Scientific Journal of Economics and Management Research Volume 4 Issue 8, 2022
ISSN: 2688-9323

# Predict the Prices based on LSTM-Based Portfolio Construction Model

Zhaoyang He*

Department of Chemistry and Chemical engineering, Shihezi University, Shihezi, China, 832003, China

*1539842778@qq.com

## Abstract

**As new types of investment portfolio, the returns and risks of gold and bitcoin are widely concerned. In this paper, we establish a novel LSTM-Based Portfolio Construction Model to predict the prices and then compute the portfolio weights of assets. Based on the forecasting model in the previous section to construct a portfolio model for gold and bitcoin, we built the Weights Investing Model, which derives from three traditional models, namely Minimum Variance Portfolio (MVP), Equal Risk Contribution Portfolio (ERCP), and Maximum Diversified Portfolio (MDP). We then use the resulting daily expected returns to determine whether the price of gold and bitcoin will go up, stay flat, or go down. Based on this we can determine the direction of our future investments.**

## Keywords

**Portfolio Construction; LSTM Networks; Trading Strategy.**

## 1. Introduction

### 1.1. Problem Background

Market traders often buy and sell volatile assets in order to maximize total returns. The soaring price of bitcoin, a new type of digital currency, has attracted widespread attention and due to its high value has led bitcoin to also be known as digital gold. Investing in a combination with traditional gold as a new investment strategy, Skew Analytics data shows that the correlation between bitcoin and gold prices has reached a record 70%.

The price of bitcoin and gold fluctuates over time, which makes it risky to a certain extent, especially for bitcoin, which is characterized by rapid price changes and high volatility, with price fluctuations of hundreds of dollars in a day being normal. For market traders, the question of when to buy, sell, or hold is a difficult one. How to predict and analyze the price of gold, bitcoin and the combination strategy between them is a challenge, but worth the effort. Therefore, we present in this paper a predictive price and optimal portfolio investment model.

### 1.2. Literature Review

With the proposal of neural networks, Sherstinsky [1] subsequently addressed the short-comings of the standard RNN by morphing the canonical RNN system into the more robust LSTM network through a series of extensions and embellishments. Clarke[2] examined the composition of Minimum Variance Portfolios with a focus on the analytic form and parameter values of individual security weights. Cipiloglu [3] used an LSTM model to predict prices of the stocks, then these predictions are used in the calculation of portfolio weights. Maillard[4] first defined ERC portfolios and analyze their theoretical properties, then compared the ERC with competing approaches and provide empirical illustrations, analyzed an alternative approach based on equalizing the risk contribution from the various components of the portfolio.

# 2. Assumptions

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

Gold vacancy statistics on non-trading days will not affect our forecast results.

The effect of trader sentiment on trading is not considered.

In a certain period, international political changes do not occur, and the implementation of national policies has less impact on the price of gold and bitcoin.

All the prices given by London Bullion Market and NASDAQ are true.

# 3. LSTM-Based Portfolio Construction

## 3.1. Conventional Weights Investing Approaches

We need to establish a model that allows measuring the weight of allocation between the two currencies and the impact of commissions on trading. We, therefore, constructed a portfolio strategy model that includes the following elements:

Have an accurate return evaluation and analysis, including a comparison of benchmarks and models.

Include the impact of risk on results.

Be able to comment on the impact of commissions on trading.

According to the research of Cipiloglu [3], the model we construct below must adhere to the following assumptions:

Buy from the first trading day to sell on the last trading day.

No short-selling allowed.

Sum of weights is 1.

