Cryptocurrency Asset Management with Proximal Policy Optimization

Jacob T. Cassady*
Johns Hopkins Whiting School of Engineering, Baltimore, Maryland, 21218

Abstract here.

I. Introduction

This paper begins in Section II with a description of the models designed for cryptocurrency trading including the feature space, reinforcement learning algorithm, action spaces, and reward functions. Section III details the python implementation of the models with a focus on the leveraged 3rd party libraries. Section IV presents the performance of the agents on historical data with comparisons to some baseline investment strategies. The paper finishes with a discussion of the results in Section V and an introduction to possible future work in Section VI.

II. Model Design

Two models were tested for cryptocurrency trading: a Buy/Sell/Hold (BSH) model and a Managed Risk model. Both models use the same feature space and reinforcement learning algorithms, but the action spaces and reward functions are different. Section II.A describes the feature space used for the models. Section II.B describes the reinforcement learning algorithm. Section II.C introduces the action space and reward function of the BSH model. Section II.D details the action space and reward function of the Managed Risk model.

A. Feature Space

The feature space for the BSH and Managed Risk models is a combination of technical analysis indicators for a target asset, technical analysis indicators for related assets, and price data for the target asset. At each time t, the technical analysis indicators described in Sections II.A.1 through II.A.8 are calculated for the target asset.

1. Simple Moving Average

The simple moving average (SMA) is a arithmetic average of the last n prices of an asset. The equation for simple moving average is shown in equation 1 where k is the number of periods used to calculate the average.

$$SMA_t = \frac{1}{k}\sum_{i=0}^{k-1} Close_{t-i} \tag{1}$$

*Graduate Student, Robotics and Autonomous Systems

2. Relative Strength Index

The relative strength index (RSI) is a momentum oscillator that is meant to indicate the strength of a financial asset market [1]. RSI is calculated using equations 2 and 2 where k is the number of periods used to calculate the average gain and average loss.

$$RSI_t = \begin{cases} 100 - \frac{100}{1+\frac{\text{Average gain}_t}{\text{Average loss}_t}}, & \text{if } t < k, \\ 100 - \frac{100}{1+\frac{(\text{Average gain}_{t-1} \cdot k)+\text{Average gain}_t}{(\text{Average loss}_{t-1} \cdot k)+\text{Average loss}_t}}, & \text{if } t \geq k. \end{cases} \tag{2}$$

3. Commodity Channel Index

The commodity channel index (CCI) compares the current price of an asset to an average price over a period of time [2]. CCI is calculated using equation 5 where k is the number of periods used to calculate the typical price and mean deviation.

$$\text{Typical Price}_t = \frac{\text{High}_t + \text{Low}_t + \text{Close}_t}{3} \tag{3}$$

$$\text{Mean Deviation}_t = \frac{1}{k} \sum_{i=0}^{k-1} \left| \text{Typical Price}_{t-i} - \text{SMA}_t \right| \tag{4}$$

$$CCI_t = \frac{\text{Typical Price}_t - \text{SMA}_t}{0.015 \cdot \text{Mean Deviation}_t} \tag{5}$$

4. Average Directional Index

[1]

$$ADX = \frac{1}{n} \sum_{i=1}^{n} \frac{\left| \text{High}_i - \text{Low}_i \right|}{\text{High}_i} \tag{6}$$

5. Moving Average Convergence/Divergence

[3]

$$EMA_{t,k} = \left(\text{Close}_t \times \frac{2}{k+1} \right) + \text{EMA}_{t-1} \left(1 - \frac{2}{k+1} \right) \tag{7}$$

$$MACD_t = EMA_{t,12} - EMA_{t,26} \tag{8}$$

6. Bollinger Bands

[4]

$$\text{Upper Band}_t = SMA_t + 2\sigma \tag{9}$$

$$\text{Lower Band}_t = SMA_t - 2\sigma \tag{10}$$

7. Average True Range

[1]

$$\text{TR}_t = \max\left(\text{High}_t - \text{Low}_t, \left|\text{High}_t - \text{Close}_{t-1}\right|, \left|\text{Low}_t - \text{Close}_{t-1}\right|\right) \tag{11}$$

$$\text{ATR}_t = \frac{1}{k}\sum_{i=0}^{k-1} \text{TR}_{t-i} \tag{12}$$

8. Rate of Change

$$\text{ROC}_{t,k} = \left(\frac{\text{Close}_t - \text{Close}_{t-k}}{\text{Close}_{t-k}}\right) \times 100 \tag{13}$$

