TAIS NextGen-Tunisian AI Society

SympactAI

# Hands-On Reinforcement Learning For Stock Trading Using FinRL


**Supervised by**

Mr.Mahdi Kallel

Mrs.Imen Ghazouani

**Presented by**

Moez KESSEMTINI

Mazen TMANI


**Saturday, August 16, 2025**

# Outline

01

Opening Hook

02

RL &
Stock Trading

03

FinRL & Trading
Environment

04

RL Workflow in
FinRL

05

Results &
Evaluation

06

Future Directions

01

# Opening Hook

# Problematic


About 95% of retail traders lose money due to behavioral biases like overconfidence and poor risk management (Medium 2022)

**Jane Street (India, 2023–2025)**: Retail investors lost $21.7B in derivatives trading (Reuters, 2025)

# Can algorithms and AI solve those Problems ?

# Motivation


- In 2023, **37%** of overall U.S. equity trading volume was executed through algorithms

- The global reinforcement learning market is expected to grow from \$2.8B in 2022 to \$88.7B by 2032

What about implementing RL algorithms in trading ?

# 02

# RL
# &
# Stock Trading

# Core Concepts of Reinforcement Learning

**Agent:** the learner or decision-maker

**Environment:** the system or market the agent interacts with

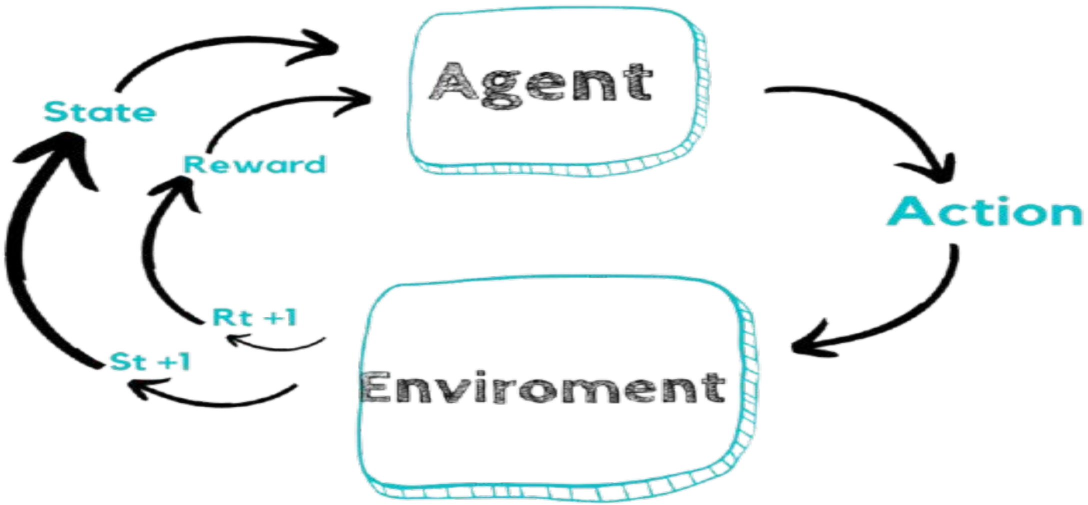
**State:** current situation or information available to the agent

**Action:** decisions or moves taken by the agent

**Reward:** feedback received based on the action's outcome

**Policy:** strategy that maps states to actions

**Value Function:** estimates expected future rewards



# Why RL is Ideal for Stock Trading

## Dynamic & Uncertain Markets

RL agents learn by trial and error, adapting to changing market conditions

Strategies continuously evolve by learning from new market data and adapting to emerging trends

## Advantages over Traditional Methods:

Adapts in real-time

Handles complex, nonlinear relationships

Incorporates risk management constraints (drawdown limits, portfolio volatility)

# 03

# FinRL & Trading Environment

# What is FinRL?

- Open-source RL framework developed by **AI4Finance**
- Built on: PyTorch, Stable Baselines3, Gym interface
- Pre-built trading environments
- Integration with financial data sources (Yahoo Finance, Quandl...)
- ideal for experimentation, research, and practical RL trading projects

# FinRL Trading Environment

## Components

- **State space**: prices, technical indicators, portfolio info
- **Action space**: buy, sell, hold (and position sizing)
- **Reward function**: portfolio value, Sharpe ratio