### 3.1.1. Minimum Variance Portfolio

In standard Markowitz portfolio theory[5], the minimum-variance portfolio has the lowest risk of all possible portfolios, geometrically at the left-most tip of the efficient frontier. The vector of optimal security weights, $\boldsymbol{w} \in \mathbb{R}^m$, only depends on the security covariance matrix,

$$\boldsymbol{\Sigma} = \frac{1}{n}\boldsymbol{X}^{\top}\boldsymbol{X},\ \ \boldsymbol{\Sigma} \in \mathbb{R}^m \times \mathbb{R}^m \tag{1}$$

and not expected security returns. Herein, $\boldsymbol{X} \in \mathbb{R}^n \times \mathbb{R}^m$ denotes the vector of observations. Specifically, the optimization problem is to minimize portfolio variance,

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \sigma_p^2 = \boldsymbol{w}^{\top}\boldsymbol{\Sigma}\boldsymbol{w} \\ s.t. \quad & \textstyle\sum_{l \le w_i \le u} w_i = 1 \end{aligned} \tag{2}$$

In these expressions, $m$ and $n$ respectively refer to the number of volatile assets and observations; $l$ and $u$ are the lower and upper bounds, respectively.

### 3.1.2. Equal Risk Contribution Portfolio

Minimum Variance Portfolios generally suffer from the drawback of portfolio concentration. A simple and natural way to resolve this issue is to attribute the same weight to all the assets considered for inclusion in the portfolio. Equally weighted portfolios are widely used in practice[6] and they are efficient out of sample[7].

In addition, if all assets have the same correlation coefficient as well as identical means and variances, the Equal Risk Contribution Portfolio is the unique portfolio on the efficient frontier.

ISSN: 2688-9323

According to the classical ERCP strategy, weights are determined in such a way that each asset contributes to the portfolio risk equally,

$$\begin{aligned} &\min_{w} \quad \sum_{i=1}^{m} \sum_{j=1}^{m} [w_i(\boldsymbol{\Sigma w})_i - w_j(\boldsymbol{\Sigma w})_j]^2 \\ &\text{s.t.} \quad \sum\nolimits_{l \le w_i \le u} w_i = 1 \end{aligned} \tag{3}$$

### 3.1.3. Maximum Diversified Portfolio

It is now increasingly popular to claim that the market capitalization-weighted indices are not efficient. Several alternative empirical solutions have been suggested. In this section, the objective of this strategy is to maximize the benefits from diversification and it uses a diversification ratio to calculate[8]. The relevant expressions are as follows,

$$\begin{aligned} &\min_{w} \quad \frac{\boldsymbol{w}^{\top}\boldsymbol{\sigma}}{\sqrt{\boldsymbol{w}^{\top}\boldsymbol{\Sigma w}}} \\ &\text{s.t.} \quad \sum\nolimits_{l \le w_i \le u} w_i = 1 \end{aligned} \tag{4}$$

## 3.2. Portfolio Construction with LSTM Network

A deep neural network is one of the mainstream investment forecasting models nowadays, and its excellent performance in investment forecasting is attributed to its powerful nonlinear fitting ability. Recurrent Neural Networks (RNN) are equipped with memory functions for time-series data by introducing directed recurrent connections between neural network layers. The model used in this paper is a modeling approach based on the Long Short-Term Memory network (LSTM), which is a modified network structure of RNN.

Based on the fluctuation data of gold and bitcoin prices, we can predict the future prices of gold and bitcoin based on LSTM.

### 3.2.1. LSTM Networks

LSTM networks are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hochreiter and Schmidhuber [9], and were refined and popularized by many people in the following work. They work tremendously well on a large variety of problems and are now widely used.

LSTM networks are explicitly designed to avoid the long-term dependency problems, and remembering information for long periods is practically their default behavior. LSTM networks have the chain structure like RNN, but the repeating module has a different structure. Instead of having a single neural network layer, there are four layers interacting in a very special way.

### 3.2.2. The Core Idea Behind LSTM Networks

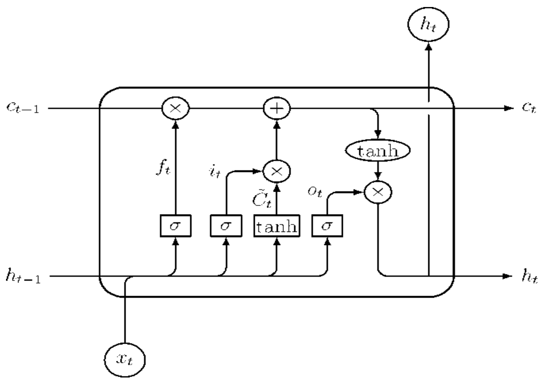


**Figure 1.** Illustration of the repeating module in LSTM networks

The key to LSTM networks is the cell state, the horizontal line running through the top of the diagram. The cell state is kind of like a conveyor belt. It runs straight down the entire chain, with only some minor linear interactions. It is very easy for information to just flow along it unchanged.

