, N is an adaptive parameter with an initial value set to 15. The meaning of N is that we need to look at the trend changes for the next N days when making decisions for the current day. If N is too small, too little future forecast data will be selected and the best time to buy and sell will not be found more accurately; if N is too large, too much future forecast data will be selected and the accuracy of the forecast data values further away from the current day will be lower, again affecting the accuracy of the decision. During the working of the model, N is automatically adjusted moderately according to the style of price movement it predicts. The ideal value of N is such that only one peak or trough occurs in the forecast data for the next N days.

The model analyzes and identifies the trends in the forecast data for the next N days, which are broadly classified into the following categories:

1. Monotonically increasing

2. Monotonically decreasing

3. Monotonic increasing curve first

4. Single-valley curves first decreasing

5. Multi-peak and multi-valley increasing

6. Multi-peak and multi-valley first decreasing

For case 3 and 4, there is only one peak or trough, which means that the N value is ideal. At this point, only a decision needs to be made: for case 3, the model recommends a buy operation. If a buy has been made previously, the model's recommendation will be ignored. For case 4, the model recommends a sell operation. The model's recommendation is ignored if it has already sold previously. The absence of peaks or troughs in case 1 and 2 indicates that N takes too small a value. The model will multiply N by a factor α to increase it appropriately after making a buy and sell recommendation respectively, in our model $\alpha = 1.1$. Case 5 and 6 shows multiple peaks and valleys, indicating that the value of N is taken too small. The model will be appropriately reduced for both N multiplied by a scale factor β after making a buy and sell recommendation, respectively. In our model, $\beta = 0.9$.

5   Implementation

5.1   Data Preprocessing and Testing

Because of the special characteristics of time series (pattern correlation between adjacent series, data are continuously generated in the time dimension), their normalization methods are also different from those of non-time series. Usually, time series analysis is based on the discovery of anomalous patterns or outliers. The choice of normalization method should also facilitate the work of subsequent algorithms or models as much as possible.

5.1.1 Data Normalization

Since the given data (*BCHAIN-MKPRU.csv* and *LBMA-GOLD.csv*) is relatively clean and complete, operations such as denoising are not needed, work is mainly focused on normalization processing, i.e., mapping all the data to a specified range (mapping data to a range of 0 to 1 or -1 to 1 for example) for processing. In this paper, **Min-Max normalization** is applied, which can be formulated by following formula:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{23}$$

The data is mapped to a specified range for processing, which is more convenient and fast. Turning dimensioned expressions into dimensionless expressions facilitates indicators of different units or magnitudes to be able to be compared and weighted. After normalization, the dimensioned data set is turned into a pure quantity, which can also achieve the effect of simplifying the calculation.
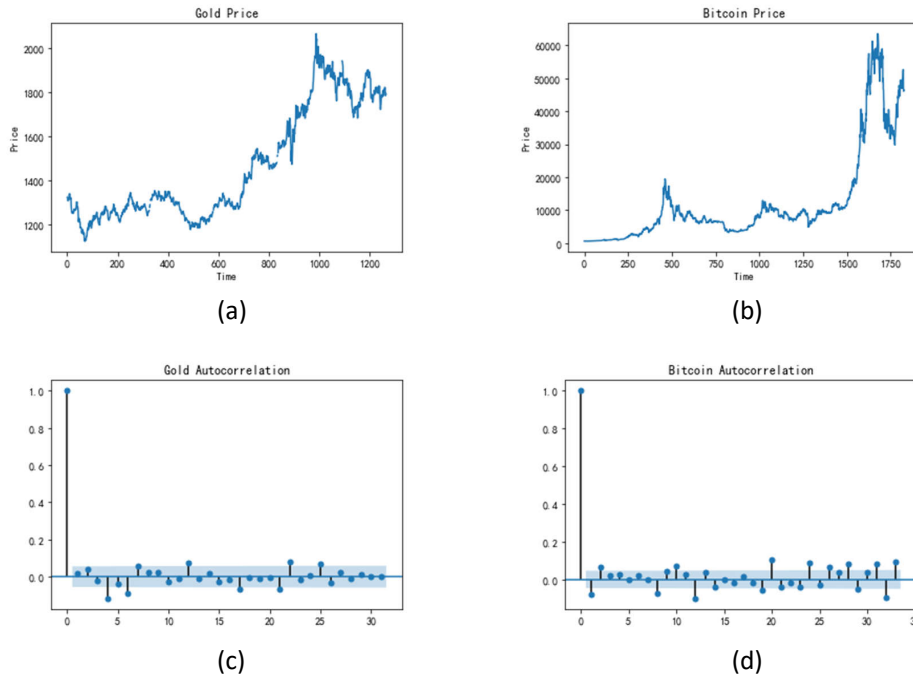
5.1.2 Stability Testing

Figure 5: (a) and (b) plot the price of gold and bitcoin. (c) and (d) plot the correlation of gold and bitcoin. (a) - (d) show that the time series needed predicting is a relatively stable sequence.

