# Portfolio Optimization Using Reinforcement Learning

## Project Statement

This project explores the use of reinforcement learning (RL) to dynamically optimize portfolio allocations across 29 diversified stocks, aiming to maximize risk-adjusted returns in an evolving financial environment. By integrating company 10-Q financial statements, news sentiment, and macroeconomic indicators into a comprehensive state representation, the study applies advanced RL algorithms—Deep Q-Network (DQN), Rainbow DQN, and Proximal Policy Optimization (PPO)—to inform quarterly rebalancing decisions.

## Data Integration & Framework

1. **Financial Fundamentals**: Fundamental ratios were calculated across activity, liquidity, solvency, and profitability dimensions to assess company performance and financial health.
2. **Market Sentiments**: News sentiment was extracted using the NYT API and analyzed with FinBERT, generating five sector-level metrics to capture sentiment trends and dynamics.
3, **Macroeconomics Context**: Quarterly macroeconomic indicators from FRED via Pandas Datareader offered insights into systematic risk through key economic measures and financial market indicators.
4. **Market Performance**: Historical daily return data from WRDS (CRSP) was used to benchmark asset performance, calculate realized volatility, and support risk-adjusted return evaluations.

## Methodology

We used three Reinforcement Learning approaches, the agent observes multi-dimensional states, takes actions through portfolio weight adjustments, and receives rewards based on risk-adjusted returns.
1. **Deep Q-Network**: Implemented using a 3-layer neural network with experience replay (10,000), epsilon-greedy strategy ($\varepsilon$=1.0→0.01), and gradient clipping for stability.
2. **Rainbow DQN**: Implemented with a dueling architecture (128/64 neurons) for better portfolio allocation, using epsilon-greedy ($\varepsilon$=0.1), $\gamma$=0.99, experience replay (10,000), and Adam optimizer over 300 episodes.
3. **Proximal Policy Optimization(PPO)**: Implemented continuous action space for portfolio weights, with rewards based on fundamental outperformance, sentiment momentum, and macroeconomic alignment.

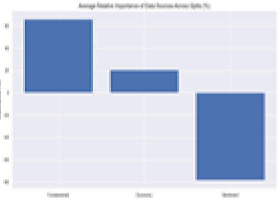
## Results

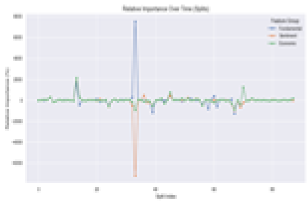
We implemented all three models and PPO outperformed other two in terms of stability, risk adjustment, and more realistic estimation.
1. **Sharpe Ratio Comparison:** The RL model achieved a higher Sharpe ratio (**1.35)** than both even-weighted and mean-variance portfolios (**1.03**), with lower variability.


2. **Portfolio Performance:** The PPO model delivered strong performance, with an average Sharpe ratio of **1.07**, annualized return of **45%**, and mean cumulative return of **2.40**. Risk was well-managed with a **-0.10** max drawdown and validated by high risk-adjusted metrics like the Calmar (**14.2**) and Sortino (**0.18**) ratios.

| | Total_Return | Sharpe_Ratio | Max_Drawdown | Win_Rate | Calmar_Ratio | Sortino_Ratio |
|---|---|---|---|---|---|---|
| mean | 0.027218 | 1.072408 | -0.103060 | 0.537140 | 14.204583 | 0.184566 |
| std | 0.119506 | 3.393978 | 0.074330 | 0.101476 | 35.070635 | 0.446076 |
| min | -0.326579 | -8.884033 | -0.345137 | 0.292683 | -4.133912 | -0.714463 |
| max | 0.398749 | 10.591986 | -0.008627 | 0.731707 | 249.147783 | 1.985231 |

3. **Contribution of Factors:** Fundamental ratios drove **70%** of portfolio performance, with economic indicators (**20%)** and sentiment momentum (**10%**) complementing short- and long-term market alignment.





4. **Dynamic Rebalancing Impact:** It improved risk-adjusted returns, boosting the Sharpe Ratio by **31.2%** over static benchmarks through adaptive asset allocation.

Model vs Baseline Performance Comparison:

| | RL_PPO | Even_Weighted | Minimum_Variance |
|---|---|---|---|
| total_return | 0.050405 | 0.040754 | 0.040754 |
| annualized_return | 0.352899 | 0.270830 | 0.270830 |
| sharpe_ratio | 1.348370 | 1.027684 | 1.027684 |
| max_drawdown | -0.158233 | -0.140422 | -0.140422 |
| win_rate | 0.658537 | 0.619048 | 0.619048 |
| cumulative_return | NaN | 0.040754 | 0.040754 |
| volatility | NaN | 0.242879 | 0.242879 |
| calmar_ratio | NaN | 1.928680 | 1.928680 |
| sortino_ratio | NaN | 0.092362 | 0.092362 |

Relative Improvement vs Even-Weighted Portfolio:

| | RL_PPO | Even_Weighted | Minimum_Variance |
|---|---|---|---|
| Sharpe_Ratio_Improvement | 0.320687 | 0.0 | 0.0 |
| Relative_Improvement(%) | 31.204795 | 0.0 | 0.0 |

5. **PPO Model - Cumulative Returns Performance**:


## Discussion

**PPO's** conservative optimization strategy of making small gradient-based changes aligns perfectly with investment principles that prioritize stability and risk management. Other model's strength and weaknesses were:

**DQN**
*Strength:* Simple and fast to implement with basic reinforcement learning components like experience replay and Q-function approximation.
*Weakness:* Produced unrealistic returns and stable allocations under volatile markets, indicating overfitting and potential data leakage.

**Rainbow DQN**
*Strength:* Achieved high Sharpe ratio and robust performance using advanced techniques like distributional Q-learning and Noisy Nets.
*Weakness:* Computationally intensive with occasional instability in late episodes and sensitivity to extreme market conditions.

**Github**: https://github.com/bhatt-prajakta/Capstone

stephaj@umich.edu, pbumich@umich.edu, jisioma@umich.edu