

Optimization of Daily Trading Strategy

Yuxin Li[#], Ziheng Yu^{*, #}, Ziqing Lian[#]

Department of Mathematics, Lancaster University, Lancaster, United Kingdom

*Corresponding author: yuz11@lancaster.ac.uk

[#]These authors contributed equally.

Abstract. This study proposes a novel method that integrates variational pattern decomposition (VMD), long short-term memory (LSTM) models, and particle swarm optimization (PSO) to optimize daily trading strategies for a portfolio. By analyzing historical price data for gold and bitcoin, we have successfully identified the optimal investment actions (buy, hold, or sell) for a portfolio comprising different assets. The VMD-LSTM model proposed by us mitigates the long-term dependence problem of machine learning algorithm RNN in financial time series prediction through batch normalization optimization, and realizes the accurate prediction of the future price of gold and bitcoin, whose R-squared values reach 0.99 and 0.98 respectively. We also applied the particle swarm optimization algorithm to a sequence of trading decisions generated by predicting prices, ultimately obtaining a maximum portfolio value of \$20,560,265.5, which is more than 14.2 times the 10-year growth of Bitcoin and 12,310.5 times the 10-year growth of gold. By verifying the universality and local optimal properties of our model, we use the Sharpe ratio as an objective function to evaluate the effect of the portfolio. After several perturbation tests, the results show that the original scheme is superior to other perturbation schemes considering the Sharpe ratio and the final return, which verifies the validity of the investment scheme selected by us as the local optimal value. These findings make our research more coherent and provide investors with a reliable basis for making decisions to achieve long-term capital appreciation and risk control.

Keywords: Investment portfolio, VMD, LSTM, PSO, Machine learning decision

1. Introduction

1.1. Background

The rise of virtual currency opened up a new field of investment for investors. As a kind of decentralized, and consensus based on peer-to-peer network mechanism of encryption currency, the currency has gained from the initial obscurity global recognition and attention [1]. Meanwhile, gold as a traditional safe havens, always plays a crucial role in determining the allocation of assets [2]. Gold and currency makes sense as a portfolio, because they represent the two different types of assets, provides a variety and diversified investment options. Gold did well in the economic uncertainties, usually seen as a refuge for preserving capital [3-4], and the currency is a kind of digital assets, has high potential returns and risks [5]. Combine them in a portfolio to balance the different types of risk, and the pursuit of better returns [6]. To fully leverage the benefits of the investment potential of gold and currency, we will develop a simple mathematical model, based on their past only price, provide daily optimal trading strategy. Through the analysis of gold and currency movements of prices, we will use mathematical model to predict the future price, and according to the forecast results generated sequence of daily trading decisions. This model will help investors make informed investment decisions, to maximize their profit potential. Such portfolio strategy can provide investors with more choices, and better asset allocation under different market conditions.

1.2. Our Model

This research aims to develop an optimized daily trading strategy by combining VMD with LSTM(Long short-term memory) model and global optimization techniques. The objective is to determine whether to make decisions on whether to purchase, retain, or divest assets within a portfolio, taking into account. The study focused on historical price data from 9/11/2016 to 9/10/2021, and used

historical data before them as training material. The VMD-LSTM model with batch normalization optimization is used to forecast the future price of gold and bitcoin. The predicted prices are then used to generate a trading decision sequence. To optimize the overall performance, PSO(Particle swarm optimization) is applied to the generated trading decision sequence, considering risk minimization and profit maximization. The results are analyzed based on portfolio value, and compared against benchmark strategies. The results indicate the efficacy of the suggested method in enhancing the daily trading strategy, outperforming traditional strategies in terms of portfolio value.

2. Establishment process

2.1. Data Processing

For the transactions of gold can only be performed when the market is open, the dates of weekends and national holidays from the data files lack the specific prices for gold. This incompleteness of the data would cause inconsistencies during data analysis and potentially break our model. As a collective agreement in the financial market, and the observation that the price of gold remains unchanged during market closure, we fill these blanks with the most recent prior price. For instance, we allocate the data from the nearest Friday to Saturdays and Sundays and 23rd December to Christmas.

Furthermore, the variations in measurements between Bitcoin and gold can have an impact on the outcomes of data analysis. Therefore, to mitigate such discrepancies and improve comparability, we employ a normalization process for both datasets. Additionally, when considering the prices of Bitcoin and GOLD on Day N, we replace the actual prices with the ratio of prices between Day N and Day N-1 as inputs to our model, aiming to achieve more meaningful results.