## Realistic constraints

- Transaction costs
- Risk management parameters

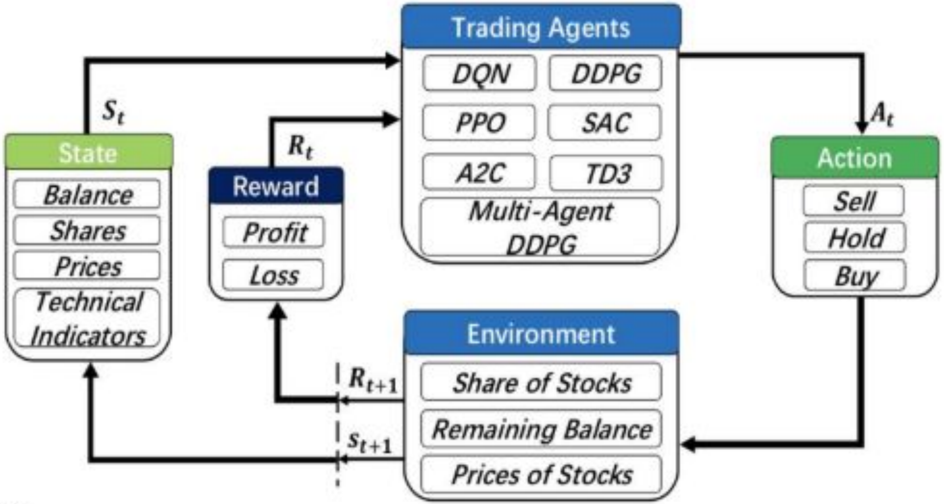
# Automated Trading in FinRL

## Agent:

- Represent the investor/trader.
- Perceive the current state in the market.
- Take an action for trading.
- Get rewards from the action.
- Learn from past trading experience.

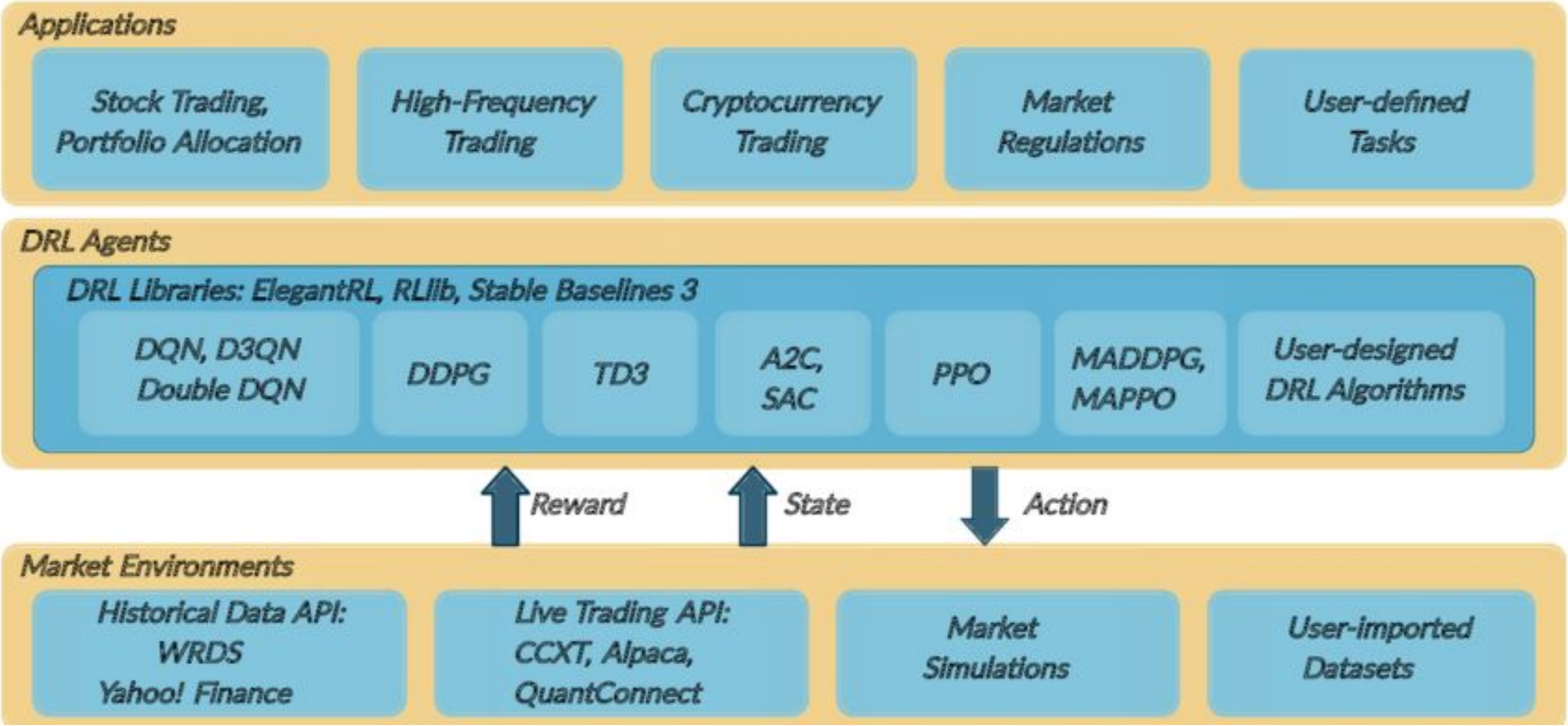
## Environment:

