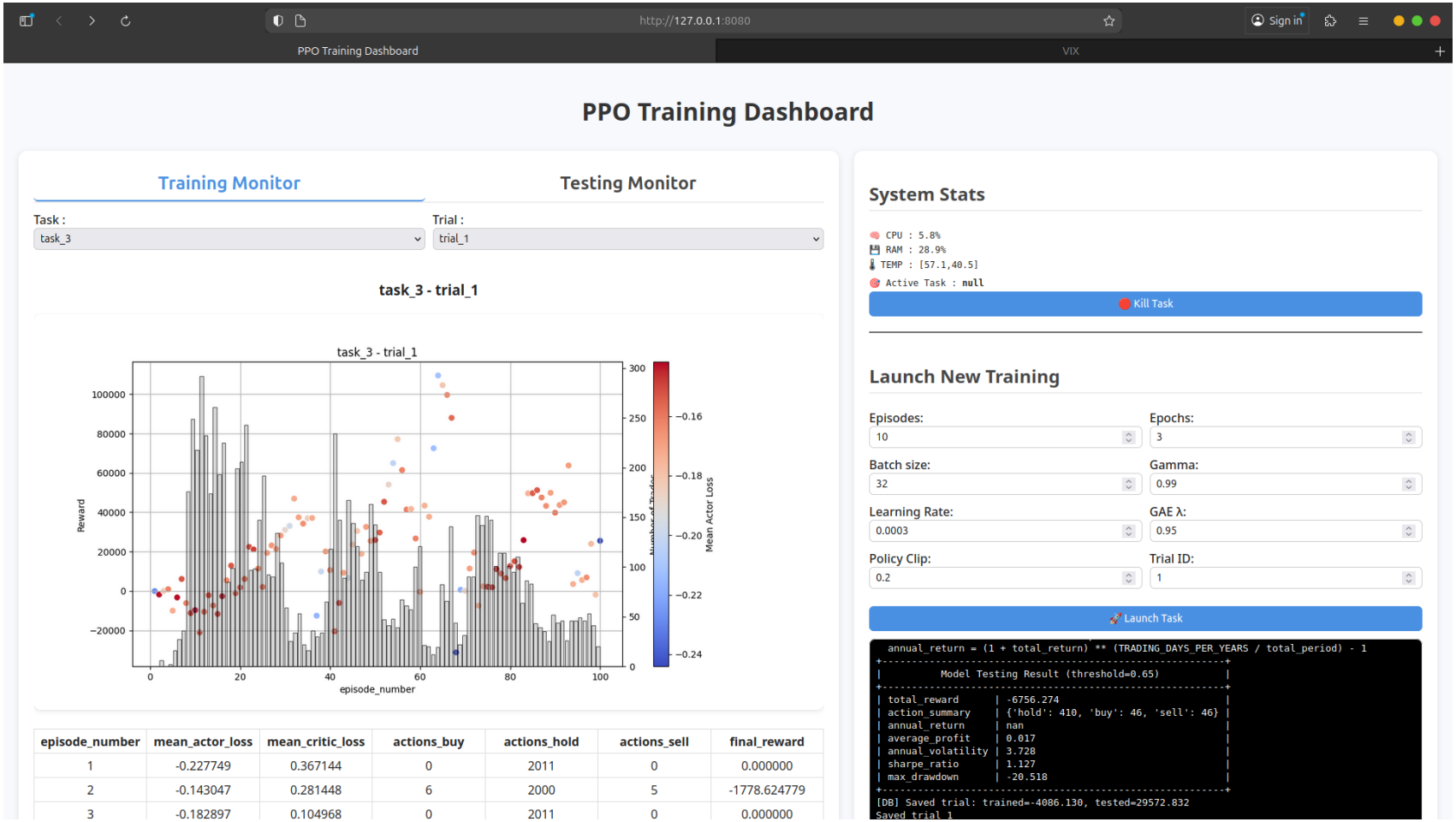
PPO Market Bot

Mathys Vinatier

Seoul National University

November 5, 2025

PPO Training API



episode_number	mean_actor_loss	mean_critic_loss	actions_buy	actions_hold	actions_sell	final_reward
1	-0.227749	0.367144	0	2011	0	0.000000
2	-0.143047	0.281448	6	2000	5	-1778.624779
3	-0.182897	0.104968	0	2011	0	0.000000