2.2. Predictive Models

2.2.1 VMD

VMD is an optimization-based, non-recursive adaptive signal decomposition method designed for effectively handling nonlinear and non-stationary signals. It offers a suitable approach to decompose such signals into their constituent components [7]. The core of VMD is to construct, analyze and solve the variational problem. The decomposition of the original sequence into K subcolumns is as follows: 1. To determine the unidirectional spectrum of each subcomponent, the Hilbert transform is employed to calculate the analytical signal for each component individually. 2. For each subcomponent, the spectrum is adjusted to the corresponding baseband according to its center frequency. 3. The signal is mediated according to the Gaussian smoothness, the frequency range of the decomposition model is estimated, and The variational problem aims to minimize the total estimated bandwidth of the model, while ensuring that the sum of all modes matches the original signal. This constraint ensures that the combined contribution of the modes accurately represents the original signal while minimizing the collective bandwidth of the subcomponents. The mathematical representation is as follows

$$\left\{ \begin{array}{l} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \sigma_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_k u_k = f \end{array} \right. \quad (1)$$

Where u_k and ω_k are the KTH mode component and its center frequency, respectively. δ_t is the unit impulse function

2.2.2 Long- Short Term Memory

LSTM is a type of recurrent neural network (RNN); thus, it is suitable for extracting time features from time series. It could solve the problem with long term time series [8]. As shown in the figure below on the structure of LSTM. The LSTM consists of three gate groups, there is the input gate i_t ,

the forget gate f_t , and the output gate O_t . The Forget Gate determines how much of the stage of units in the last period C_{t-1} is stored to the current period C_t ; the Input Gate determines how much of the current input X_t is preserved in current unit C_t . The Output Gate determines how much of the stage in current unit C_t is transmitted to the current output h_t . As shown in Figure 1.

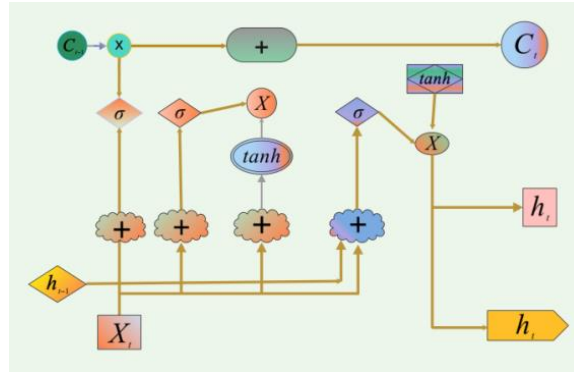


Figure 1. The Structure of LSTM

The functions of the three gates, long-term memory, and short-term memory are as follows:

$$\text{Input Gate: } i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\text{Forgot Gate: } \begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \end{cases} \quad (3)$$

$$\text{Output Gate: } O_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (4)$$

$$\text{Long-term Memory: } C_t = f_t * C_{t-1} + i_t * C_t \quad (5)$$

$$\text{Short-term Memory: } h_t = O_t * \tanh(C_t) \quad (6)$$

2.2.3 VMD-LSTM

Initially, we employ VMD to perform a decomposition of the time series data, resulting in the generation of a collection of intrinsic mode functions (IMFs). These IMFs help reveal the underlying structure of the data. The LSTM model is then utilized to predict future prices, with each IMF=5 serving as a separate dataset. The training set comprises the first 700 days, and the test set covers the remaining 300 days. In this way, the data over ten years are gradually predicted. All input and output data are normalized for improved training efficiency. The LSTM model consists of two hidden layers with 50 and 80 neurons, respectively. We set a maximum of 200 training cycles, a learning rate of 0.01, and reduce the learning rate by 30% every 50 cycles. After training, the model is used to make predictions on the test data and future data. The process is shown in Figure 2.