- Represent the trading market.
- State is a snapshot of current market conditions.
- Receive the trading action and step into a new state.
- Return the reward for the trading action.



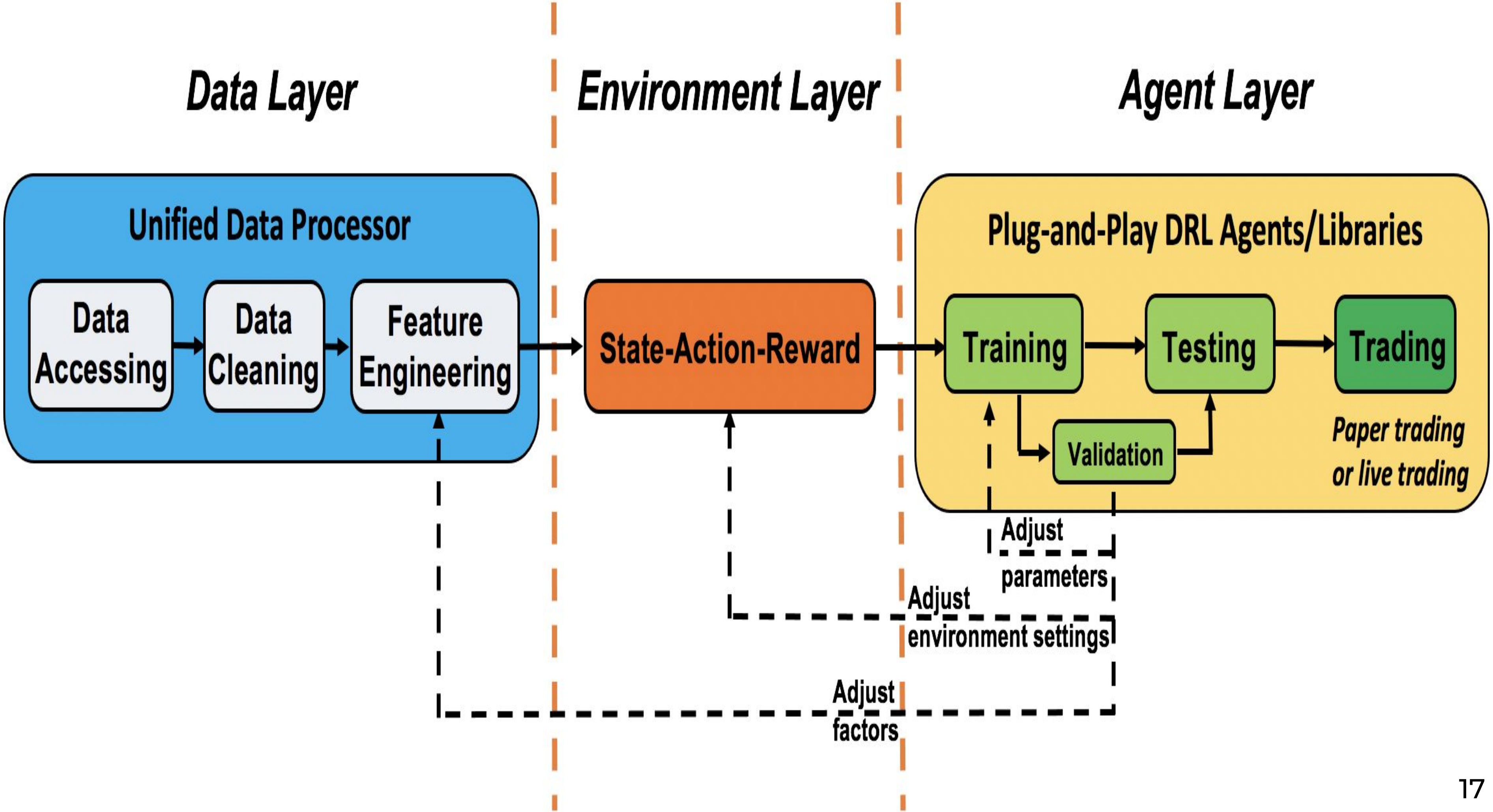

Overview of automated trading in FinRL, using deep reinforcement learning.

# FinRL Layers :



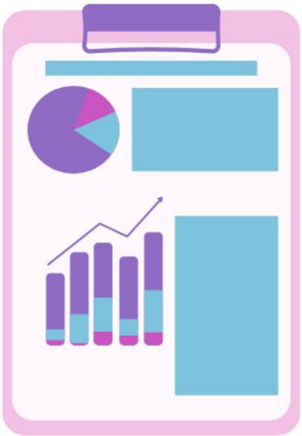
04

# RL Workflow in FinRL

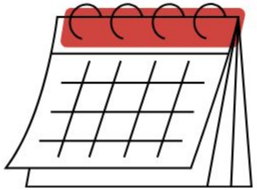


Data Layer
Environment Layer
Agent Layer
Unified Data Processor
Data Accessing
Data Cleaning
Feature Engineering
State-Action-Reward
Plug-and-Play DRL Agents/Libraries
Training
Testing
Trading
Validation
Paper trading or live trading
Adjust parameters
Adjust environment settings
Adjust factors

# Data Collection



Yahoo Finance
(daily OHLCV data)



2010 → 2025

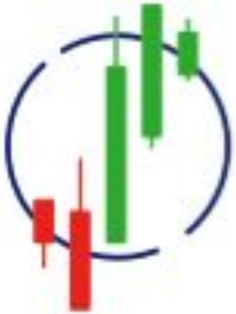
30 Dow Jones
Industrial Average stocks

| Price | date | close | high | low | open | volume | tic | day |
|---|---|---|---|---|---|---|---|---|
| 0 | 2010-01-04 | 6.424604 | 6.439314 | 6.375672 | 6.407193 | 493729600 | AAPL | 0 |
| 1 | 2010-01-04 | 39.913254 | 40.016977 | 39.111116 | 39.159521 | 5277400 | AMGN | 0 |
| 2 | 2010-01-04 | 32.637966 | 32.781535 | 32.215237 | 32.550232 | 6894300 | AXP | 0 |
| 3 | 2010-01-04 | 43.777550 | 43.941189 | 42.702201 | 43.419101 | 6186700 | BA | 0 |
| 4 | 2010-01-04 | 39.403461 | 39.834174 | 38.703553 | 38.797774 | 7325600 | CAT | 0 |