There are two tests for the smoothness of time series: the graph test and the unit root test, respectively.

Time-series plot   According to the statistical characteristics of a smooth time series, the time series plot of a smooth series always fluctuates randomly around a constant value, and the fluctuation range

is bounded. If the series shows a monotonic trend and periodicity, then it is non-stationary.

Autocorrelation plot   Smoothed series usually have short-term correlation, i.e., the autocorrelation coefficient decays quickly toward zero as the number of delay periods increases, and conversely the correlation coefficient of a non-smoothed series usually decays toward zero with a lower speed.

As shown in Fig. 5, (a) and (b) plot the price of gold and bitcoin. (c) and (d) plot the correlation of gold and bitcoin. (a) - (d) show that the time series needed predicting is a relatively stable sequence.

5.1.3   Pure Randomness Testing

The absence of any correlation between sequence values means that the sequence is a sequence with no memory, i.e., past behavior has no effect on the future, and such a sequence is called a purely random sequence such as white noise sequence. The main tests for pure randomness are the Q statistic for large samples and the LB statistic for small samples.
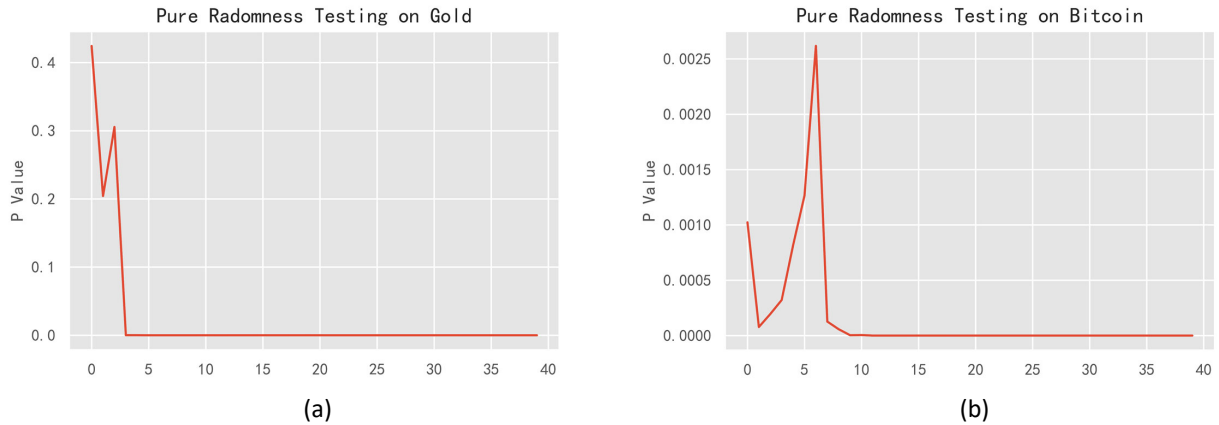

(a)   (b)

Figure 6: (a) shows the P value of the gold randomness testing, and (b) shows the P value of the bitcoin randomness testing. It can be observed that most of the P values are significantly less than 0.05, which shows that these sequence are not pure random sequences and provide evidence for the following prediction.

5.2   Prediction using X-Prophet model

The traditional time series prediction task uses a longer time series from the past to predict the change in that series over a certain period of time in the future. This requires only one training and one prediction. In contrast, this task requires forecasting price changes for each day using data from before that day. This requires training a separate model for each day's decision and predicting price changes over time as needed for the decision model. Therefore, the description of the forecasting model in the following paragraphs will be divided into two parts: single-step forecasting discussed in Sec. 5.2.1 and global forecasting discussed in Sec. 5.2.2.

