MAFS6010Z (L1) - Artificial Intelligence in Fintech Final Project—Cryptocurrency Trading

Li Zhihan, Huang Fan, Zhang guangyu, Shangguan Wenfang, Wang Siyu

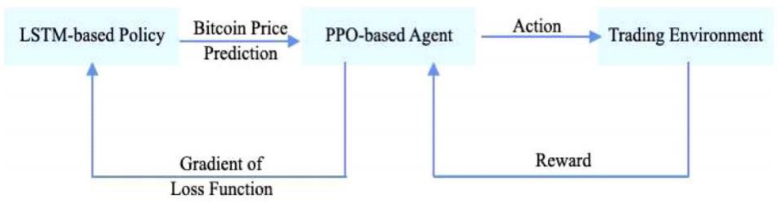
1. Introduction

In this project, we aim to use historical minute-level OHLCV (opening, high, low, close, volume) data of 4 major cryptocurrencies(BTC,ETH, LTC and XRP) to simulating trading process. Our main strategy is using a risk parity investment portfolio as the initial asset allocation, trading is conducted using a reinforcement learning agent program based on the cyclic PPO algorithm and LSTM strategy network.

2. Methodology

Reinforcement learning with LSTM (Long Short-Term Memory)

The learning network for reinforcement learning uses LSTM. Depending on the learning method of reinforcement learning, an action is determined by the LSTM network or randomly. We update the deep reinforcement learning network to prove the transaction outcome by changing the state according to the following circumstances. We also compare the difference between the behavior of a given transaction and the appropriate reward.



3. Feature Engineering

We have extracted many indicators through feature engineering and we selected weaker-correlated indicators.
From the heatmap of indicators'correlation. we divided them into two parts:

- Across varying-time period indicators:

Amihud,Amihud_30min,SLOWK_30min,VROCP6_5min,VROCP6_15min, amihud_30min.

- Non-across varying-time period indicators:

PLUS_DM, WILLR, RSI, APO, BOP, ADX, ADOSC, SAR.

4. Trading Strategy

- **Trading Strategy:**

We combines a Risk Parity Portfolio as the initial asset allocation with a Reinforcement Learning (RL) agent for dynamic trading decisions.

- **Environment and RL Integration:**

We have implemented this strategy using the OpenAI Gym package, which provides a standardized environment for training and testing.

- **Algorithm and Network Architecture:**

The Recurrent PPO algorithm allows our agent to learn from past experiences, enabling it to make more informed decisions. For the policy network, we use an LSTM architecture.

- **Action:**

The RL agent's actions are defined as a vector encompassing both trade type (buy, sell, hold) and trade size. The trade size is represented as a percentage (0% to 100%) of the current cash balance.

- **Reward Metric:**

Use the Sortino Rate as the reward metric.

- **Hyperparameter Optimization:**

For hyperparameter tuning, we utilize the 'optuna' package to perform an efficient search for the optimal set of hyperparameters that maximize the mean reward achieved by our RL agent.

5. Backtest

Setting Optimal Weights:[0.50477571 0.3363203 0.1448375 0.01406649] for BTC,ETH, LTC and XRP.

Analysis:

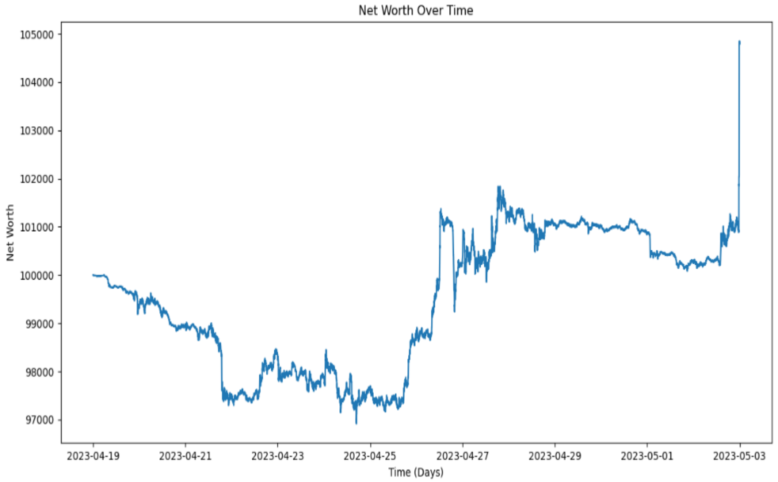
The results showcase a robust performance of the cryptocurrency investment strategy. The strategy performance show in the table:

Metric	Value
Annualized Total Return	86.86%
Annualized Sigma	16.37%
Sharpe Ratio	5.2636
Max Drawdown	-3.08%
Turnover Rate	269.32%

MAFS6010Z (L1) - Artificial Intelligence in Fintech Final Project—Cryptocurrency Trading

Li Zhihan, Huang Fan, Zhang guangyu, Shangguan Wenfang, Wang Siyu

The Sharpe Ratio of 5.24 indicates an attractive risk-adjusted return profile, indicating strong performance relative to the risk taken. 269% Turnover Rate implies a frequent repositioning of assets within the portfolio. Since we are doing high frequency Cryptocurrency trading, the expected turnover rate might be high.



From the net worth over time graph, it performs poor for a while at the beginning. However, after half time, it begins to perform good and finally achieve ideal results. This process is quite like how reinforcement learning work, as reinforcement learning has distinctive characteristics of trial and error, exploration and delayed rewards. Our outcome shows that the algorithm has achieve a satisfactory balance between exploration and exploitation of RL.

6. Conclusion

➢ **Advantages**

- From our simulating trading process, the **Recurrent PPO algorithm** has shown excellent performance in Cryptocurrency market.
- The strategy driven by it has generated impressive return and has done a good job on risk-return tradeoff.
- The algorithm also utilizes the advantages of **reinforcement learning** quite well, as it shows clear exploration and exploitation process and achieve good balance.

➢ **Disadvantages**

- The risk of **overfitting** is a concern as the model may become too closely tailored to the training data and perform poorly on unseen data.
- **The lack of dynamic weight adjustments** is another drawback. It is crucial for the model to adapt to changing conditions.
- Additionally, the training **duration** poses a challenge and there are **high noise levels** present in Bitcoin price data.

In conclusion, the price fluctuations can be influenced by various factors, including market sentiment and external events. These unpredictable variations can impede the convergence of the model, requiring a longer training period to accurately capture the underlying patterns.

7. Improvement

To improve our strategy and make it possible to put it into practice, we ca attempt to do the following:

➢ Do some research on control the turnover rate to reduce transaction cost.

➢ The weights can be changed from a fixed percentage to a dynamic adjustment based on the volatility and returns of the portfolio over time.

➢ Try different reward to see if we can further improve the balance between exploration and exploitation.

8. Reference

- Lee J, Koh H, Choe H J. Learning to trade in financial time series using high-frequency through wavelet transformation and deep reinforcement learning[J]. Applied Intelligence, 2021, 51: 6202-6223.
- Liu F, Li Y, Li B, et al. Bitcoin transaction strategy construction based on deep reinforcement learning[J]. Applied Soft Computing, 2021, 113: 107952.

9. Contribution

- Li Zhihan, Shangguan Wenfang: Coding;
- Wang Siyu, Huang Fan, Zhang Guangyu: Poster, powerpoint