# Data Preprocessing

# Stock Trading Environment

291 Dimensional vector



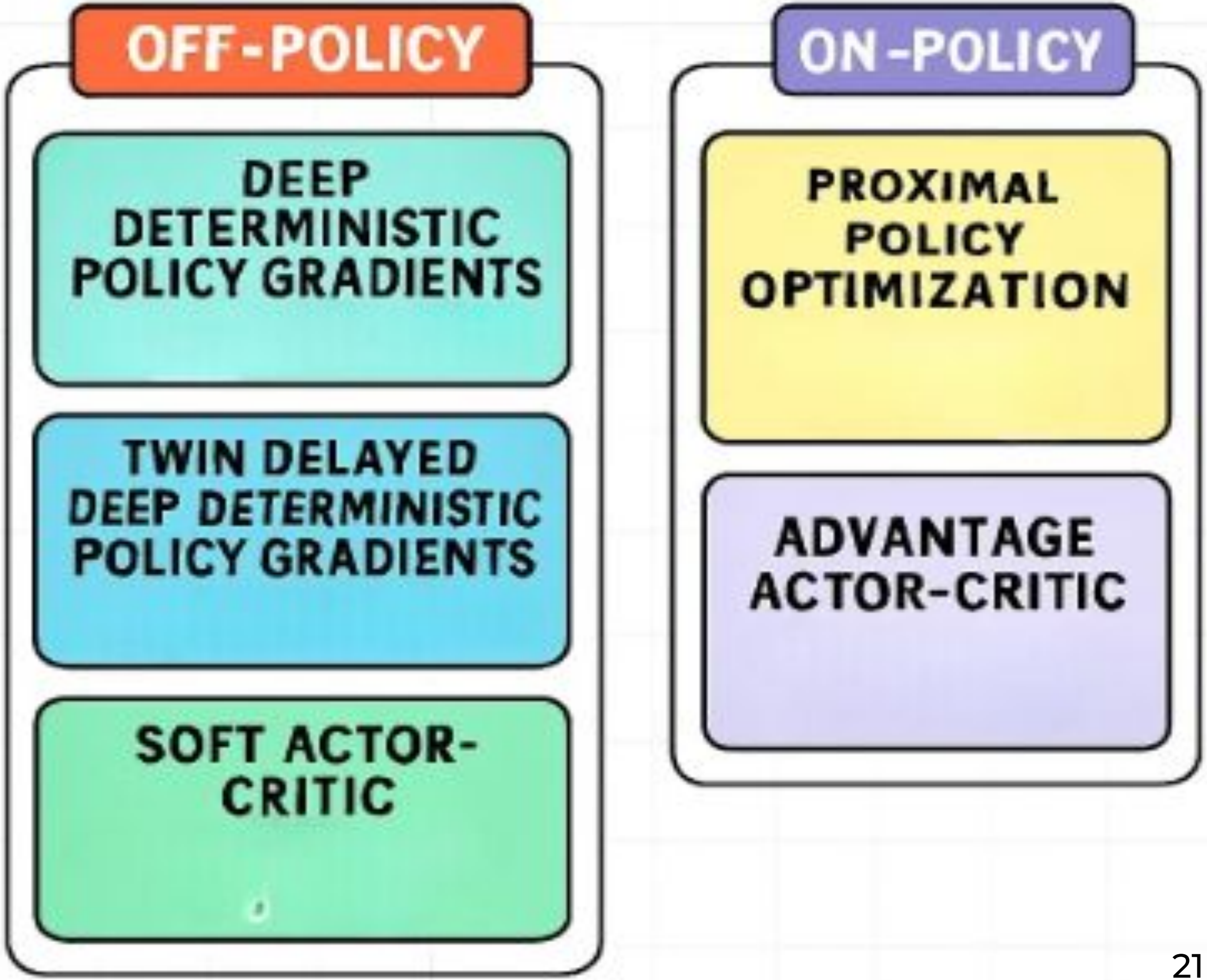
[-1 .. 1]

Reward Function

Transaction cost (0.001)


StockTradingEnv

# RL Algorithms and Libraries



| Library | Pros | Cons |
|---|---|---|
| Stable-Baselines3 ✓ | Simple to set up, well-documented, large community, stable & reproducible results, widely used in finance | Limited distributed training |
| ElegantRL | Fast GPU training, flexible for custom algorithms | Smaller community, less documentation |
| RLlib | Highly scalable, supports massive parallelism | Complex setup, slower for small/medium tasks |

