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

Use this template to begin typing the first page (summary page) of your electronic report. This
template uses a 12-point Times New Roman font. Submit your paper as an Adobe PDF electronic file (e.g. 1111111.pdf), typed in English, with a readable font of at least 12-point type.

Do not include the name of your school, advisor, or team members on this or any page.

Papers must be within the page limit specified in the problem statement.

Be sure to change the control number and problem choice above.
You may delete these instructions as you begin to type your report here.

Follow us @COMAPMath on Twitter or COMAPCHINAOFFICIAL on Weibo for the most up to date contest information.

To Harvest More: Achieving Best Trading Strategies on Gold and Bitcoins

Contents

1 Introduction

1.1 Problem Background

As the popularity of financial knowledge grows, many of the population, more or less, 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 enjoys 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. Regarding to trading rules, Gold is only traded on days the market is open while bitcoins are traded everyday. For both of them, commissions are charged to make each transaction.

Figure 1: Gold and bitcoin daily prices, 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

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 transaction costs affect the strategy and results.

1.3 Our Approach

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 market trader sells all of the gold and bitcoins by the end of the five-year trading period, i.e. 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

Table 1: Symbols and Definitions.

Notations	Description
η	1
ξ	
P	
r	
x	
X	
N	
n	

2.3 Symbols and Definitions

3 Mathematical Models

3.1 Basic Model

$$\sum_{t} \tag{1}$$

According to Equation (1)

Notations	Description
a	Persuasion of comments
$s(X \to Y)$	Degree of support between X and Y, indicating how often the rules can be used for analysis
$c(X \to Y)$	Confidence between X and Y, indicating the frequency of transactions in Y containing X
X	Promotion/The 'verified purchase' is 'N'
$\overline{X}$	No promotion or The 'verified purchase' is 'Y'
Y	Poor feedback
$\overline{Y}$	Favourable feedback
Z	Poor evaluation support rate
$\overline{Z}$	Favourable support rate
f_V	Amount of platform commentators
$f_{\overline{V}}$	Amount of common customers
a_V	Support rate of comments written by writers
$a_{\overline{V}}$	Support rate of comments written by non writers
a_T	Overall weighted support rate
$Q_\mu(v)$	Amount of comments, dependent variable in multiple linear regression
μ_i	Regression coefficient of multiple linear regression, $\{i = 0, 1, 2, 3\}$
v_i	Independent variable of multiple linear regression, $\{i = 0, 1, 2, 3\}$
v_1	Amount of no promotions in monthly reviews
v_2	Number of disapproval of poor feedback and approval of favorable feedback in each month
v_3	Frequency of good keywords in each month
ξ	Random error term of multiple linear regression
r^2	Sample determination coefficient discrimination coefficient
SSR	Regression sum of squares
SST	Sum of squares of total variation
T	Weighted mean value of star rating in the train set
$\widetilde{T}$	Weighted mean value of star rating in the testing set
std	Standard deviation of the result in training set and testing set
D	Future value of products
φ	Weighted star rating
δ	The rate of positive keywords in reviews

$$\begin{cases} \frac{dS_2}{dt} = -R_0 \cdot S_2(I_1 + I_2) \\ \frac{dI_2}{dt} = R_0 \cdot S_2(I_1 + I_2) - \frac{I_2}{r} \\ \frac{dS_1}{dt} = \rho \left[1 - \frac{S_1 + (1 + v/r) I_1}{K_1} \right] - R_0 \cdot S_1(I_1 + I_2) - v \cdot S_1 \\ \frac{dI_1}{dt} = R_0 \cdot S_1(I_1 + I_2) - \frac{I_1}{r + v} \end{cases} \tag{2}$$

3.2 Improved Model

Additional assumptions for the model improvement

Algorithm 1: Competitive selection

Input: the set of data patterns $\mathbb{X}$
Output: the set of prototype seeds $\mathbb{S}^*$
Compute the Euclidean distance $dist\left(\mathbf{x}_m, \mathbf{x}_n\right)$
Compute the density $D(\mathbf{x}_m) \geq \gamma$
Select eligible $\mathbf{x}_m$ for the candidate seed set $\mathbb{C}^0 \leftarrow \{\mathbf{x}_m \mid C(\mathbf{x}_m, \gamma) = 1\}$
Initialize $\mathbb{C}^* \leftarrow \mathbb{C}^0$
while $\mathbb{C}^* \neq \phi$ **do**
    Initialize $\mathbb{S}^j \leftarrow \mathbb{S}^*$
    Select the winning seed from the candidate set $\mathbf{x}_s^j \leftarrow \arg\max D(\mathbf{x}_m)$, $\mathbf{x}_m \in \mathbb{C}^j$
    Update $\mathbb{S}^* \leftarrow \mathbb{S}^j \cup \{\mathbf{x}_s^j\}$
    Update $j \leftarrow j + 1$
end
return $\mathbb{S}^*$

4 Results and Solutions

Result analysis

Discussions

5 Model Evaluation and Sensitivity Analysis

5.1 Model Evaluation

Figure 2: Figure illustration.

5.2 Sensitivity Analysis

6 Strength and Weakness

6.1 Strength

The models have the following strengths:

- Advantage 1
- Advantage 2

6.2 Weakness

The models have the following weaknesses:

- Weakness 1
- Weakness 2

7 Conclusions

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

Write the letter here.

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()
```