9. On-Balance Volume

[5]

$$\text{OBV}_t = \text{OBV}_{t-1} + \text{Volume}_t \cdot \text{Sign}\left(p_t - p_{t-1}\right) \tag{14}$$

10. Stochastic Oscillator

$$\text{Stochastic Oscillator} = \left(\frac{\text{Close}_t - \text{Low}_k}{\text{High}_k - \text{Low}_k}\right) \tag{15}$$

B. Model Selection

Model reference. [6]

A2C [7].

TRPO [8].

PPO [9].

SAC [10].

TD3 [11].

DDPG [12].

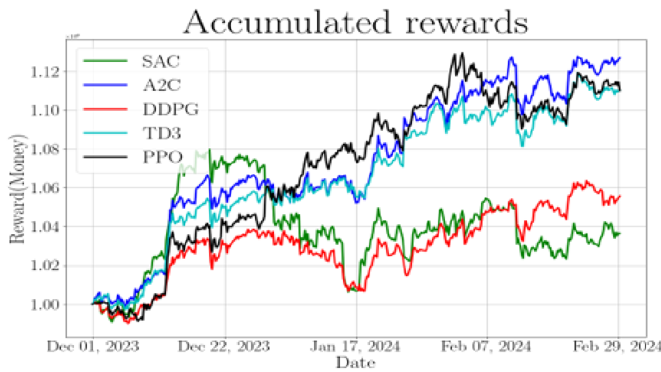

Fig. 1 Performance of Reinforcement Learning Strategies on Financial Asset Trading [6]

C. Buy/Sell/Hold Agent

1. BSH Action Space

Action	Description
0	Invested in cash.
1	Invested in asset.

Table 1 Buy/Sell/Hold Agent Action Space

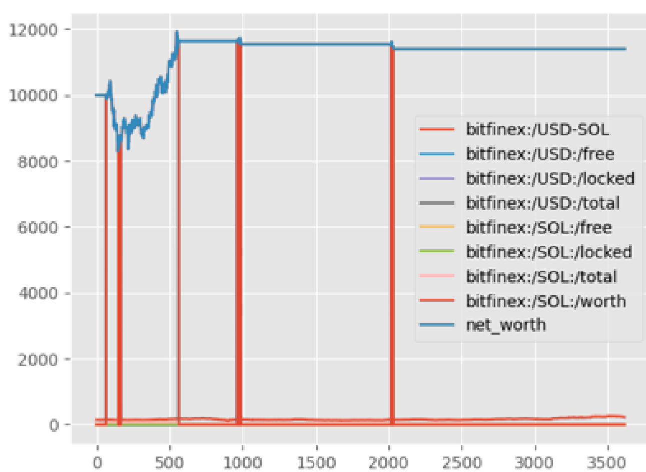

Fig. 2 BSH Action Plot Example

2. Position-Based Reward Function

$$R_t = (\text{Close}_t - \text{Close}_{t-1}) \cdot x_t \tag{16}$$

D. Managed Risk Agent

1. Limit Order Action Space

Action Parameter	Description	Values used
Stop Loss	Maximum loss before selling	[2%, 5%, 10%]
Take Profit	Maximum profit before selling.	[1%, 5%, 10%, 15%]
Trade Size	Size of next trade.	$[\frac{1}{16}, \frac{2}{16}, \ldots, 1]$