# Agents Training

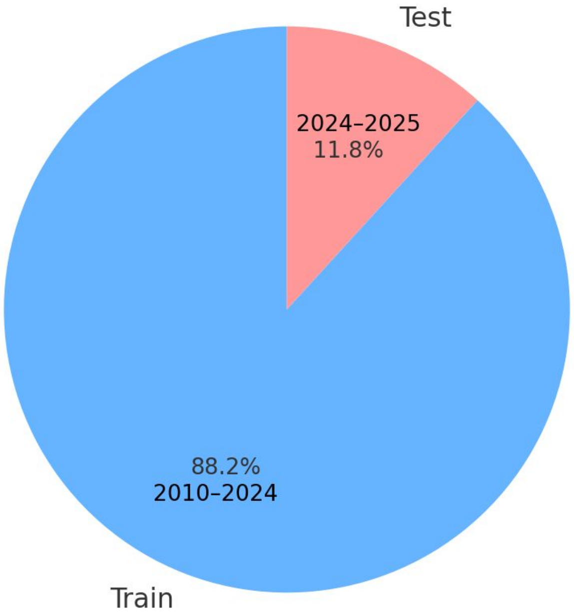
Training Paradigms

**Static Training:** Train once on historical data, then test on future unseen data.

**Rolling Training:** Periodically retrain with new data to adapt to evolving market conditions.

# Data Distribution :

Static Training




Rolling Training

Retrain periodically during 2024–2025 to capture market regime shifts

# Trade-offs

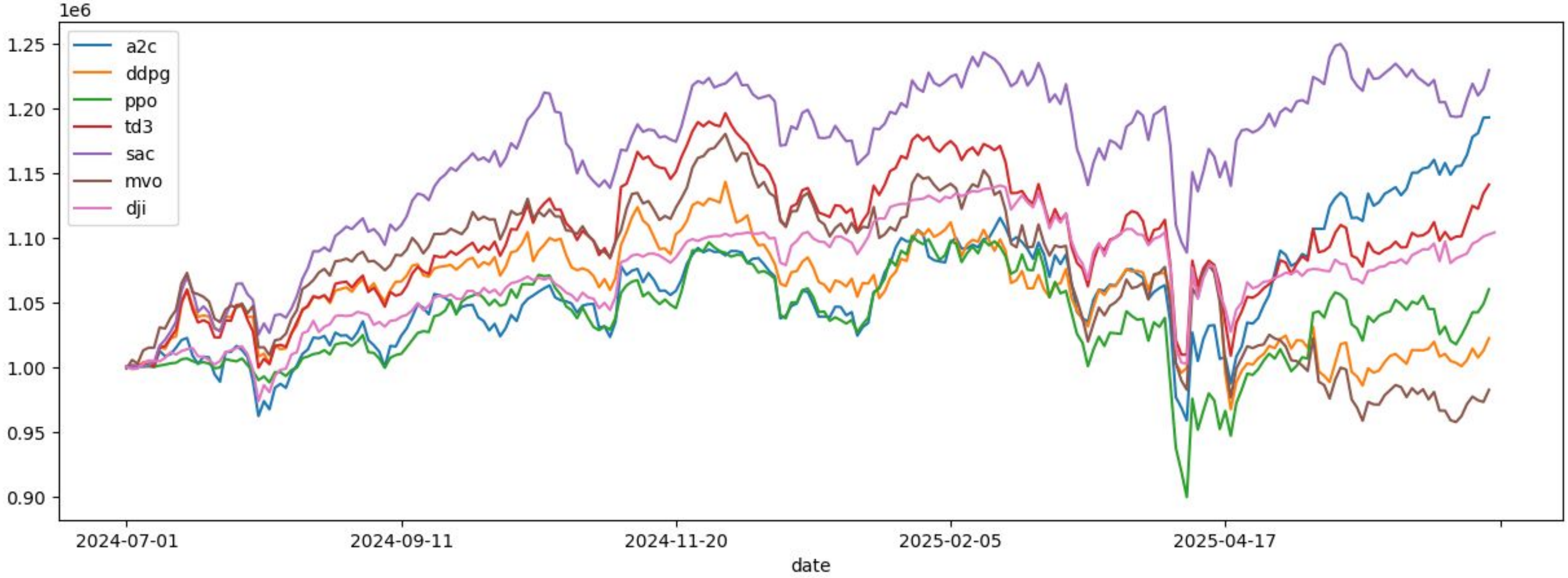
**Static**: Simple, low computational cost, but not adaptive to market changes

