Reinforcement Learning with Proximal Policy Optimization for Strategic Betting in Counter-Strike 2 Esports

Nathan Ho
Drexel University
Philadelphia, PA, USA
nlh55@drexel.edu

Matthew Protacio
Drexel University
Philadelphia, PA, USA
mp456@drexel.edu

Alexey Kuraev
Drexel University
Philadelphia, PA, USA
ak789@drexel.edu

Abstract

This project explores the application of Proximal Policy Optimization (PPO), a reinforcement learning algorithm, to develop an intelligent betting agent for esports matches. We evaluate the agent's performance and decision-making in a simulated betting environment using historical match data.

1 Introduction

Sports betting is a widely practiced recreational activity in which individuals place wagers on specific outcomes of sporting events. Bet types range broadly, including predicting specific occurrences during a game or determining the ultimate winner of a match. This paper specifically explores *moneyline* bets, where wagers are placed solely on the final result of a contest.

Despite its popularity, sports betting remains underrepresented as a quantitative research domain, partly due to its association with gambling, which contributes to limited academic inquiry and systematic study. Predicting winners accurately poses considerable challenges, as match outcomes are inherently stochastic due to significant variability in team performance and individual player dynamics.

Effective betting strategies often hinge on identifying edges, which involve detecting discrepancies between bookmaker odds and bettors' valuations. Precisely computing match odds from fundamental analyses demands extensive computational resources. However, by considering established bookmakers' odds as a reliable proxy for the market's perceived fair value, we circumvent extensive calculations while still gaining actionable insights.

Applying these concepts to professional esports introduces unique complexities. State representation in video games can lead to state explosion due to numerous exogenous variables. Moreover, continual game updates and modifications create unstable environments where strategies effective in one period may become obsolete in the next.

Nonetheless, the esports title *Counter-Strike 2 (CS2)* offers specific advantages for quantitative modeling. Economic management significantly influences gameplay outcomes, with team bankroll serving as a critical predictive indicator. Within CS2, team economics are determined by several well-defined factors, including weapon and equipment purchases, income from winning rounds, and earnings from eliminating opponents. Matches typically span best-of-13 rounds, allowing temporal economic analysis across discrete intervals.

While raw economic indicators provide foundational insight, their direct application can suffer from excessive noise and limited predictive value. To enhance signal strength, we compute more sophisticated financial metrics, including teams' Return on Investment (ROI), implied probability derived from bookmaker odds, ROI based on betting odds, and Cost Per Kill (CPK). Interpreting a CS2 team as an evolving market asset allows us to aggregate performance metrics over time, offering valuable temporal context for predictive modeling.

This paper applies reinforcement learning—specifically, Proximal Policy Optimization (PPO)—to leverage these economic indicators for betting decisions. We further explore reward function formulations beyond simple correctness, investigating the impact of Expected Value (EV) calculations and the Kelly Criterion on betting efficacy. In doing so, we examine the limitations of traditional betting methodologies, exploring whether integrating financial and probabilistic principles yields improved betting outcomes in esports scenarios.

2 Introduction

Hello!