# A Supplementary Materials

## A.1 Baselines

- **Buy & Hold** evenly distributes the capital to all assets at the beginning and sells them all at the end of the backtest.
- **UCRP** is similar to "Buy & Hold" but maintains the allocation weights of assets throughout the period. Both strategies follow the market trend.
- **Mean-Variance** takes into account the risk and return of each portfolio and uses a convex objective function to find a portfolio on the Markowitz portfolio efficient frontier.
- **OLMAR** is a mean reversion trading algorithm that considers multiple periods when determining the mean and explicitly predicts the next price relatives ($p_{t+1}/p_t$). It also uses online learning to manage portfolios.
- **WMAMR** is an extension of OLMAR and takes into account an $\epsilon$-insensitive loss. It keeps the portfolio and waits for prices to bounce back if the return is less than $\epsilon$.
- **DPM** is a model-free DRL method based on the deterministic policy gradient. The agents are trained to trade assets and aim to maximize returns.
- **SARL** considers augmented asset states to boost cumulative returns in PM. The states can be news or predicted price movements that are beneficial for decision-making.

## A.2 Metrices

We evaluated the performance of the investment strategies using standard evaluation measures: portfolio value, Sharpe ratio, and maximum drawdown. The following provides a brief description of these measures.

- **Portfolio value (PV)** is a measure of the return of an investment period [0, T], which is calculated by:

$$\mathrm{PV} = P_T/P_0 = \Pi_{t=1}^{T} \mu_t \mathbf{a}_t \cdot \mathbf{y}_t, \tag{17}$$

  where $P_T$ is the total reward defined in Equation 3. A higher PV is considered better, however, this measure does not take into account risks during the investment period. A loss in a certain sub-period could be severe even if the overall return of the whole investment period is positive.

- **Sharpe Ratio (SR)** [3] measures the mean return relative to the total risk and is calculated as:

$$\mathrm{SR} = \frac{\mathbb{E}(r_t)}{\sqrt{\mathrm{Var}(r_t)}}, \tag{18}$$

  where $r_t = (P_t/P_{t-1}) - 1$ is the return rate at time $t$. A high SR value is traditionally preferred as it indicates a high mean return rate and low variance of the rates. However, a sudden high return can decrease the Sharpe Ratio as it increases the standard deviation in the denominator.

- **Maximum drawdown (MDD)** measures the risk by calculating the maximum loss from a peak to a trough. The MDD is calculated as:

$$\mathrm{MDD}_T = \max_{\tau \in [1,T]} \left( \max_{t \in [1,\tau]} \left( \frac{P_t - P_\tau}{P_t} \right) \right). \tag{19}$$

  A smaller MDD value is considered better. MDD is determined based on a single loss throughout the period and does not take into account returns. Allocating all capital in cash would achieve zero MDD, but it would also earn low returns.

## A.3 Comparison to other Traditional Trading Methods

In addition to comparing our method to DRL methods, we also compared it to several traditional trading algorithms on the same dataset. These methods can be classified into three categories: (1) follow-the-loser: Anticor[4], CWMR[5] and PAMR [6], (2) follow-the-winner: EG[7], ONS[8] and UP[2], and (3) pattern matching: BK[1]. The results are shown in Table 1. It is worth noting that the numbers in this table are comparable to the numbers in Table 1 in the main paper, as the experimental setting is identical.

# References


[1] L. Györfi, G. Lugosi, and F. Udina, "Nonparametric kernel-based sequential investment strategies," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 16, no. 2, pp. 337–357, 2006.

[2] T. M. Cover, "Universal portfolios," in *The Kelly Capital Growth Investment Criterion: Theory and Practice*, 2011, pp. 181–209.

[3] W. F. Sharpe, "The sharpe ratio," *Journal of portfolio management*, vol. 21, no. 1, pp. 49–58, 1994.

[4] A. Borodin, R. El-Yaniv, and V. Gogan, "Can we learn to beat the best stock," *Advances in Neural Information Processing Systems*, vol. 16, 2003.

[5] B. Li, S. C. Hoi, P. Zhao, and V. Gopalkrishnan, "Confidence weighted mean reversion strategy for online portfolio selection," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 7, no. 1, pp. 1–38, 2013.

[6] B. Li, P. Zhao, S. C. Hoi, and V. Gopalkrishnan, "Pamr: Passive aggressive mean reversion strategy for portfolio selection," *Machine learning*, vol. 87, no. 2, pp. 221–258, 2012.

[7] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth, "On-line portfolio selection using multiplicative updates," *Mathematical Finance*, vol. 8, no. 4, pp. 325–347, 1998.

[8] A. Agarwal, E. Hazan, S. Kale, and R. E. Schapire, "Algorithms for portfolio management based on the newton method," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 9–16.