Table 2 Limit Order Agent Action Space

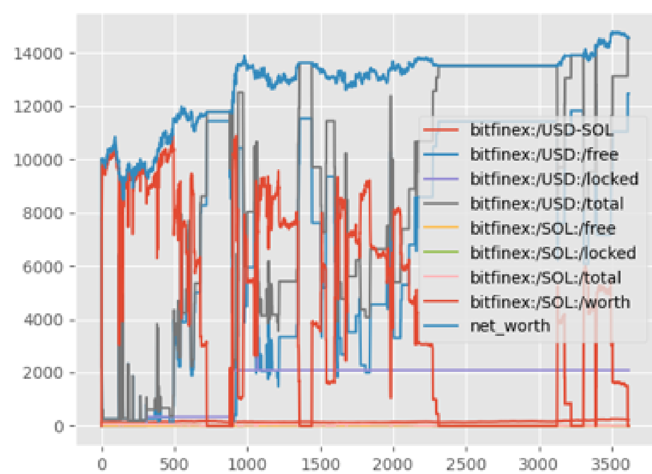

Fig. 3 Limit Order Action Plot Example

2. Sortino Ratio Reward Function

[13]

$$R_t = \frac{\frac{x_t - x_{t-1}}{x_{t-1}} - r_f}{\sigma_d} \tag{17}$$

III. Implementation

The models described above were implemented in Python by leveraging third party libraries shown in table 4. The hyperparameters for the technical analysis indicators are shown in table 3.

Technical Analysis Indicator	Hyperparameters	Target Asset	Related Asset(s)
Relative Strength Index	$k = 14$	✓	✓
Consumer Commodity Index	$k = 30$	✓	
Average Directional index	$k = 30$	✓	
Simple Moving Average	$k = 30$	✓	
Simple Moving Average	$k = 60$	✓	
Moving Average Convergence/Divergence	$fast = 12, slow = 26$	✓	✓
Bollinger Bands	$k = 20$	✓	
Average True Range	$k = 20$	✓	
Rate of Change	$k = 20$	✓	
On-Balance Volume	$k = 20$	✓	
Stochastic Oscillator	$k = 20$	✓	

Table 3 Technical Analysis Indicator Hyperparameters

3rd Party Libraries	Usage
NumPy [14]	
pandas [15]	
pandas_ta [16]	
TensorTrade [17]	
stable_baselines3 [18]	

