Problem Chosen	2022 MCM/ICM Summary Sheet	Team Control Number
C		2208487

Traders buy and sell volatile assets to maximize their income. Recently, with the growing interest of the public in cryptocurrencies, gold and bitcoin become a feasible combo. Our team is requested to determine the trading strategies for a trader which uses only the past daily prices.

As for problem 1, First we plot the price data and found that they show non-stationarity in the sense of mean. We therefore utilize ARIMA to predict the prices of gold and bitcoin. The predictor shows good capability to accurately predict the price of future. Price data from 2016 to 2019 is used to train the model. Then for position management, Markowitz mean-variance model is used to decide to sell, buy or hold each asset.

As for problem 2, we

As for problem 3, we

Finally

Key words:

To Harvest More: Best Trading Strategies on Gold and Bitcoins

Contents

1 Introduction

1.1 Problem Background

With the popularity of cryptocurrencies and the simplification of trading methods, more of the population become market traders. Some of them expect to outperform inflation, while others want to create wealth. By buying and selling volatile assets frequently, market traders pursue a goal to maximize their total return. Gold and bitcoins enjoy great popularity these days for their complementary characteristics in risk and value. Gold is stable in price and has lower risk while the value of bitcoins varies greatly and thus has a higher risk, as is shown in Figure 1.

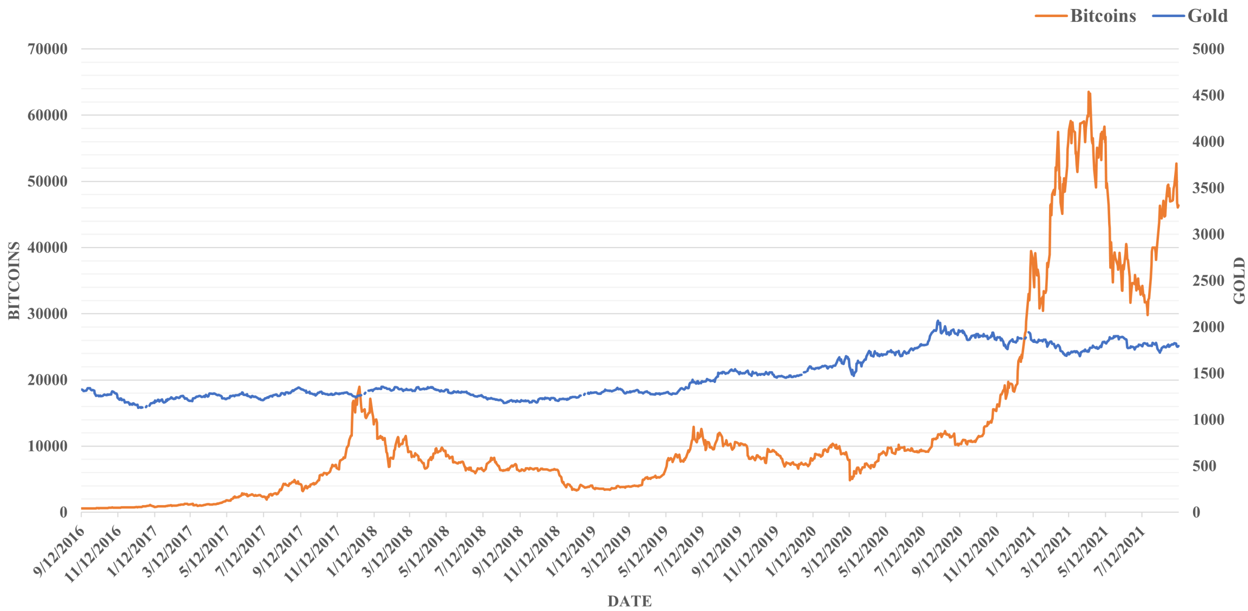

Figure 1: Gold and bitcoin daily prices, in U.S. dollars per troy ounce and U.S. dollars per bitcoin. **Source:** London Bullion Market Association, 9/11/2021 and NASDAQ, 9/11/2021

Regarding trading rules, Gold is only traded on days the market is open while bitcoins are traded every day. Commissions are charged to make each transaction. For market traders to achieve their goals, they need to build a model to determine the strategies to manage their portfolios well.

1.2 Restatement of the Problem

- Develop a model that gives the best daily trading strategy based only on price data up to that day, and calculate how much the initial $1000 investment is worth on 9/10/2021 using the model and strategy.

- Present evidence that your model provides the best strategy.

- Determine how sensitive the strategy is to transaction costs and analyze how the costs affect the strategy and results.