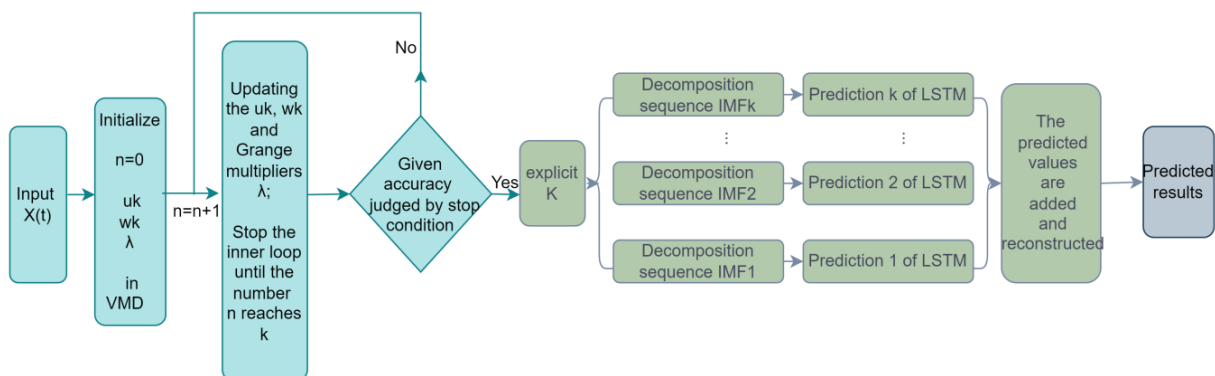
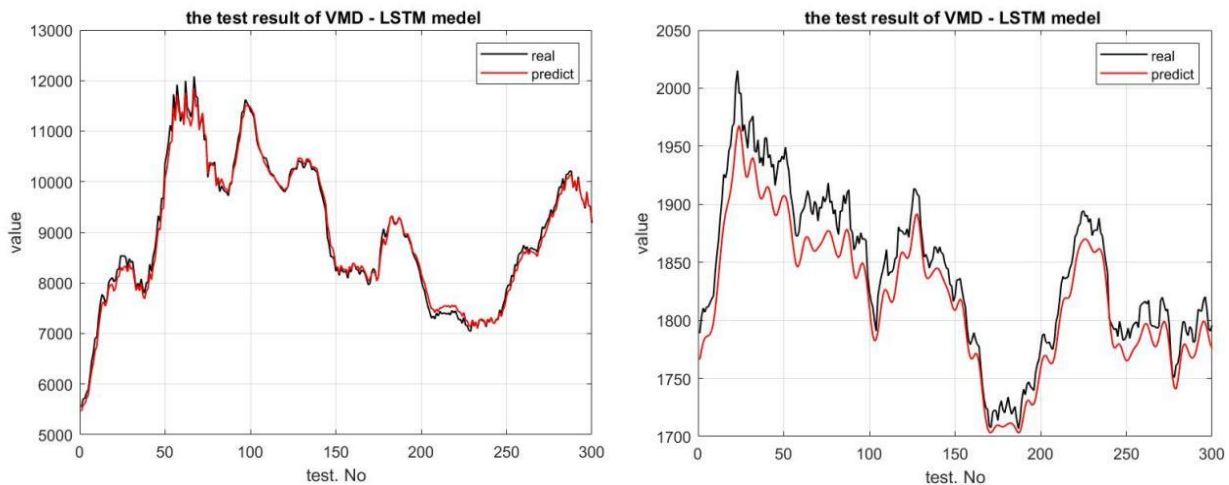


Figure 2. Procedure of the VMD-LSTM algorithm

2.2.4 predicted results



(a) The Test Result of GOLD in Last 300 days (right) (b)The Test Result of Bitcoin in Last 300 days (left)

Figure 3. The Test Result of VMD-LSTM Model

The predicted results are visualized in Figure 3. We utilize the R-square coefficient as a metric to assess the model's predictive performance, because it measures the degree to which the model explains the variance of the observed data. In other words, the R-square coefficient represents the proportion of the dependent variable's variability that can be explained by the model, relative to the total variability. The R-square coefficient ranges from 0 to 1, where a value of 0 for the R-square coefficient suggests that the model is unable to explain the variability observed in the dependent variable, and 1 indicates that the model fully explains the variability of the dependent variable. Typically, a higher R-square coefficient indicates a stronger predictive performance of the model, the better the predictive performance of the model. Here, the R-square results of gold and bitcoin are 0.99 and 0.98 respectively, close to 1, indicating that our model has excellent predictive performance.

