

PPO Market Bot

Mathys Vinatier

Seoul National University

November 25, 2025



New Environment

Modifications

Added **5-minute tick** to the state space

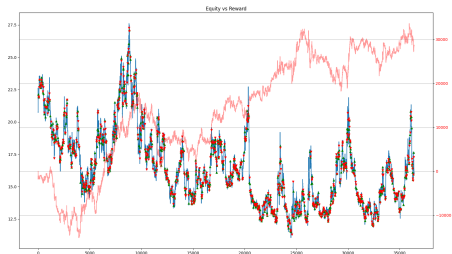
Enabled **short positions**

Optuna optimization on **max drawdown** and **average profit**

Issues Encountered

RAM saturation: long training runs crashed the OS

VRAM limitation: sequence length reduced to **1 day** (288 samples)



Training dataset

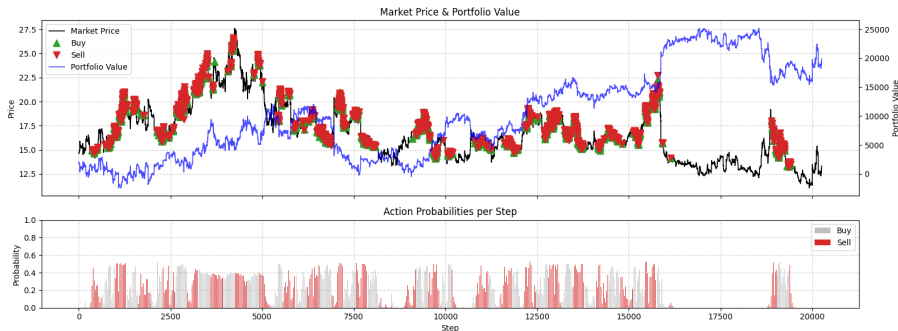


Figure 2: Agent performance on the training dataset



First results

Model Training Result (threshold=0.65)	
total_reward	163,491,909
action_per_day	6
annual_return	0.29
average_profit	-0.08%
annual_volatility	22.92
sharpe_ratio	-0.90
max_drawdown	-1.49

Figure 3: Training phase results



First results

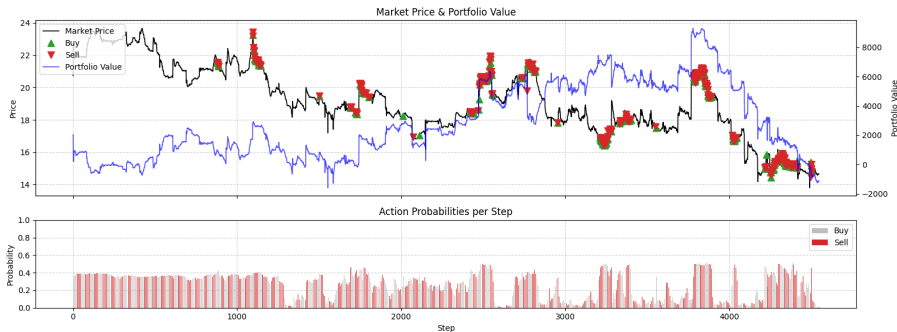


Figure 4: Agent performance on the testing dataset



First results

Model Testing Result (threshold=0.65)	
total_reward	3,960,964
action_per_day	4
annual_return	nan
average_profit	-0.186%
annual_volatility	195.909
sharpe_ratio	-0.240
max_drawdown	-1.547



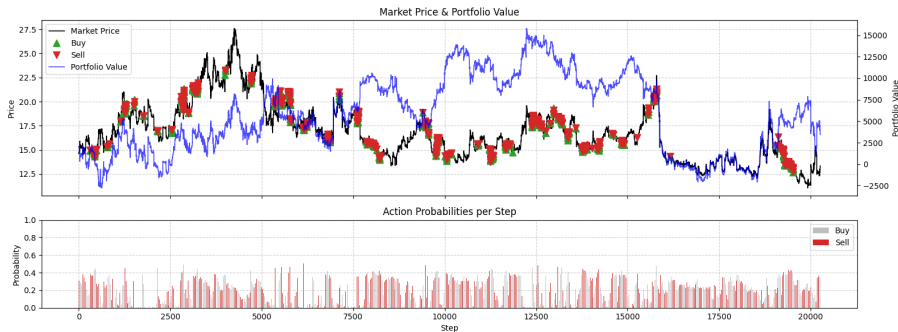


Figure 5: Agent performance on the training dataset





Figure 6: Agent performance on the training dataset



Model Testing Result (threshold=0.65)	
total_reward	6,431,756
action_per_day	6
annual_return	0.113
average_profit	-0.034%
annual_volatility	54.216
sharpe_ratio	-0.203
max_drawdown	-1.398



Thank You