Figure 1: Dashboard - Training Monitor

PPO Training API

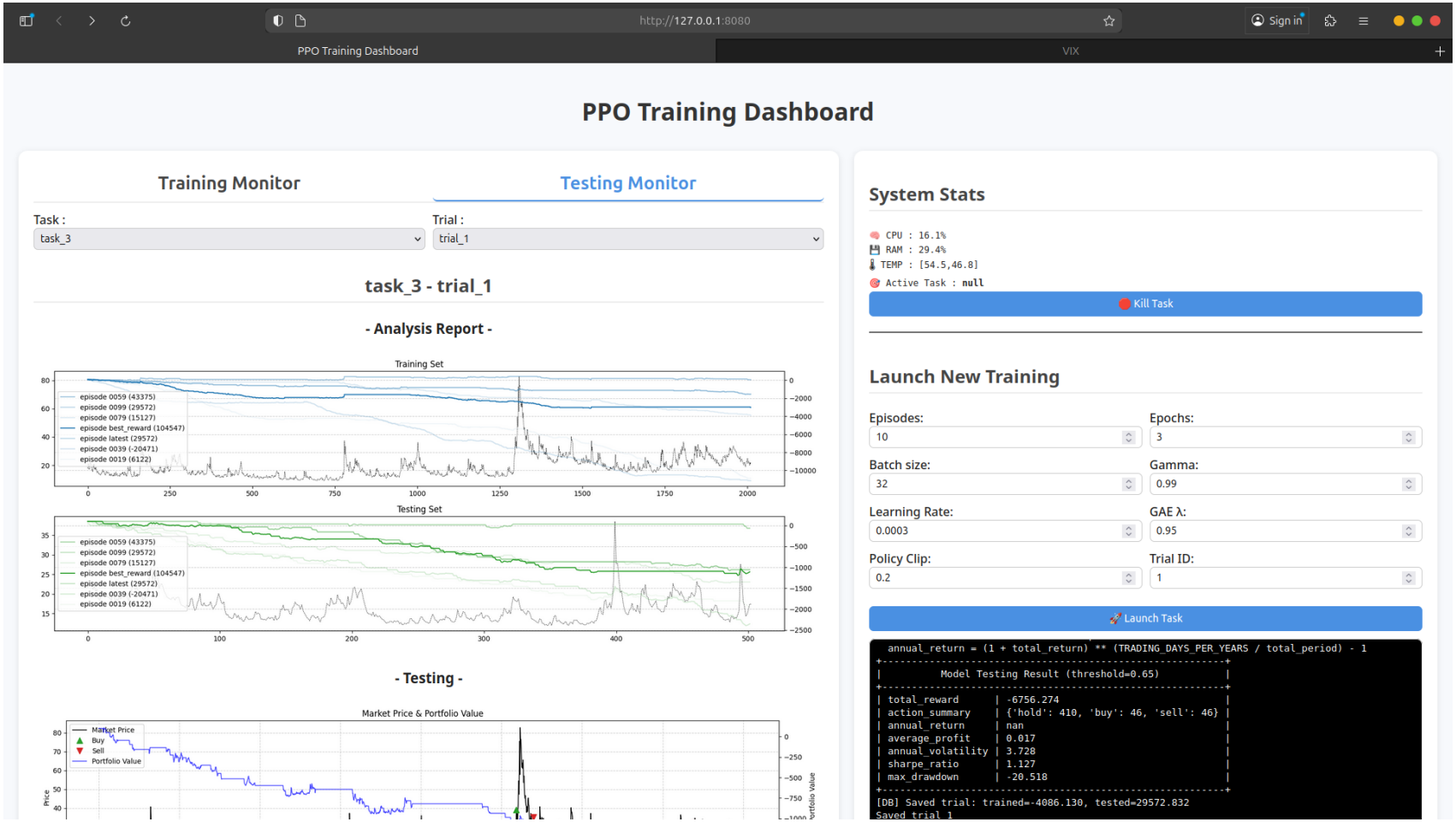
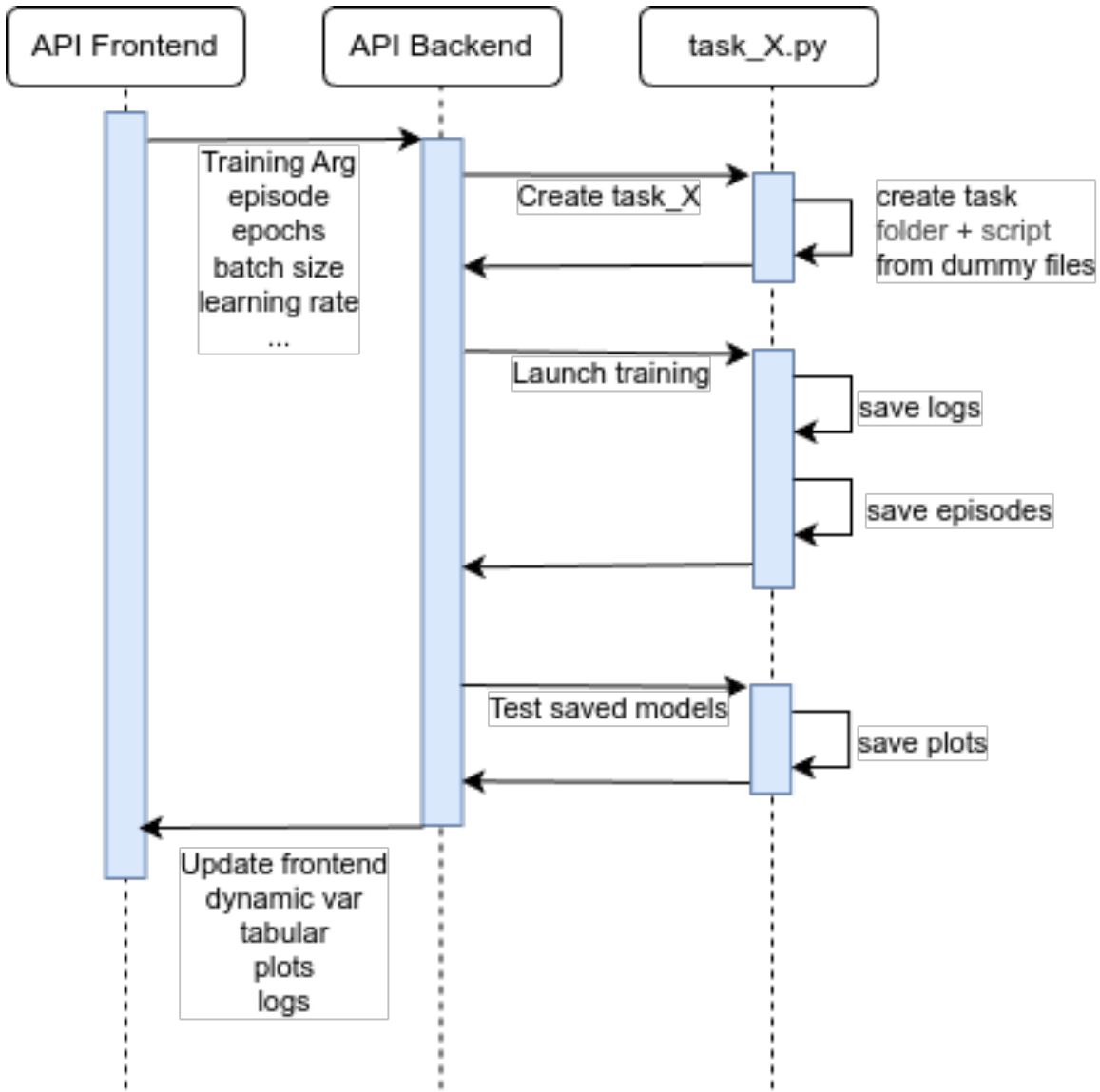