5.2.1 Single Period Prediction

As the method shown in Sec. 4.1.3, Prophet is used to make predictions, its back-end system is programmed using Stan, a probabilistic programming language, which facilitates Prophet to exploit many of the advantages of Bayesian algorithms, such as Giving the model a simple and easily interpretive periodic structure. What's more, the prediction results include confidence intervals derived from the full posterior distribution, i.e., Prophet provides a data-driven risk estimate. In addition, Prophet also provides simple, easy-to-interpret forecast results for components of the time series (e.g., days of the week, or days of the year), as shown in Fig. 7.
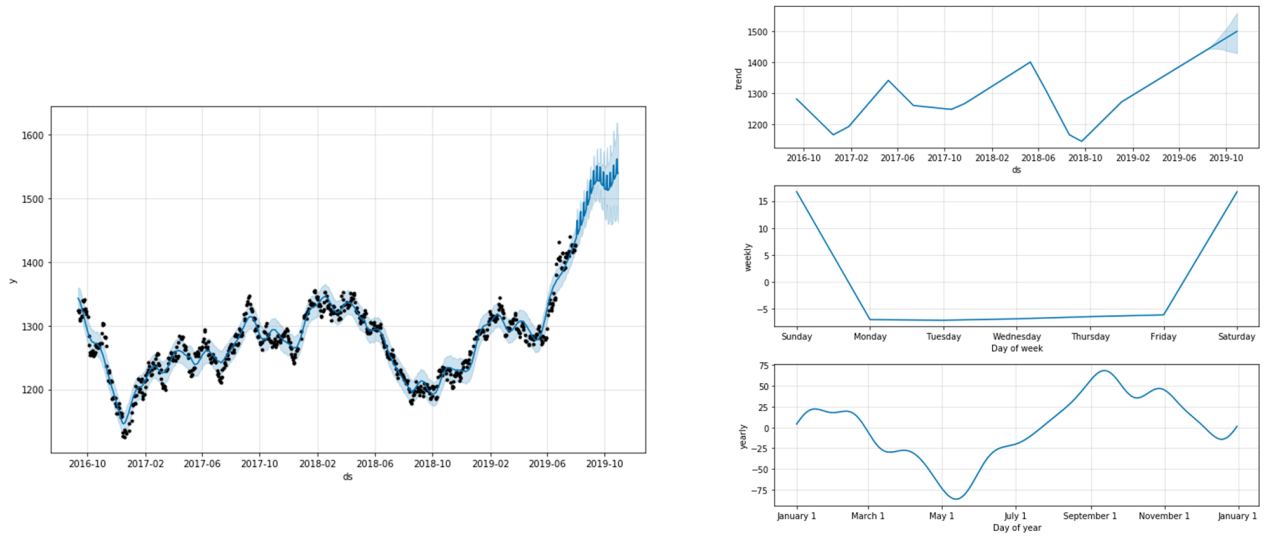

Figure 7: The visualization of a single step in our method.

Note that, based on the fast-changing nature of Bitcoin, we innovatively propose a Prophet model based on a sliding window mechanism, the basic idea of which is shown in Figure 8. For sequence data with less than one sliding window, we use all data from the beginning to the current one. If the length of the currently known time series is larger than the length of the sliding window, we use only the data of the most recent sliding window size for prediction. Interestingly, this design is highly consistent with the way real-life economics experts predict bitcoin prices, which can be seen as a combination of machine learning and human intuition. Based on the sliding-window mechanism, the accuracy of the prediction model for our predictions significantly outperforms models that use full-window data for prediction.

5.2.2 Global Forecast

In this section, the way to predict the overall gold and bitcoin price will be discussed. In fact, the global forecasting model is composed of a number of single-step forecasting models. For each day, the future data is predicted based on the data from a previous window size. A practical parameter setting is to set the window size to be 90, which means using the price of previous season to predict the future price. As shown in Fig. 9, the first day price of the prediction is plotted. It is observed that the
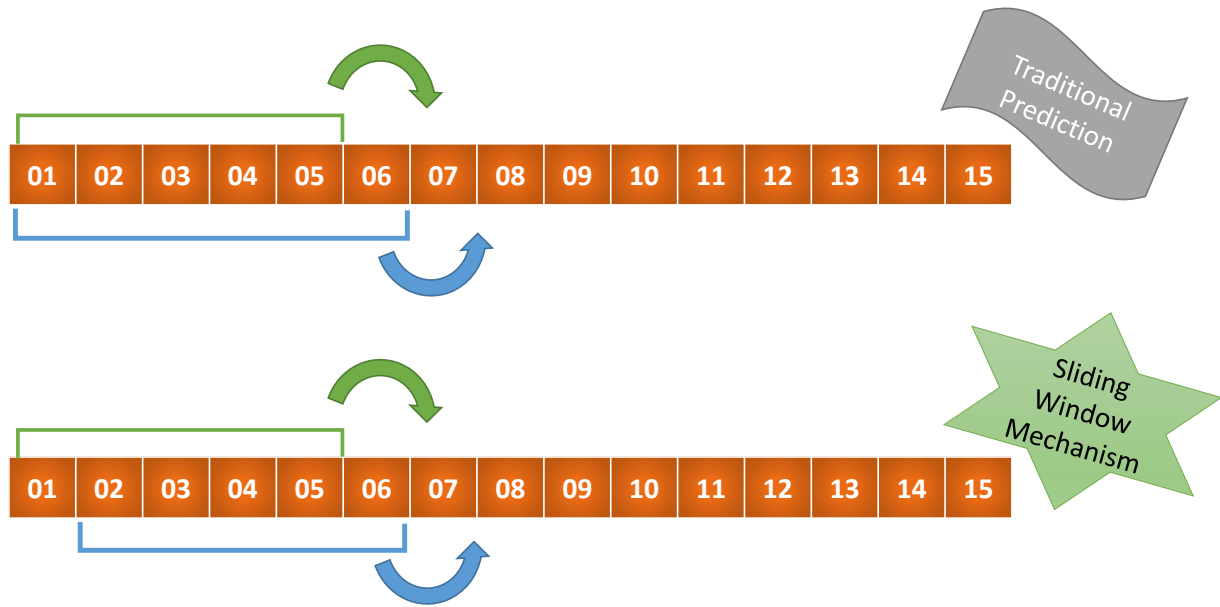
Figure 8: The illustration of the proposed sliding window mechanism of prediction model.

price of gold is less variable relative to bitcoin and has a higher prediction accuracy overall. The high intensity variation of Bitcoin, on the other hand, places a higher demand on the accuracy of the model predictions. After the predictions are made, the decision model can make buy and sell operations for bitcoin and gold based on these predicted data.

