

Risk-Aware Deep Reinforcement Learning with Hierarchical Adaptation for XAU/USD Trading

XAU/USD trading is the process of speculating on the future price movements of gold (XAU) against the US dollar (USD). This speculation relies on historical price data, which exhibits constantly evolving dynamics due to shifting market conditions. Traditional trading, limited by human biases, struggles with inconsistent decisions, poor risk management, and adapting to volatile markets, while on-policy RL methods like PPO and A3C fail to leverage historical data for generalizable patterns, critical for adapting to shifting conditions with limited data. And discrete-action algorithms cannot dynamically adjust trade and leverage sizes, and existing methods lack robust exploration to handle changing market regimes (trending, volatile, flat)

To address these limitations, we propose a hierarchical reinforcement learning framework using Soft Actor-Critic (SAC) to enable adaptive, risk-conscious trading. We begin by setting up the trading environment with 1-hour XAU/USD candlestick data (2010-2020), incorporating transaction costs and slippage for realism. State representations are constructed using Gramian Angular Field heatmaps to capture price patterns, volatility indicators like ATR to quantify market conditions, and binary flags for economic events to signal regime shifts, all normalized for stability. The low-level SAC agent is trained on these states to optimize trading actions (position size, leverage) using a risk-aware reward function balancing profit, drawdown, and exploration, while the meta-controller learns on meta-states, such as aggregated volatility or CUSUM test results, to detect market regime shifts (trending, volatile, flat) and select sub-policies (e.g., adjusting leverage for aggressive or conservative mode) to guide the SAC agent. Hyperparameters are optimized via Optuna to maximize Sharpe ratio.

Subsequently, online adaptation is enabled through periodic or triggered retraining on 2020-2025 data, updating the replay buffer to incorporate recent market dynamics for continuous learning. The framework is then backtested on 2020-2025 data across trending, volatile, and flat regimes, evaluating Sharpe ratio and maximum drawdown, and compared against benchmarks like moving average crossovers (rule-based), DQN (discrete-action), and PPO (on-policy) to ensure robust performance in non-stationary markets.

This work offers a scalable framework for volatile assets like XAU/USD, with potential for real-time trading and extension to other markets, advancing algorithmic trading research and practice.

References

1. Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft Actor-Critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *Proceedings of the 35th International Conference on Machine Learning (ICML)*, 1861-1870.
2. Nachum, O., Gu, S., Lee, H., & Levine, S. (2018). Data-efficient hierarchical reinforcement learning. *Advances in Neural Information Processing Systems (NeurIPS)*, 3303-3313.
3. Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2017). Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3), 653-664.