Figure 2: Dashboard - Testing Monitor

PPO Training API



Training (no broker) - batch_size = 64, epochs = 112, episode = 10

Figure 3: Task 3 - Training plot (no broker)

Figure 4: Task 3 - Testing plot (no broker)

Training (no broker) - batch_size = 32, epochs = 100, episode = 100

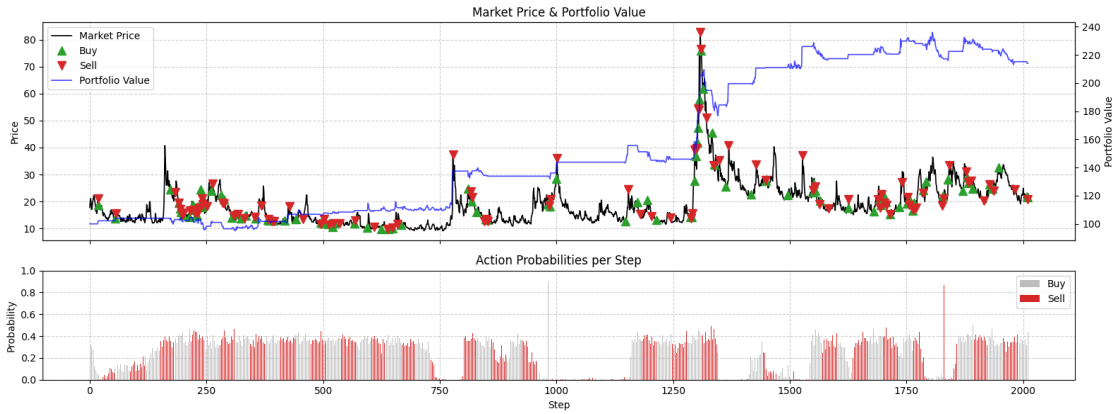


Figure 5: Task 4 - Training plot (no broker)

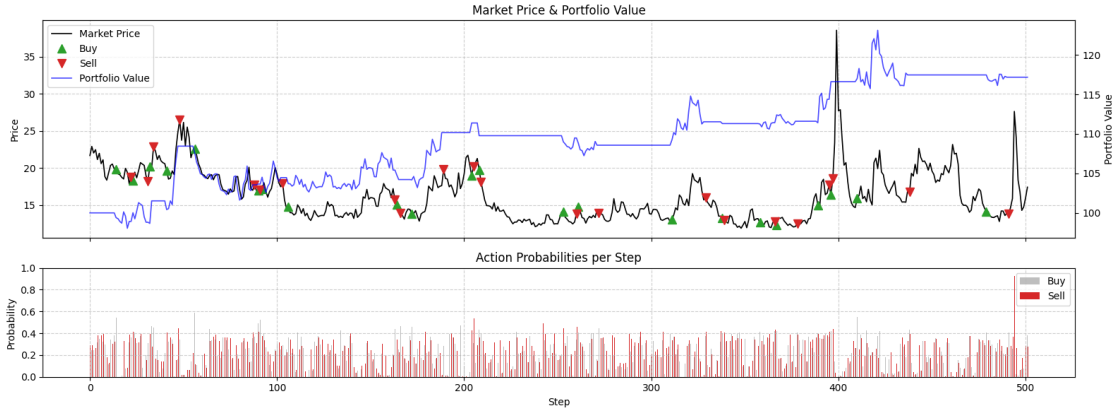


Figure 6: Task 4 - Testing plot (no broker)

Training (no broker) - Conslusion

annual return	avg profit	annual vol.	max drawdown
0.082	0.0003	0.11	6%

Episodes	Epochs	Tick	Batch Size	Learning Rate	Final Profit on test
100	100	Daily	32	0.0003	17%