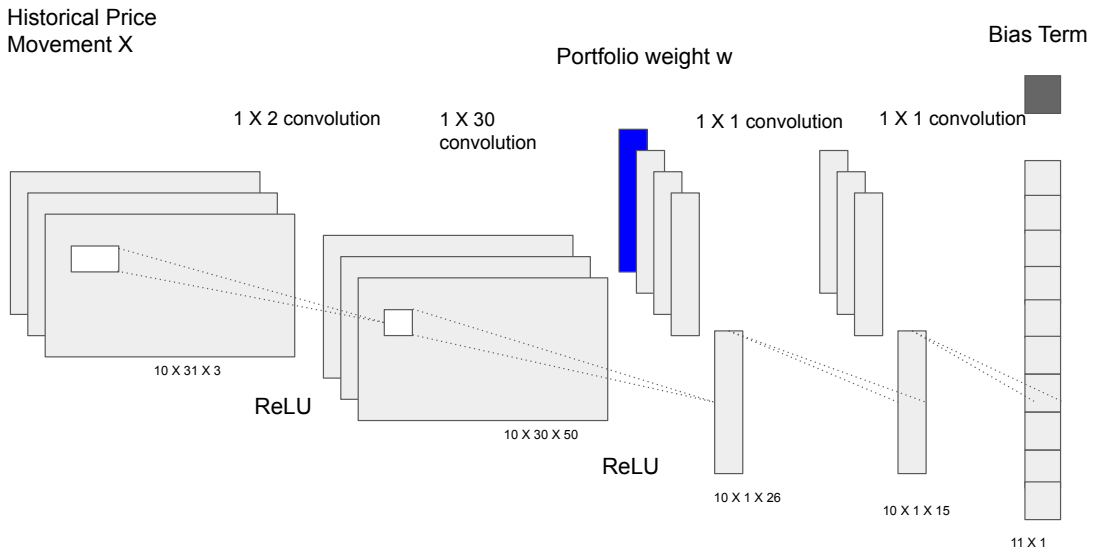


Figure 1: The illustration shows the network design, which includes four convolution layers. The kernel dimensions are $1\times2$, $1\times30$, $9\times1$, $1\times1$, $9\times1$ and channel sizes are 10, 5, 9, 15, 10. The resolution of each output layer is $9\times30\times10$, $9\times1\times5$, $9\times1\times9$, $9\times1\times15$, $10\times1$. The convolution kernels either consider one or all assets at a time and the order of assets does not have any effect.

U.S. Stock Market

| Method | Increase | | | No change | | | Decrease | | |
|---|---|---|---|---|---|---|---|---|---|
| | PV ↑ | SR ↑ | MDD ↓ | PV ↑ | SR ↑ | MDD ↓ | PV ↑ | SR ↑ | MDD ↓ |
| ANTICOR | 1.51 | 2.32 | 0.07 | 0.96 | 0.01 | 0.28 | 0.90 | 0.11 | 0.60 |
| BK | 1.25 | 0.72 | 0.29 | 0.88 | -0.09 | 0.36 | 0.58 | -0.99 | 0.63 |
| CWMR | 0.42 | -2.21 | 0.59 | 0.30 | -2.57 | 0.72 | 0.49 | -0.47 | 0.70 |
| EG | 1.38 | 2.28 | 0.05 | 1.17 | 0.71 | 0.16 | 0.97 | 0.18 | 0.54 |
| ONS | 1.34 | 2.22 | 0.06 | 1.10 | 0.47 | 0.14 | 1.20 | 0.56 | 0.55 |
| PAMR | 0.45 | -2.05 | 0.56 | 0.33 | -2.37 | 0.69 | 0.48 | -0.51 | 0.70 |
| UP | 1.38 | 2.28 | 0.05 | 1.16 | 0.71 | 0.16 | 0.97 | 0.17 | 0.54 |

Cryptocurrency Market

| Method | Increase | | | No change | | | Decrease | | |
|---|---|---|---|---|---|---|---|---|---|
| | PV ↑ | SR ↑ | MDD ↓ | PV ↑ | SR ↑ | MDD ↓ | PV ↑ | SR ↑ | MDD ↓ |
| ANTICOR | 2.16 | 3.73 | 0.37 | 1.06 | 0.93 | 0.36 | 0.62 | -2.87 | 0.52 |
| BK | 1.30 | 1.85 | 0.59 | 0.27 | -4.81 | 0.79 | 0.45 | -3.87 | 0.75 |
| CWMR | 0.01 | -15.31 | 0.99 | 0.01 | -13.75 | 0.99 | 0.01 | -16.71 | 0.98 |
| EG | 2.16 | 4.71 | 0.36 | 1.10 | 1.22 | 0.19 | 0.76 | -3.49 | 0.34 |
| ONS | 1.63 | 3.03 | 0.43 | 1.27 | 2.44 | 0.21 | 0.64 | -3.05 | 0.43 |
| PAMR | 0.01 | -14.09 | 0.99 | 0.02 | -12.13 | 0.98 | 0.03 | -14.57 | 0.97 |
| UP | 2.16 | 4.71 | 0.36 | 1.10 | 1.18 | 0.19 | 0.76 | -3.49 | 0.33 |

Table 1: The table describes statistics of portfolio values, Sharpe ratios, and maximum drawdowns for traditional methods on both the U.S. stock market and the cryptocurrency market. Upward arrows indicate higher values are better and downward arrows indicate lower values are better. The Sharpe ratios presented are annualized.