Table 4 3rd Party Libraries

IV. Analysis

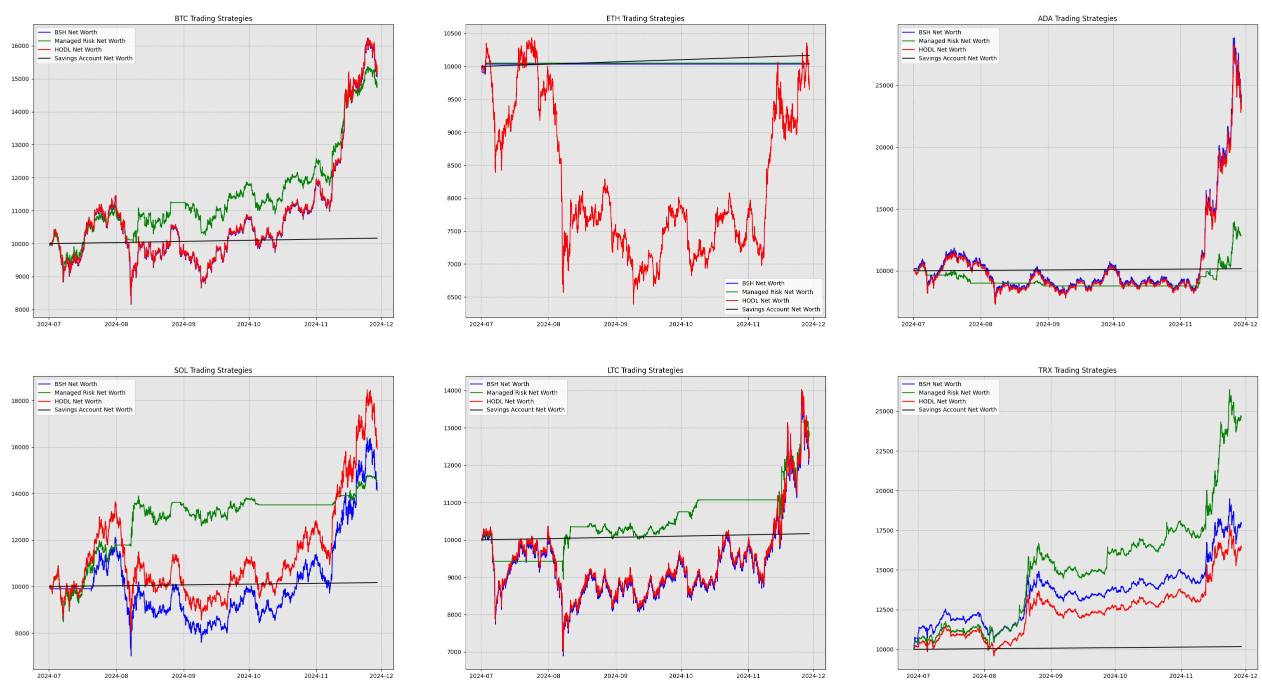

Fig. 4 Results by Cryptocurrency

Fig. 5 Aggregate Results

V. Discussion

VI. Future Work

Appendix

Acknowledgements

References

[1] Wilder, J., *New Concepts in Technical Trading Systems*, Trend Research, 1978. URL `https://books.google.com/books?id=WesJAQAAMAAJ`.

[2] Lambert, D. R., "Commodity channel index: Tool for trading cyclic trends," *Technical Analysis of Stocks & Commodities*, Vol. 1, 1983, p. 47.

[3] Appel, G., *Technical analysis: power tools for active investors*, FT Press, 2005.

[4] Bollinger, J., *Bollinger on Bollinger bands*, McGraw-Hill New York, 2002.

[5] Granville, J., *New Key to Stock Market Profits*, Martino Publishing, 1963.

[6] Mohammadshafie, A., Mirzaeinia, A., Jumakhan, H., and Mirzaeinia, A., "Deep Reinforcement Learning Strategies in Finance: Insights into Asset Holding, Trading Behavior, and Purchase Diversity," , June 2024. https://doi.org/10.48550/arXiv.2407.09557.

[7] Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T. P., Harley, T., Silver, D., and Kavukcuoglu, K., "Asynchronous Methods for Deep Reinforcement Learning," , 2016. URL https://arxiv.org/abs/1602.01783.

[8] Schulman, J., Levine, S., Moritz, P., Jordan, M. I., and Abbeel, P., "Trust Region Policy Optimization," , 2017. URL https://arxiv.org/abs/1502.05477.

[9] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O., "Proximal Policy Optimization Algorithms," , 2017. URL https://arxiv.org/abs/1707.06347.

[10] Haarnoja, T., Zhou, A., Hartikainen, K., Tucker, G., Ha, S., Tan, J., Kumar, V., Zhu, H., Gupta, A., Abbeel, P., and Levine, S., "Soft Actor-Critic Algorithms and Applications," , 2019. URL https://arxiv.org/abs/1812.05905.

[11] Fujimoto, S., van Hoof, H., and Meger, D., "Addressing Function Approximation Error in Actor-Critic Methods," , 2018. URL https://arxiv.org/abs/1802.09477.

[12] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D., "Continuous control with deep reinforcement learning," , 2019. URL https://arxiv.org/abs/1509.02971.

[13] Sortino, F. A., and Price, L. N., "Performance measurement in a downside risk framework," *the Journal of Investing*, Vol. 3, No. 3, 1994, pp. 59–64.

[14] Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E., "Array programming with NumPy," *Nature*, Vol. 585, No. 7825, 2020, pp. 357–362. https://doi.org/10.1038/s41586-020-2649-2, URL https://doi.org/10.1038/s41586-020-2649-2.

[15] pandas development team, T., "pandas-dev/pandas: Pandas," , Feb. 2020. https://doi.org/10.5281/zenodo.3509134, URL https://doi.org/10.5281/zenodo.3509134.

[16] "pandas-ta," , 2024. URL https://github.com/twopirllc/pandas-ta.

[17] "TensorTrade," , 2024. URL https://www.tensortrade.org/en/latest/index.html.

[18] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., and Dormann, N., "Stable-Baselines3: Reliable Reinforcement Learning Implementations," *Journal of Machine Learning Research*, Vol. 22, No. 268, 2021, pp. 1–8. URL http://jmlr.org/papers/v22/20-1364.html.