In general, our prediction models receive two limitations: the accuracy of the prediction and the generalization of the model. For other prediction models such as LSTM, it is possible to get good results on the given dataset with the tuning of hyperparameters. However, the same model and hyperparameters may perform poorly on another similar dataset, i.e., the problem of poor generalizability. In our model, on the one hand, prophet is based on a strong statistical theory and the model has a strong interpretability, which is more easily accepted by the market. On the other hand, the model has better generalization ability and can get a good result for most of the data, saving the time and effort of tuning.
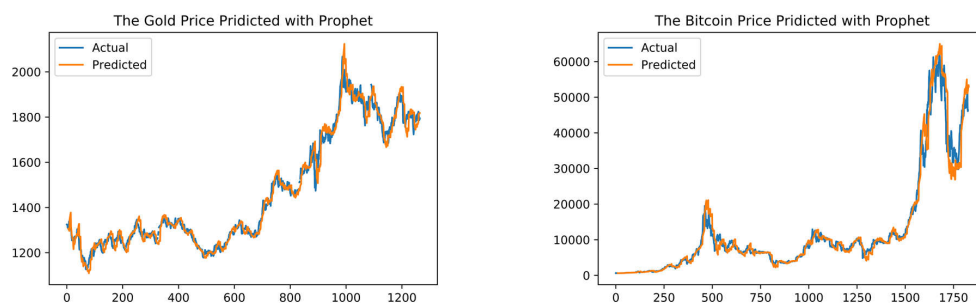

Figure 9: The final prediction of the gold and bitcoin price using X-Prophet model.

5.3   Get Best Profit by Decision Model

Based on the price data predicted by X-Prophet model, it is possible to make brilliant decisions about the day's operations. However, forecasting models cannot guarantee complete accuracy, especially when the amount of previous data is small. Therefore, it is very important to introduce the "5-day SMA" method to advise and correct the decision. When there is less amount of data available, the model prefers to base its decisions on the "5-day SMA"; when more data is available, the model prefers to base its decisions on the forecast data for the next N days.

   We apply the above model to daily decisions. Using the decision model everyday, decision can be made on buying or selling gold and bitcoin for the day. We decide to trade $\delta\%$ of the initial principal in gold and the remaining $1 - \delta\%$ in bitcoin. The return on the last day changes with δ. The variation is shown in 5.3.
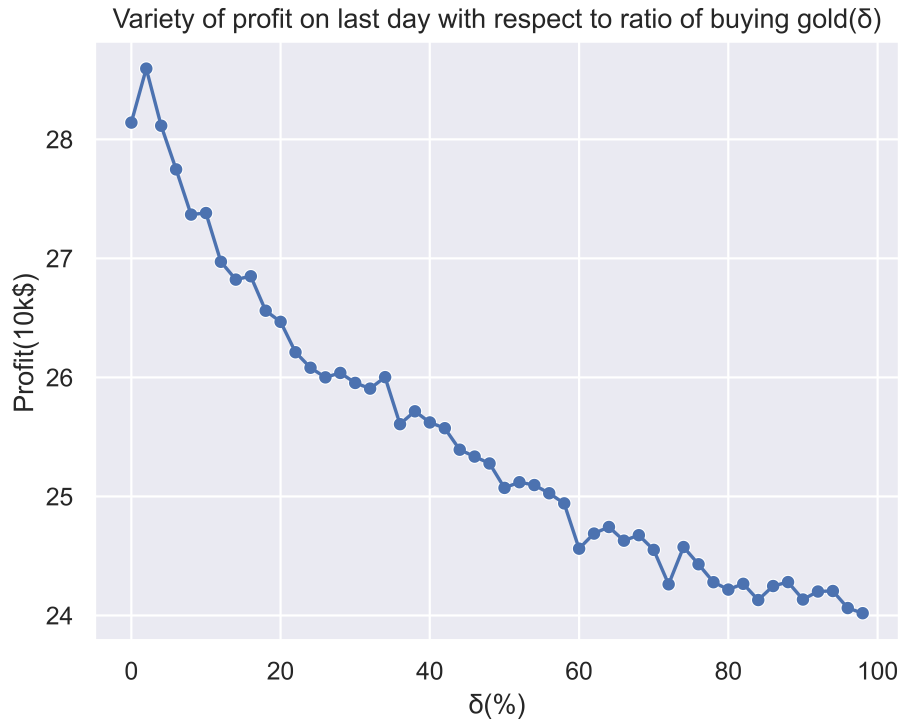

Figure 10: Variety of profit on last day with respect to radio of buying gold($\delta\%$)

   The figure shows that the final return is maximum for $\delta = 2$. Therefore, 2% of the initial principal is used for gold trading and the remaining 98% is used for bitcoin trading. Using our prediction model and decision model, with a principal of $1000, the final total assets will be $285,647.