The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates. Gates are a way to optionally let information through. They are composed out of a sigmoid neural net layer and a pointwise multiplication operation. The sigmoid layer outputs numbers between zero and one, describing how much of each component should be let through. A value of zero means we should let nothing through, while a value of one means we should let everything through.

An LSTM has three of these gates, to protect and control the cell state.

### 3.2.3. Step-by-Step LSTM Walk Through

The first step in our LSTM is to decide what information we are going to throw away from the cell state.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \tag{5}$$

The next step is to decide what new information we are going to store in the cell state. This has two parts.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{6}$$

$$\vec{C_t} = \tanh\,(\mathrm{W_c} \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{b_c}) \tag{7}$$

The previous steps already decided what to do, we just need to actually do it. We multiply the old state by ft, forgetting the things we decided to forget earlier.

$$C_t = f_t \times C_{t-1} + i_t \times \vec{C_t} \tag{8}$$

Finally, we need to decide what we are going to output.

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \tag{9}$$

$$h_t = o_t \times \tanh\,(C_t) \tag{10}$$

In summary, LSTM neural network is composed of memory units and control gate units, and control gate units enable cell state updates. At the same time, three gates are set in this cell, namely the forget gate, the Input gate, and the Output gate.

## 3.3. Trading Amount Model

Based on the portfolio strategy that tells us the weighting of the amount we should allocate to bitcoin and gold each day. However, due to commission and the fact that we only have a limited amount of cash on hand, the actual amount of exchange that takes place between the bitcoin and gold investments is very complicated to calculate exactly how much to allocate to bitcoin and gold so that the value is exactly equal to the weighting percentage.

The principles for determining the supplemental value of the weights are:

There is no sign change after the weight change plus the weight supplement value, which means that it cannot change from positive to negative, and from negative to positive.

based on the previous item, it is as large as possible, and generally, we take it as 5 times the corresponding order of magnitude that does not change the sign.

At the same time, the prediction model based on LSTM will give the predicted closing price of each day. this value is compared with the forecast closing price cB on the day t to obtain the daily forecast increase. This is represented by the following formula. For gold, just replace all subscripts with G.

$$r^t = (c^t - c^{t-1})/c^{t-1} \tag{11}$$

After getting the change in the allocated amount minus the corresponding transaction costs, we can get the daily settlement amount through the increase. In summary, its mathematical expression is expressed in the following form:

$$\Delta_B^t = (B^{t-1} + G^{t-1})(w_B^t - w_B^{t-1} + w_C) \tag{12}$$

$$\Delta_B^t = (B^{t-1} + G^{t-1})[1 - (w_B^t - w_B^{t-1} + w_C)] \tag{13}$$

$$C_B^t = \Delta_B^t \alpha_B + (\Delta_B^t - \Delta_B^t \alpha_B)\alpha_B \tag{14}$$

$$C_g^t = \Delta_G^t \alpha_G + (\Delta_G^t - \Delta_G^t \alpha_G)\alpha_G \tag{15}$$

$$B^t = (B^{t-1} + \Delta_B^t - C_B^t) r_B^t + B^{t-1} \tag{16}$$

$$G^t = (G^{t-1} + \Delta_G^t - C_G^t) r_G^t + G^{t-1} \tag{17}$$

At the same time, for the first day, since there is no settlement amount for bitcoin and gold, we must define an initial state for it. Looking at the forecast data, we can see that on the first day, bitcoin is rising, while gold on non-trading days, so we will buy bitcoin with the full initial amount.