2.3. Transaction model

2.3.1 Model assumption and Planning model building

We employ dynamic programming to revise the daily trading strategy and optimize future profits. Then we consider that the allocations of funds, gold holdings, and bitcoin holdings within the overall investment portfolio on day k are represented by $[c_k, g_k, b_k]$, subject to the equation $c_k + g_k + b_k = 1$. The decision variables $[\Delta G_k, \Delta B_k]$ are the changes in gold and bitcoin holdings. The increases in cash, gold, and bitcoin on day k are denoted as $[0\%, G_k \%, B_k \%]$, based on the prediction results. We adjust the proportions of gold and bitcoin to maximize future profits while keeping the total asset unchanged. At the end of each day, we settle the gains or losses. The increments in cash, gold, and bitcoin lead to new proportions $[c_k, (1 + G_k \%)g_k, (1 + B_k \%)b_k]$. However, due to the change in the total amount, the sum of these proportions may no longer be equal to 1. Therefore, we normalize the proportions after all changes, allowing us to proceed with the operations for the next day.

Next, we execute trades, adjust positions, and settle transaction costs. Transaction costs affect the change in the cash proportion. We assume transaction costs of 0.01 for gold trades and 0.02 for bitcoin trades. After settling the transaction costs, we determine the new proportions of cash, gold, and bitcoin in the total asset, i.e.

$$[c_k - (1 \pm 0.01) \Delta G_k - (1 \pm 0.02) \Delta B_k, (1 + G_k \%)g_k + \Delta G_k, (1 + B_k \%)b_k + \Delta B_k] \quad (7)$$

Due to this planning model is established to update the trading strategy based on historical data and maximize future profits. The constraints require the proportions of cash, gold, and bitcoin in the total asset to be greater than or equal to 0, i.e.

$$\begin{cases} c_k + (1 \pm 0.01)\Delta G_k + (1 \pm 0.02)\Delta B_k \geq 0 \\ (1 + G_k \%)g_k - \Delta G_k \geq 0 \\ (1 + B_k \%)b_k - \Delta B_k \geq 0 \end{cases} \quad (8)$$

2.3.2 Programming model solving

During the model building process, we set up an objective function $f(k)$ that requires us to adjust our positions that day and remain unchanged for the next three days. In reality, however, fixed positions can be problematic for day-to-day decision making due to frequent fluctuations in the market's ups and downs. Therefore, we need to extend the decision-making scope from the three days of the current position to the global position. However, the global position strategy involves many decision variables and the scope is not clear, which leads to the problem of slow convergence of the model. To address this problem, we adopt the particle swarm optimization algorithm (PSO) [9] to enhance the convergence speed of the model. By optimizing the programming model under appropriate initial conditions, the convergence speed can be significantly improved. Assuming we have 6 decision variables on day k , representing the changes in gold and bitcoin on day N , $N+1$, and $N+2$, this implies a six-dimensional particle swarm operation. In particle swarm optimization, each particle is represented by a position vector $x_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ and a velocity vector $v_i = (v_{i1}, v_{i2}, \dots, v_{iN})$, existing in a multidimensional space without mass or volume. We assume an initial particle count of 100. Each particle has a moderate value determined by its objective function f_k , which guides the particle in finding its optimal position $p_i = (p_{i1}, p_{i2}, \dots, p_{iN})$ and the optimal position of neighboring particles $g_i = (g_{i1}, g_{i2}, \dots, g_{iN})$. By continuously moving, updating, and iterating, particles obtain new position parameters, following velocity and position update formulas.

$$v_{ij}(t + 1) = \omega v_{ij}(t) + c_1 r_1 (P_{iN}(t) - x_{iN}(t)) + c_2 r_2 (g_{iN}(t) - x_{iN}(t)) \quad (9)$$

$$x_{ij}(t + 1) = \omega v_{ij}(t) + v_{ij}(t + 1) \quad (10)$$

Where: ω is the inertial parameter, that is $\omega > 1$ is acceleration of particle motion and $\omega < 1$ is deceleration of particle motion, the value ranges from 0.4 to 0.9. c_1, c_2 is the learning parameter, generally $c_1 = c_2 = 2$; r_1, r_2 is a random number of $[0, 1]$. The position parameter $x_{ij}(t + 1)$ is the i particle iterating $t + 1$ times in the j TH dimension, and the velocity parameter $v_{ij}(t + 1)$ of the i particle iterating $t + 1$ times in the j TH.

2.3.3 Dynamic programming model results

In the end, the total value of the investment with the goal of maximizing profit is 20,560,265.5, which is 14.2% higher than the growth of the value of Bitcoin itself over the decade, that is, 18,000,000, and 12310.5% higher than the growth of gold over the decade.