Training (with broker) - batch_size = 128, epochs = 50

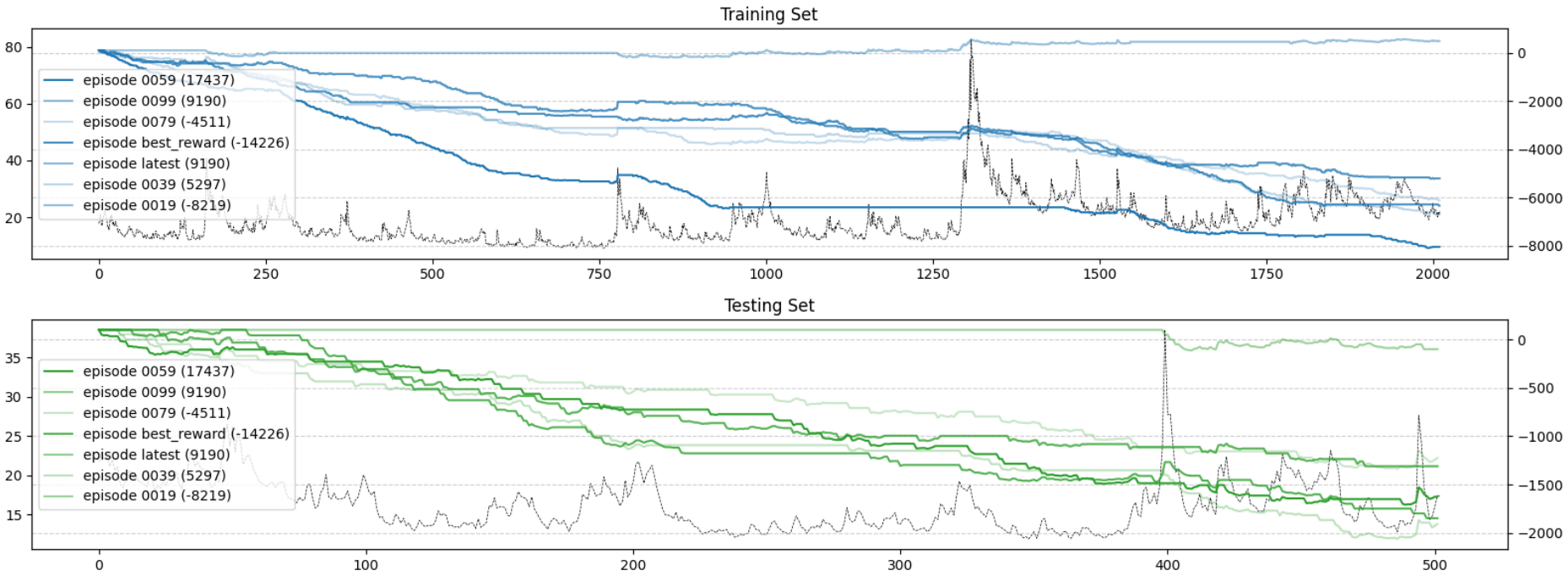


Figure 7: Task 4 - Analysis plot

Training (with broker) - batch_size = 128, epochs = 50, episode = 19

Figure 8: Task 6 - Episode 19 Test

Episodes	Epochs	Tick	Batch Size	Learning Rate	Final Profit on test
19	50	Daily	256	0.0003	17%