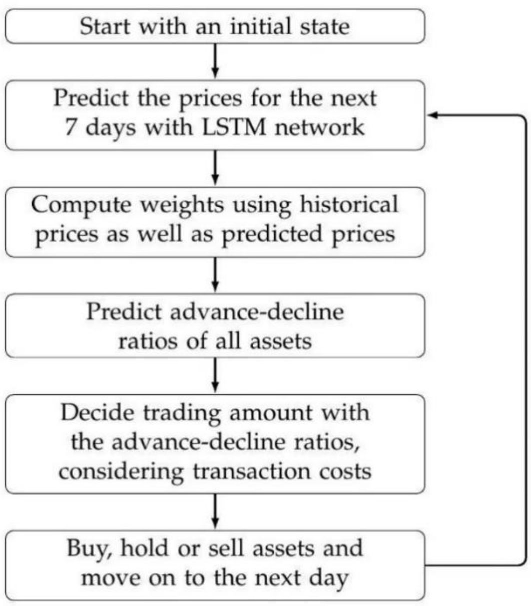


**Figure 2.** Illustration of schematic overview

## 3.4. Model Evaluations

In finance, the Sharpe ratio (also known as the Sharpe index, the Sharpe measure, and the reward-to-variability ratio) measures the performance of an investment such as a security or portfolio compared to a risk-free asset, after adjusting for its risk. It is defined as the difference between the returns of the investment and the risk-free return, divided by the standard deviation of the investment returns. It represents the additional amount of return that an investor receives per unit of increase in risk, and it can serve as a key indicator for our evaluation of the portfolio strategies generated by the above construction model. It is defined as follows,

ISSN: 2688-9323

$$S = \frac{E(R_P) - R_f}{\sigma_P} \tag{18}$$

where $E(RP)$ denotes the expected annualized payoff of the portfolio, $Rf$ denotes the annualized risk-free rate, and $\sigma p$ denotes the standard deviation of the annualized payoff of the portfolio.

The purpose is to calculate how much excess payoff the portfolio will generate for each unit of total risk taken. If it is greater than 1, it means that the investment payoff is greater than the volatility risk; if it is less than 1, it means that the investment operational risk is greater than the payoff. In this way, each portfolio can calculate the Sharpe Ratio, which is the ratio of investment return to excess risk, the higher the ratio, the better the portfolio.

At the same time, due to the volatility in price movements, especially in practice there are some significant ups and downs, and for new currencies like bitcoin the above situation is even more pronounced, in order to measure the degree of impact of this situation in our chosen portfolio strategy model and further define the magnitude of risk. We introduce the MDD measure, whose value is the difference from a peak to a minimum, representing the maximum level of loss. Although the MDD measurement is not a general performance measure, it is necessary in our evaluation due to the specificity of this new currency.

## References


[1] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," Physical D: Nonlinear Phenomena, vol. 404, p. 132306, 2020.

[2] R. Clarke, H. De Silva, and S. Thorley, "Minimum-variance portfolio composition," The Journal of Portfolio Management, vol. 37, no. 2, pp. 31-45, 2011.

[3] Z. Cipiloglu Yildiz and S. B. Yildiz, "A portfolio construction framework using Istm- based stock markets forecasting," International Journal of Finance & Economics, 2020.

[4] S. Maillard, T. Roncalli, and J. Teiletche, "The properties of equally weighted risk contribution portfolios," The Journal of Portfolio Management, vol. 36, no. 4, pp. 60-70, 2010.

[5] H. M. Markowitz, "Portfolio selection, 1959," Journal of Finance, vol. 7, p. 7791,1952.

[6] S. Benartzi and R. H. Thaler, "Naive diversification strategies in defined contribution saving plans," American economic review, vol. 91, no. 1, pp. 79-98, 2001.

[7] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?," The review of Financial studies, vol. 22, no. 5, pp. 1915-1953, 2009.

[8] Y. Choueifaty and Y. Coignard, "Toward maximum diversification," The Journal of Portfolio Management, vol. 35, no. 1, pp. 40-51, 2008.

[9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.