The daily trading strategy diagram is shown in Figure 4, and the inter-time trading strategy diagram is shown in Figure 5. And the wealth value obtained from each decision as shown in Figure 6.

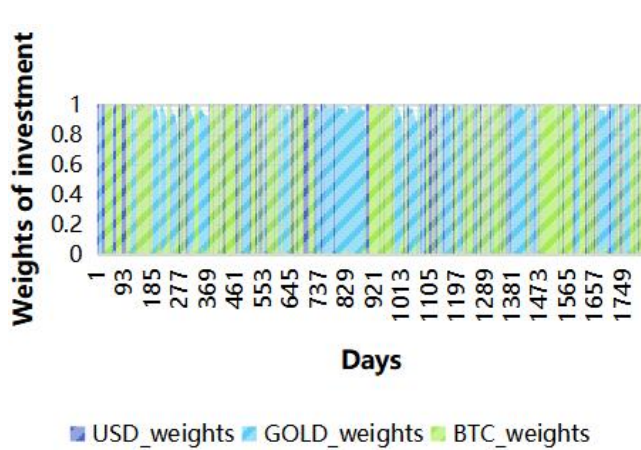


Figure 4. Daily investment strategy (left)

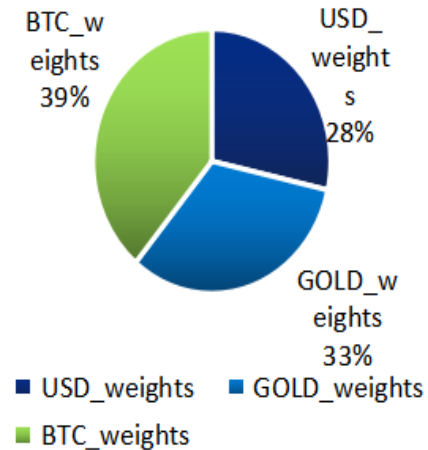


Figure 5. Weights of USD,BTC,GOLD in the overall investment strategy (right)

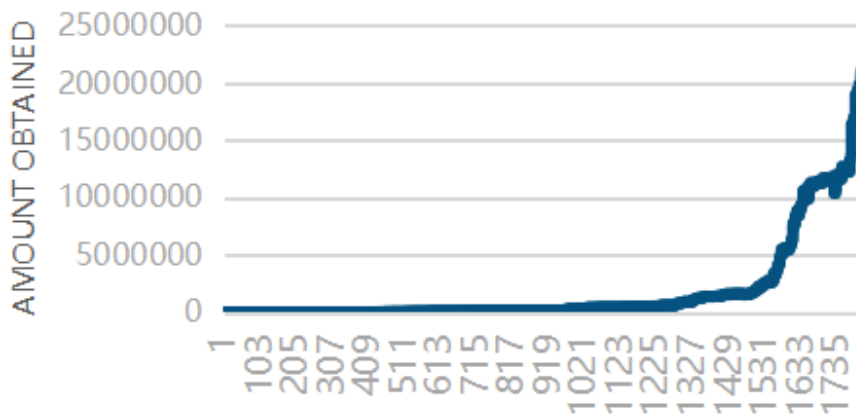


Figure 6. The wealth value of daily decisions

3. Verification of sensitivity and universality

In order to verify the generality of our model and that its solution is a local optimal solution, the $f(k)$ we construct can set any goal to obtain a strategy that meets the maximum profit. For instance, to effectively measure the correlation between daily investment risk and return, the Sharpe ratio is employed as a risk-adjusted return indicator. This ratio plays a crucial role in assessing asset portfolio performance, evaluating capital market efficiency, constructing optimal asset portfolios, and facilitating investment decision-making. Mathematically, it is calculated as the standard deviation of the portfolio's expected return minus the risk-free rate of return, divided by the portfolio's rate of return [10].