**Rolling**: Adapts to new data and market shifts, but higher computational cost.

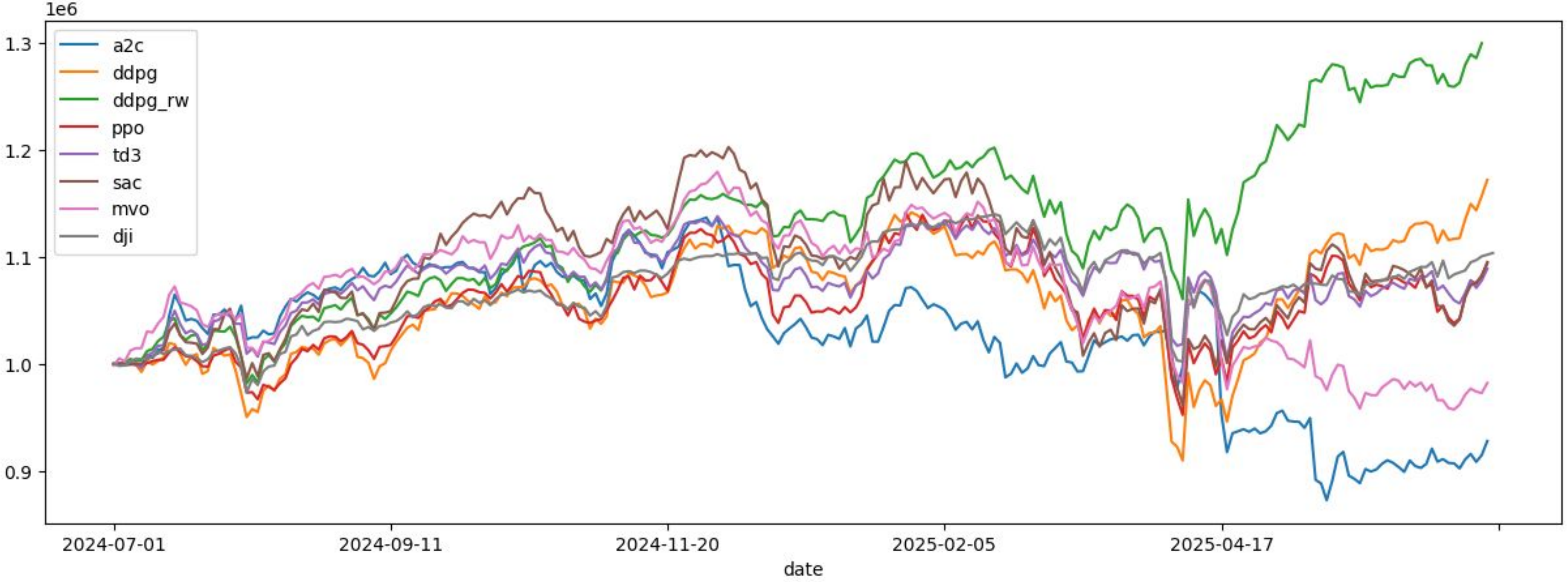
# 05

# Results & Evaluation

# Cumulative returns for DRL agents and benchmarks



# Performance with DDPG retrained via expanding rolling window



# Selected performance metrics for all Agents

| | PPO | DDPG | DDPG_RW | A2C | TD3 | SAC |
|---|---|---|---|---|---|---|
| Annual return | 0.095 | 0.174 | **0.306** | -0.073 | 0.090 | 0.097 |
| Sharpe ratio | 0.603 | 0.926 | **1.649** | -0.268 | 0.640 | 0.571 |
| Calmar ratio | 0.575 | 0.858 | **2.591** | -0.313 | 0.768 | 0.481 |
| Max drawdown | -0.166 | -0.203 | **-0.118** | -0.233 | -0.117 | -0.201 |
| Stability | 0.143 | 0.187 | **0.793** | 0.501 | 0.067 | 0.003 |

**Agent performance varies by algorithm and settings**

**Using rolling window training, DDPG outperforms all other RL algorithms**

**The rolling window strategy proved effective in boosting returns and reducing risk**

# 06

# Future Directions

- **Integration with News & Sentiment Analysis**: leverage financial news and social media for improved decision-making

- **Real-Time Trading Deployment**: deploy RL agents in live or paper trading environments

- **Ensemble of RL Agents**: combine multiple RL strategies for more robust performance

# Thank you for your attention