1.3 Our Approach

1. **To predict the prices of bitcoin and gold and make decision, we use *ARIMA* (Autoregressive integrated moving average).** To predict future prices with existing data is a difficult problem because too many factors may influence the prices. The international situation, national policies, and even social media can have a considerable impact on the prices. Moreover, the data shows non-stationarity. To take as many factors as possible and predict accurately according to their inner laws, we adopt this approach to predict the prices, which is proved to give satisfying results.
2. **To make decision on trading, we adopt *Markowitz mean-variance model* and *Sharp ratio*.**

2 Assumptions and Justifications

2.1 Assumptions

To simplify the problem stated above, we make following assumptions, each of which is justified properly:

1. **The trader will have $1000 in the beginning, and the transaction commissions for gold and bitcoin are $\alpha_{gold} = 1\%$ and $\alpha_{bitcoin} = 2\%$, respectively.**
2. **The market trader sells all of the gold and bitcoins by the end of the five-year trading period, i.e. on 9/10/2021.** Generally, investors cares about funding liquidity. Among cash, gold and bitcoins, only cash can circulate unhindered in the market. So we make this assumption and thus measure the outcome in cash.

2.2 Symbols and Definitions

2.3 Symbols and Definitions

3 Mathematical Modeling

3.1 Data Preprocessing

There are some missing value in the provided data of gold prices, as is shown in Figure 2.

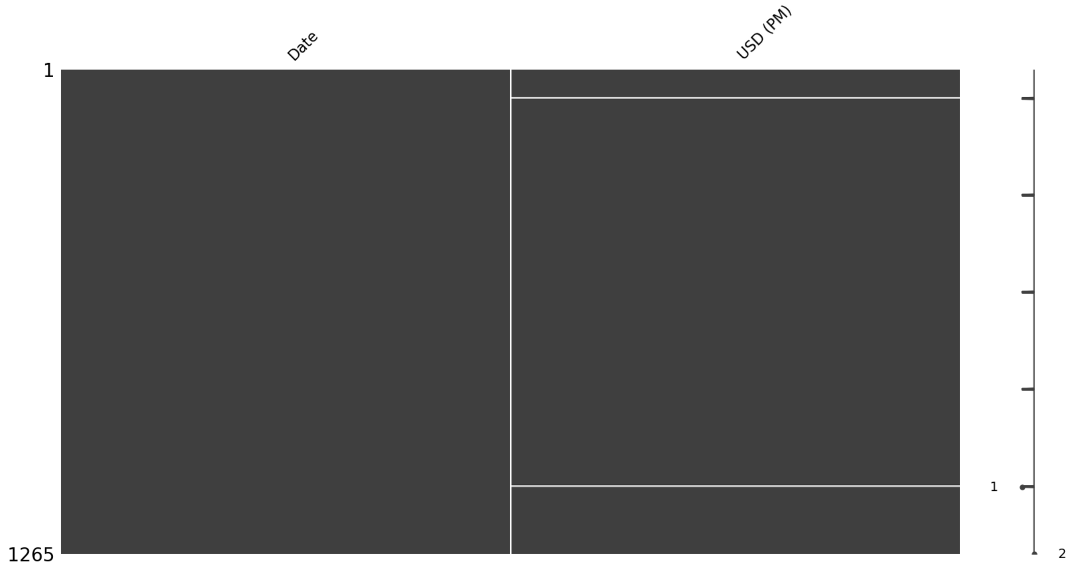

Figure 2: Data-dense view

The data-dense view reveals that in the price column named USD (PM), there are some missing values roughly in the end of each year. Carefully inspecting the data, we found there is two missing values each year, one is the last day and one is the fourth day from the last.