Generally speaking, it is a good decision for the Sharpe ratio to be greater than three [10]. In order to meet the requirement of the Sharpe ratio, and to establish the objective function, we define it as the return after three days minus 0.8 times the risk assessment value σ_k . This formulation aims to stabilize the Sharpe ratio within the range of 3-4, which effectively quantifies the relationship between daily investment risk and return. To determine the overall objective function f_k , we compute the mean of all local f_k directly. Next, we introduce a perturbation to the obtained results and demonstrate that the perturbed outcomes no longer maximize the objective function. Specifically, we introduce a fluctuation range of 5% to the previously determined buying and selling strategies for gold and Bitcoin. Subsequently, we reconfigure the buying and selling schemes accordingly. The results from multiple tests are illustrated in Figure 7:

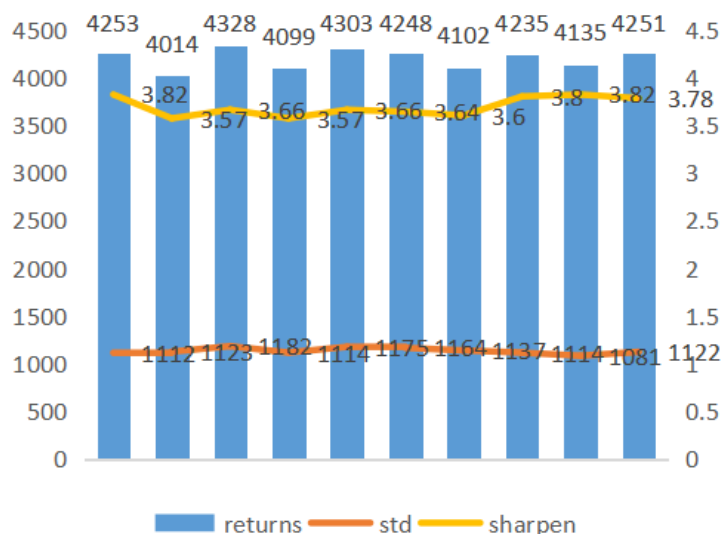


Figure 7. The mean, Sharpe value and total wealth value of the model after each disturbance

It can be seen that while considering Sharpe ratio and final return, the final results of other disturbances are not as good as the original scheme. Therefore, the local optimal value of the investment scheme we choose is effective.

4. Model deficiency

Based on LSTM, this model has the advantage of processing time series data and mitigating gradient disappearance or explosion problems. However, LSTM still has some limitations when dealing with very long sequences, and it requires a lot of computation. Specifically, our model only uses the historical prices of gold and bitcoin as training data and may lack real-world reference value. In practical applications, introducing more features, such as opening prices and trading volumes, may improve accuracy. In addition, because of smart algorithms and floating-point encoding, the output may contain tiny operations that are difficult to avoid. In practice, we adjust the transaction amount to achieve a more reasonable transaction.

5. Conclusion

Our findings provide investors with a coherent and reliable basis for making informed decisions to achieve long-term capital appreciation and effectively manage risk. The integration of VMD, LSTM, and PSO techniques demonstrates their effectiveness in optimizing daily trading strategies and highlights the potential for enhancing portfolio performance in the financial market.

Reference

- [1] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Retrieved from <https://bitcoin.org/bitcoin.pdf>
- [2] Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review*, 45(2), 217-229.
- [3] Bouoiyour, J., Selmi, R. and Wohar, M., 2018. Measuring the response of gold prices to uncertainty: An analysis beyond the mean. 3rd ed. pp.105-116.
- [4] BullionByPost. 2022. Bull Market. Available at: (<https://www.bullionbypost.co.uk/index/market-commentary/bull-market/>).
- [5] Cheah, E. T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130, 32-36.

- [6] Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar—A GARCH volatility analysis. *Finance Research Letters*, 16, 85-92.
- [7] Jiang, M., Jia, L., Chen, Z., et al. (2022). "The two-stage machine learning ensemble models for stock price prediction by combining mode decomposition, extreme learning machine and improved harmony search algorithm." *Annals of Operations Research*, 309, 553-585. <https://doi.org/10.1007/s10479-020-03690-w>.
- [8] Jian Cao, Zhi Li, Jian Li, 2019, 'Financial time series forecasting model based on CEEMDAN and LSTM', *Physica A: Statistical Mechanics and its Applications*, Volume 519, Pages 127-139, ISSN 0378-4371, available at: (<https://doi.org/10.1016/j.physa.2018.11.061>).
- [9] Federico M, Beata W (2015). "Particle swarm optimization (PSO): A tutorial." *Chemometrics and Intelligent Laboratory Systems*, 149(Part B), 153-165. ISSN 0169-7439.
- [10] Sharpe, W., 1994. The Sharpe Ratio. 'The Journal of Portfolio Management'. Available at: (<http://web.stanford.edu/~wfsarpe/art/sr/sr.htm>).