6   Evidence of model optimality

   The proposed model can give the best strategy of trading every day due to two reasons: (1) the proposed model can always give the optimal decision; (2) the proposed model also consider about risk,
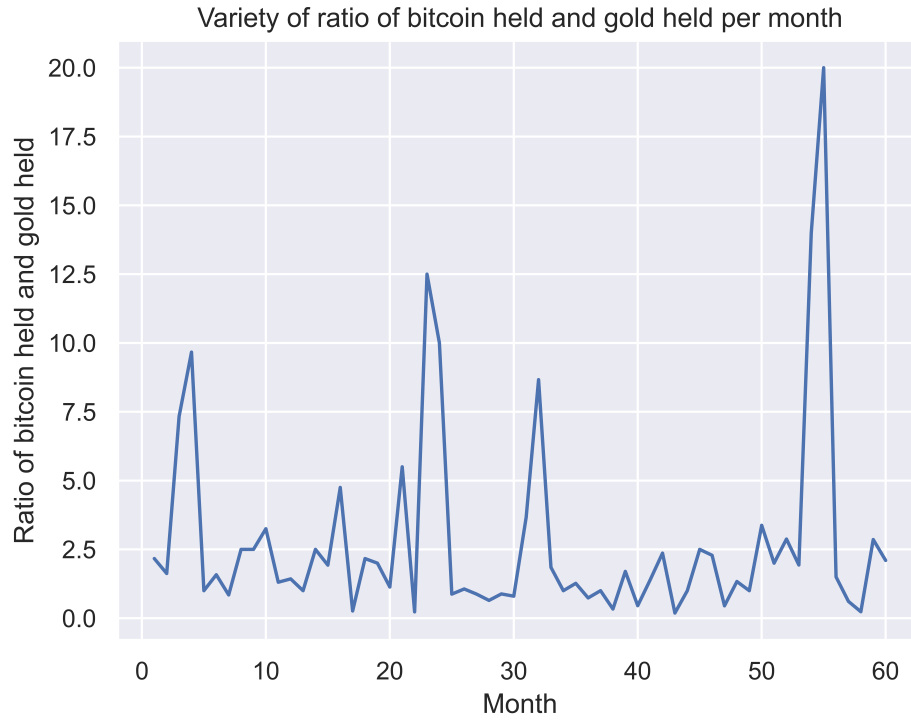
Figure 11: Variety of ratio of bitcoin held and gold held per month

making balance between profit and possible risk. Thus, the strategy given by the model is safe and excellent.

In the following few paragraphs, the evidence that the proposed model can provide the best strategy is shown by proving the above two reasons. Before that, the procedure of calculation of final profit and how the proposed model play its part in this procedure will be first stated.

6.1 Description of Final Profit Calculation Procedure

The objective of this problem is to maximize the return on the last day. Define the number of days from the current date to the first day as i, the number of day of the last day from the first day as $lastday$, initial cash $money = \$1000$, the ratio of buying gold as δ, the price of bitcoin on day i as b_i, the price of gold on day i as g_i (if gold is not traded on that day, the price on that day is the price on the nearest trading day to that day), the trading operation for gold on day j as $op_{j,g}$, and for bitcoin is $op_{j,b}$, the total assets on the jth day as P_j, the transaction cost of gold as c_g, the transaction cost of bitcoin as c_b. The final profit decision algorithm is shown in algorithm1.The goal of our model is to find an optimal policy $Policy$, such that:

$$Policy = argmax(P_{lastday}(Policy)) \qquad (24)$$

$$Profit = P_{lastday}(Policy) - money \qquad (25)$$

where the total assets on i^{th} day P_i can be calculated as:

$$P_i = MoneyForGold_i + MoneyForBitcoin_i + gold_i \times g_i + bitcoin_i \times b_i \tag{26}$$

$MoneyForGold_i$ represents the money invested in gold trading on the ith day while $MoneyForBitcoin_i$ represents the money invested in bitcoin trading on the ith day. $gold_i$ represents the amount of gold own on the ith day, while $bitcoin_i$ represents the amount of bitcoin own on the ith day.