Figure 3: Filling the missing data around 12/22/2017 and 12/29/2017

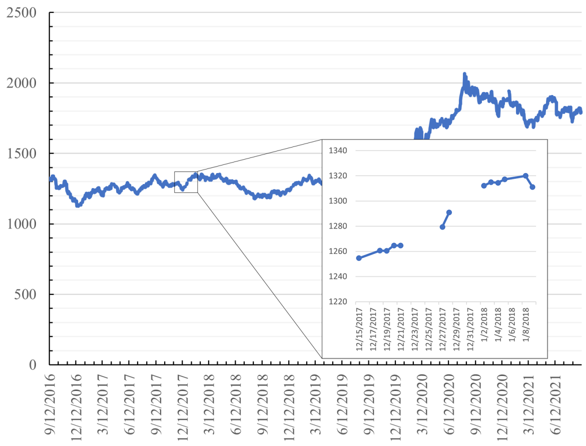

Figure 4: Before

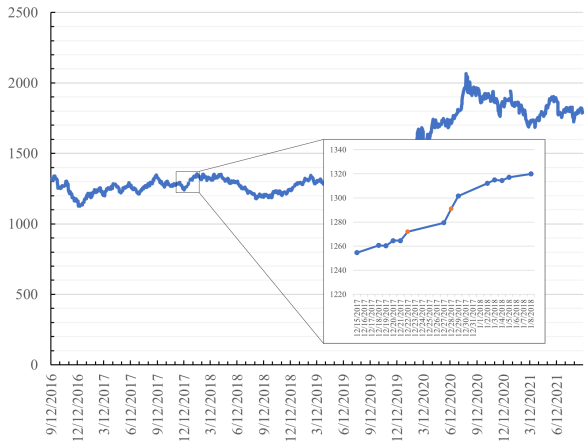

Figure 5: After

To make up the them, moving median method is used, where the length of window is 5. For each missing value v_0, we chose a window

$$v_{-2}, v_{-1}, v_0, v_1, v_2$$

Where v_{-2}, v_{-1} are the predecessors of v_0 and v_1,v_2 are the successors of v_0. Then, we let the value of v_0 be the *median* of $\{v_{-2}, v_{-1}, v_1, v_2\}$.

The method is commonly used when only a few values are missing, which is fast and fit in the trend of the provided data well, as is shown in the Figure 3.

3.2 Metric Calculating

3.3 ARIMA Predictor

3.4 Differentiating

4 Results and Solutions

Result analysis

Discussions

5 Model Evaluation

5.1 Annualized Return

5.2 Retracement

5.3 Sharp Rate

6 Sensitivity Analysis

7 Strength and Weakness

7.1 Strength

The models have the following strengths:

- Advantage 1
- Advantage 2

7.2 Weakness

The models have the following weaknesses:

- Weakness 1
- Weakness 2

8 Conclusions

In this paper, we build a ARIMA based predictor to enhance the income of trading cash, gold, and bitcoins. First, we preprocess the data and fill the missing values. Then, we finish the feature engineering by finite difference and feature store to boost the machine learning effect. After that, we use a RNN-based ARIMA to predict the price. Finally, we reach the optimality of our model by compare different trading strategies. Sensitive tests and evaluation are carried out. We measured our model by annualized return, retracement and sharp rate. The influence of commissions are experimented to test the sensitivity. Our model yields satisfying results.

References

[1] P. Y. Pawar and S. H. Gawande, "A Comparative Study on Different Types of Approaches to Text Categorization," *International Journal of Machine Learning and Computing*, vol. 2, no. 4, pp. 423-426, 2012.

[2] B. Pang and L. Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales," *Proceedings of the 43rd Meeting of the Association for Computational Linguistics*, pp. 115–124, 2005.

[3] J. McAuley, J. Leskovec, "Hidden factors and hidden topics: understanding rating dimensions with review text", *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys 2013)*, pp. 165-172, October 2013.

[4] N. Hu, P. Pavlou, and J. Zhang, "Overcoming the J-shaped distribution of products reviews," *Communications of the ACM*, vol. 52, pp. 144-147, 2009.

[5] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30-37, 2009.

[6] Y. F. Hu, Y. Koren, C. Volinsky, "Collaborative filtering for implicit feedback datasets," *Proceedings of IEEE International Conference on Data Mining (ICDM 2008)*, pp. 263-272, 2008.

[7] E. W. Frees, *Regression Modeling with Actuarial and Financial Applications*, Cambridge, UK: Cambridge University Press, 2009. ISBN: 978-0521135962

Memorandum to the Trader

Considering the intensely changing financial markets and the difficulty of handling your portfolio, using an appropriate model to make predictions and strategies to trade are of vital importance to improve income. It is a great honor for us to develop the model for you to buy, hold and sell your assets. Here is our model based on **ARIMA** and strategies for you to trade your assets effectively.

1.

2.

3.

We appreciate this opportunity to help you to build up a trading strategy for cash, gold and bitcoins, and we firmly believe that our model can be utilized in maximizing your total income. Feel free to contact us for further information on the proposal.

Sincerely yours

MCM 2022 Team

Appendix: Programs and Codes

If you do not want to provide program codes, delete this appendix section.

```python
import numpy as np
import matplotlib.pyplot as plt

x = np.array([-3.0, -2.9, -2.8, -2.7, -2.5, -2.4, -2.3, -2.2, -2.1, -2])
y = np.sin(x)

plt.figure()
plt.xlabel("x axis")
plt.ylabel("y axis")
plt.plot(x, y, '-^')
plt.xlim(min(x)-0.05, max(x)+0.05)
plt.ylim(min(y)-0.05, max(y)+0.05)
plt.legend(loc='best')
plt.show()
```