Latest Results - batch_size = 256, epochs = 50, episode = 99

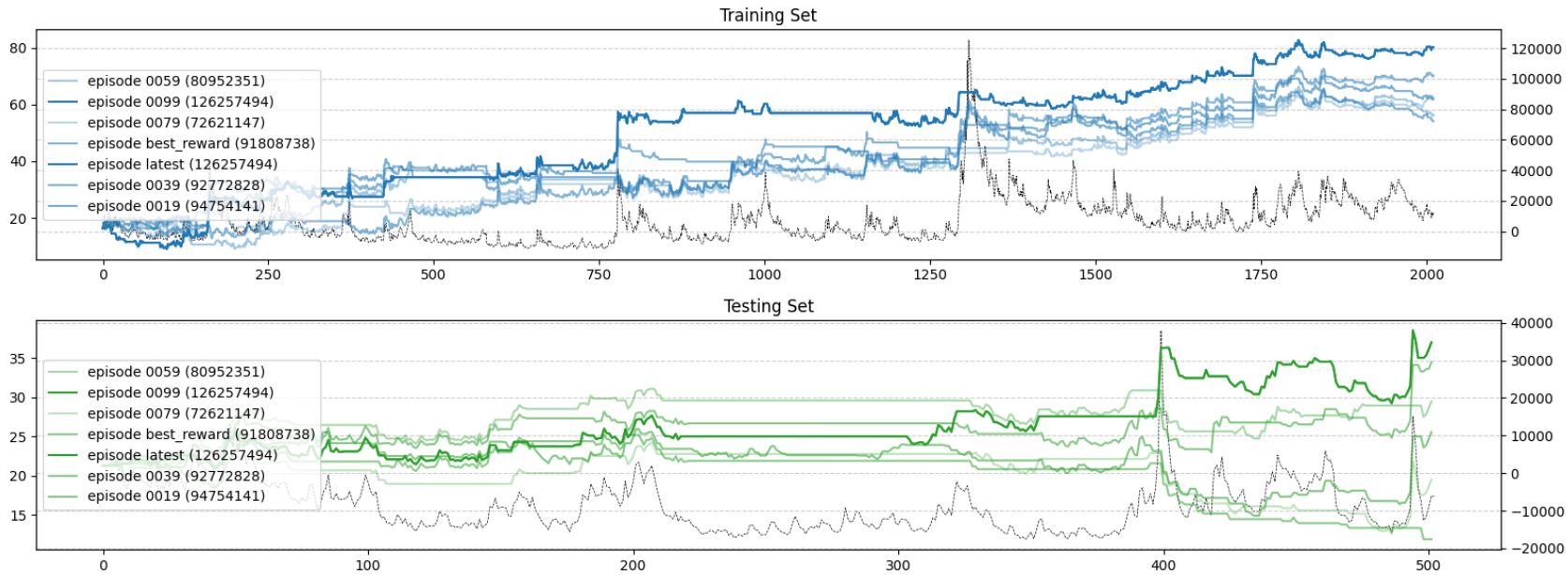


Figure 9: Task 0 - Analysis

Latest Results - batch_size = 256, epochs = 50, episode = 99

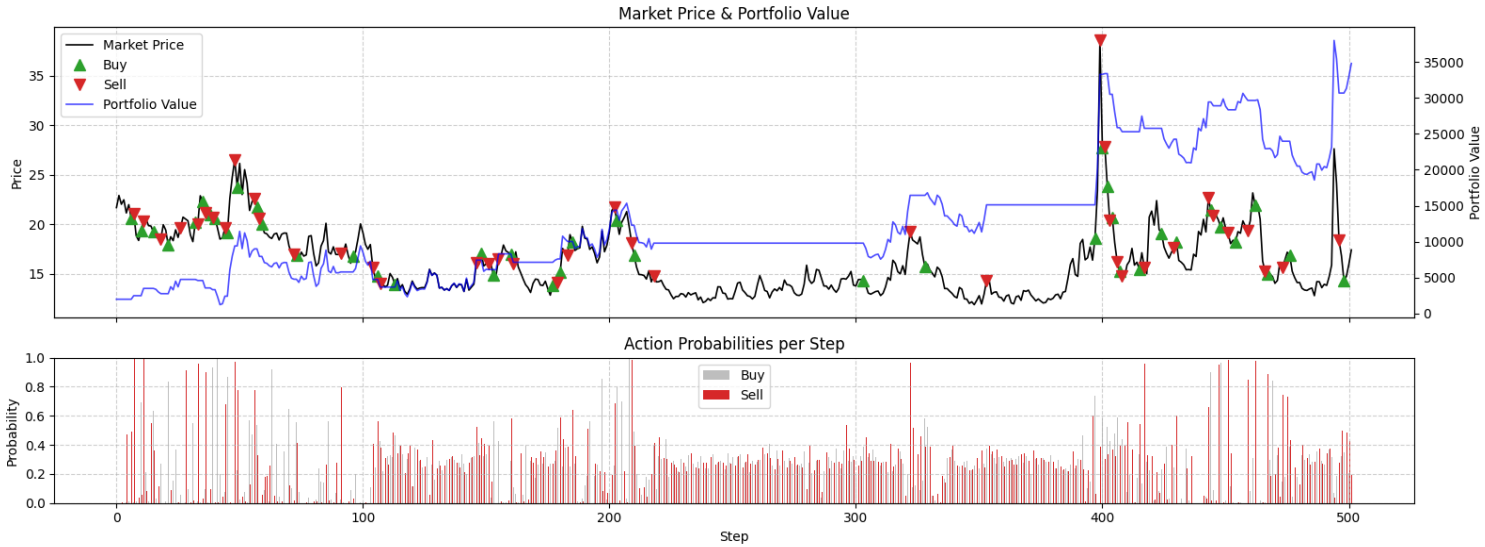


Figure 10: Task 0 - Test

Episodes	Epochs	Tick	Batch Size	Learning Rate	Final Profit on test
99	50	Daily	256	0.0003	15%

annual return	avg profit	annual vol.	max drawdown
0.067	0.001	3.12%	3%

Workflow - What's next

- Changing the dataset tick to a 5-minutes basis
- Add the S&P500 market and VVIX index
- It has been suggested to use the following signal :

(a)

$$\begin{aligned} a_{\text{signal}} &= VVIX_{\text{index}} \times VIX_{\text{index}} \times a_{\text{signal}} \\ &= \frac{VVIX_{\text{index}}}{VIX_{\text{index}}} \times a_{\text{signal}} \\ &= VIX_{\text{index}} \times VVIX_{\text{index}} \,(\text{ratio}) \end{aligned}$$

(b)

$$b_{\text{signal}} = \ln\left(VVIX_{\text{index}}\right) - \ln\left(VIX_{\text{index}}\right)$$

- Fine-tune another time with the new training features

Thank You