Algorithm 1 Final Profit Decision Algorithm

Require: $i \in [0, lastday]$
Ensure: $P_{lastday}$
  $money \leftarrow \$1000$
  $P_0 \leftarrow money$
  $MoneyForGold \leftarrow money \times \delta\%$
  $MoneyForBitcoin \leftarrow money \times (1 - \delta\%)$
  $gold \leftarrow 0$
  $bitcoin \leftarrow 0$
  for $j = 1$ to $lastday$ **do**
    $op_j, g, op_j, b \leftarrow$ DecisionFromProposedModel
    if $op_{j,g} ==$ buy **and** $MoneyForGold > 0$ **then**
      $MoneyForGold = 0$
      $gold = MoneyForGold \times (1 - c_g)/g_i$
    end if
    if $op_{j,g} ==$ sold **and** $MoneyForGold == 0$ **then**
      $MoneyForGold = gold \times g_j \times (1 - c_g)$
      $gold = 0$
    end if
    if $op_{j,b} ==$ buy **and** $MoneyForBitcoin > 0$ **then**
      $MoneyForBitcoin = 0$
      $bitcoin = MoneyForBitcoin \times (1 - c_b)/b_i$
    end if
    if $op_{j,b} ==$ sold **and** $MoneyForBitcoin == 0$ **then**
      $MoneyForBitcoin = bitcoin \times b_j \times (1 - c_b)$
      $bitcoin = 0$
    end if $P_j = MoneyForBitcoin + MoneyForGold + bitcoin \times b_j + gold \times g_j$
  end for

6.2 The Proposed Model Gives Optimal Decision

As mentioned above, the X-Prophet model has high accuracy in predicting future price changes, which makes it feasible for us to use the future forecast prices to make decisions about the current operation. The higher the accuracy of the forecasting model, the greater the probability that we will make an optimal decision. However, even if our forecasting model is quite good, it still cannot achieve 100% accuracy. Therefore, we introduce the 5-day SMA method, which is proposed by expert research, to

make certain corrections to the decisions made by the forecasting model, thus making the decisions more optimal.

The **5-day SMA** is a method that has been proven to be effective by combining a large number of price curves with big data analysis and using the prices of the past 5 days to determine price movements. The **5-day SMA** is the short-term average of the SMA system, reflecting the short-term trend of the price or index. Some people also call the **5-day SMA** an attack line. Regardless of whether the short-term trend of the price or index is up or down, the steepness of the 5-day SMA's angle of operation represents the magnitude of the short-term upward or downward trend of the price or index.

Therefore, by constructing an excellent forecasting model and using empirical rules, it is possible to make a fairly accurate judgment and forecast of the price trend for the current day and beyond, and thus make scientifically accurate decisions.

6.3   The Proposed Model Considers About Risks

The price of gold is less volatile and the price of bitcoin is highly volatile. The risk of investing in gold is lower than the risk of investing in bitcoin. Although the price of bitcoin has risen much more than gold, it would be riskier to invest all the principle on bitcoin.As shown in fig.5.3,the optimal solution is not to use all of the principle to buy bitcoin, it is of high risk and earn less profit. Instead, our model select a compromise solution: use a moderate amount of principle to invest in gold and use a larger portion of the remaining principle to invest in bitcoin. This will effectively reconcile risk and return to maximize final profit.

7   Sensitive Analysis

In order to determine the sensitivity of our model against differing transaction cost, a series of sensitivity analysis is performed. In the analysis, the transaction cost of gold is taken varied from 0.00005 to 0.0030, while the transaction cost of bitcoin is taken varied from 0.0001 to 0.0030. The upper limit of transaction cost of gold and bitcoin are both taken as 0.0030, with the rule that the maximum brokerage trading commission shall not exceed 0.3% of the transaction amount.

The result of sensitivity analysis is shown in fig. 12. It can be seen that bitcoin affects the total return more significantly with the increase in transaction cost relative to gold. That's because the price of bitcoin is much higher than gold, and transaction cost can easily affect the decision of bitcoin. Therefore, bitcoin is more sensitive to transaction cost than gold.

