Risk-Aware Deep Reinforcement Learning for Crypto and Equity Trading Under Transaction Costs

Author: Ekantheswar Bandarupalli
Date: October 2025

Abstract

We present a reinforcement learning (RL) trading agent that optimizes risk-adjusted returns in volatile markets by explicitly penalizing drawdowns and turnover. Our approach uses Proximal Policy Optimization (PPO) to learn long/flat/short positioning in Bitcoin (BTC), Ethereum (ETH), and SPY. The learned RL policy achieved a Sharpe ratio of 1.23 versus 1.46 for a buy-and-hold benchmark, with a final NAV of 1.916 compared to 2.213. Although underperforming in raw and risk-adjusted returns, this outcome highlights the sensitivity of RL performance to risk–reward balance.

Introduction

This work explores the use of reinforcement learning in risk-aware financial trading. Unlike profit-only optimization, our PPO-based agent explicitly balances drawdown, volatility, and transaction cost, modeling real-world capital efficiency.

Methodology

Data from 2020–2024 was used for BTC, ETH, and SPY. The RL agent's reward includes transaction costs and volatility penalties. Training used PPO with λ=0.1 and 5bps transaction cost. Out-of-sample evaluation was conducted on 2024 data.

Results

The RL model underperformed buy-and-hold (NAV 1.916 vs 2.213; Sharpe 1.23 vs 1.46). Drawdown was -30.16% vs -26.18%. This indicates over-penalization of risk, leading to conservative trading behavior.

Discussion

The results reveal the importance of balancing stability and profitability in reward design. Excessive risk penalties can suppress exploration, while minimal penalties may lead to overtrading. Ongoing work focuses on adaptive volatility-based penalties and regime-dependent reward scaling.

Conclusion

Although the PPO agent did not outperform benchmarks, it demonstrated interpretable, risk-sensitive decision behavior. The study contributes to understanding how reward shaping influences agent behavior and provides a transparent basis for future research.