The result of the analysis shows that, during changing transaction cost of gold in the given range, the maximum fluctuation amount of final profit is 4.86%; during changing transaction cost of bitcoin in the given range, the maximum fluctuation amount of final profit is 0.35%. Thus, as the transaction cost of gold and bitcoin changing, the total returns do not fluctuate significantly, which shows that the proposed model has a pretty good robustness in transaction cost for both gold and bitcoin.
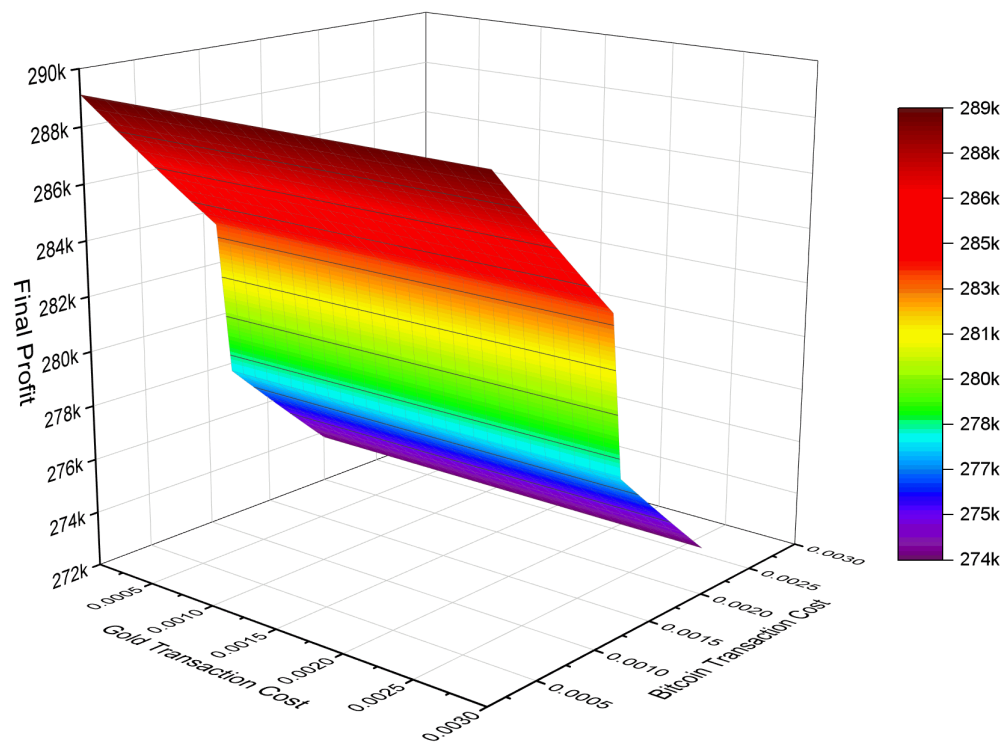
Figure 12: Sensitive Analysis on gold and bitcoin transaction.

8 Strengths and weaknesses

8.1 Strengths

1. We combine and compare a lot of methods to figure out a more effective algorithm to predict the price of gold and bitcoin, and the prediction result is accurate as you can see above.

2. We use X-Prophet model with sliding window, which is a model that combines the advantages of Prophet and XGBoost, making our prediction model result more closely to the real data.

3. Our model takes amounts of factors into consideration to obtain a universal situation, and the result shows that the performance of our model is good enough in many aspects.

4. The decision model we develop can get a pretty high total return for market trader, and we also present some evidences to demonstrate that our model can provide the best strategy.

8.2 Weaknesses

1. Since only the gold daily prices and bitcoin daily prices are available, but their prices are affected by many factors. Which may it hard to predict the prices of gold and bitcoin.

2. We do some simplifications in our model to make the calculation more easily, but it may cause additional error effects.

8.3 Further Improvements

In view of the imperfections of our model and work, we hope to complete a more comprehensive mathematical model with the help of deeper researches. Up to now, our team still have a poor knowledge in gold and bitcoin markets. We greatly hope that our model can become more perfect with the help of the deeper understanding of the relevant background knowledge. For example, we can add more factors that affect the price of gold and bitcoin to our model to make our prediction more accurate. As we had mentioned before, the research on gold and bitcoin is of great significance, therefore, we will try our best to do more research and improve the performance of our model.

9 Conclusion

In our article, we design a model that use only the past stream of daily prices up to date to determine the strategy that market trader should employ on given day. The model can determine if the trader should buy, hold, or sell their assets in their portfolio for each day, maximizing their total return. Our model can separate into two major phases, prediction model and decision model. For the prediction model, we compare a lot of algorithms, and finally use the improvement Prophet model call X-Prophet, and we also propose a Sliding Window Mechanism based on the instability of bitcoin price. For the decision model, we use some programming methods to determine the trading strategy. We also provide evidence that our model can achieve the maximum return balancing the risk. What's more, we perform a sensitivity analysis to verify the stability of the model and get practical guidance significance. However, due to the limitation of information given by the data, there is still a certain gap between the reality and our model. We'll keep trying to improve our model in the future.

References

[1] Taylor SJ, Letham B. 2017. Forecasting at scale. PeerJ Preprints 5:e3190v2 https://doi.org/10.7287/peerj.preprints.3190v2.

[2] Wu, Wenpei, Song, Yalin, Wei, Shangfei. Research on electricity consumption prediction based on improved Prophet model[J]. Computer Simulation,2021,38(11):473-478.

[3] Five-Day SMA Trading Method: https://caifuhao.eastmoney.com/news/

[4] A study of the Facebook time series prediction algorithm Prophet, https://zhuanlan.zhihu.com/p/52330017

[5] Q. M. Xia, C. H. Lan, H. J. Gu, and P. C. Guo. Research on Internet money fund based on Prophet model[C]//The winning collection of papers of 2021 (7th) National Student Statistical Modeling Competition (II). [publisher unknown],2021:2014-2051.

[6] Bitcoin on air, 57% of young Americans believe bitcoin will boost global economy in 2022, https://sv.bitcoinonair.com/57-of-young-americans-believe-bitcoin-boosts-global-economy

10　Memorandum

To: Market trader
From: Team# 2200688
Date: February 22, 2022
Subject: The best daily trading strategy of gold and bitcoin

Dear Sir or Madam:

　　Investors frequently make predictions and decisions about whether to buy and sell gold and bitcoin to maximize their total return. In our work, we have established PADRRI model in hope to provide you the best daily trading strategy based solely on price data up to that day. We're excited to write this memo to show you how our model works, then we'll recommend the best strategy for you, and finally the model's results will be presented.

　　Now we would like to show you our model, which is divided into two major phases, prediction and decision. For the prediction model, we only use the price data of gold and bitcoin up to that day to predict their future price. The prediction model is the most important procedure of our work, since the accuracy of prediction will have a direct impact on our final strategy. To improve prediction accuracy, we studied and experimented with a variety of existing algorithms such as LSTM based on neural networks, ARIMA, ARCH and so on. Finally, we discovered that Prophet performed admirably in our prediction work with higher prediction accuracy. Due to the limit of Prophet, it may lead to overfitting during some specific periods of given sequence. Therefore, we utilized an improvement model called X-Prophet, which combines the advantages of Prophet and XGBoost. What' more, based on the unstable price of bitcoin, we proposed a novel sliding window mechanism. Using the prediction model, we can predict the price trend with excellent generalization ability and interpretability.

　　For the decision model, through our market research on gold and bitcoin, we found that gold and bitcoin conform to 5-day SMA Principle of price patterns. In order to better adapt to the price trends of gold and bitcoin, we decided to combine the price prediction with 5-Day SMA Principle to make decisions on buying and selling gold and bitcoin. The results of the improved model are based solely on price data prior to the current day, i.e., the price forecast for the next N days is based entirely on past data. Evidence shows that our decision model is able to achieve optimal results when the forecasting model performs absolutely well.

　　After introducing our model, we apply the above model to day-by-day decisions to maximize the final profit. The decision model is used daily to make decisions about buying and selling gold and bitcoin for the day in conjunction with the current and previous gold and bitcoin prices. From the results mentioned in the above paper, we can see that we can maximize the final return when we use 2% of the initial principal for gold trading and the remaining 98% for bitcoin trading. As shown of the results, applying our prediction model and decision model with a principal of $1000, you will end up with total assets of $285, 647.

　　We hope that our model can help you make the best daily trading strategy and achieve maximum total return. Investment has risks, and enter the market need to be careful. Good luck with your risk aversion and gains!

　　Best wishes!

Your Sincerely,
Team#2200688

Appendix 1

```python
def five_days_sma(day, thing):
    """
    5-day SMA
    """
    five_day_avg = count_five_day_line(day, thing)
    if price[thing][day][1] <= five_day_avg:
        return 0
    else:
        if price[thing][day][1] > (1 + five_day_rate) *
            five_day_avg:
            return 0
        else:
            return 1


def gold_decision(day, period):
    """
    Decision on gold trading using forecast data in n days' time
    """
    end = min(len(gf), day + period)
    goldf = gf[day : end]
    maxp, minp = [], []
    for i in range(1, len(goldf) - 1):
        if goldf[i] >= goldf[i - 1] and goldf[i] >= goldf[i + 1]:
            maxp.append(i)
        if goldf[i] <= goldf[i - 1] and goldf[i] <= goldf[i + 1]:
            minp.append(i)

    shorten, should_buy = False, False
    if len(minp) + len(maxp) >= 2:
        shorten = True
    if maxp and not minp:
        should_buy = True
    if minp and not maxp:
        should_buy = False
    if minp and maxp:
        if minp[0] < maxp[0]:
            should_buy = False
        else:
            should_buy = True
    if len(goldf) < 2:
        return (False, shorten)
```

```python
43      if not minp and not maxp:
44          if goldf[0] > goldf[1]:
45              should_buy = False
46          else:
47              should_buy = True
48
49      return (should_buy, shorten)
50
51  def bitcoin_decision(day, period):
52      """
53      Decison on bitcoin trading using forecast data in n days' time.
54      """
55      end = min(len(bf), day + period)
56      bitf = bf[day : end]
57      maxp, minp = [], []
58      for i in range(1, len(bitf) - 1):
59          if bitf[i] >= bitf[i - 1] and bitf[i] >= bitf[i + 1]:
60              maxp.append(i)
61          if bitf[i] <= bitf[i - 1] and bitf[i] <= bitf[i + 1]:
62              minp.append(i)
63
64      shorten, should_buy = False, False
65      if len(minp) + len(maxp) >= 2:
66          shorten = True
67      if maxp and not minp:
68          should_buy = True
69      if minp and not maxp:
70          should_buy = False
71      if minp and maxp:
72          if minp[0] < maxp[0]:
73              should_buy = False
74          else:
75              should_buy = True
76      if len(bitf) < 2:
77          return (False, shorten)
78
79      if not minp and not maxp:
80          if bitf[0] > bitf[1]:
81              should_buy = False
82          else:
83              should_buy = True
84
85      return (should_buy, shorten)
```