

Optimizing Stock Trading Strategy with Reinforcement Learning

Project Aim:

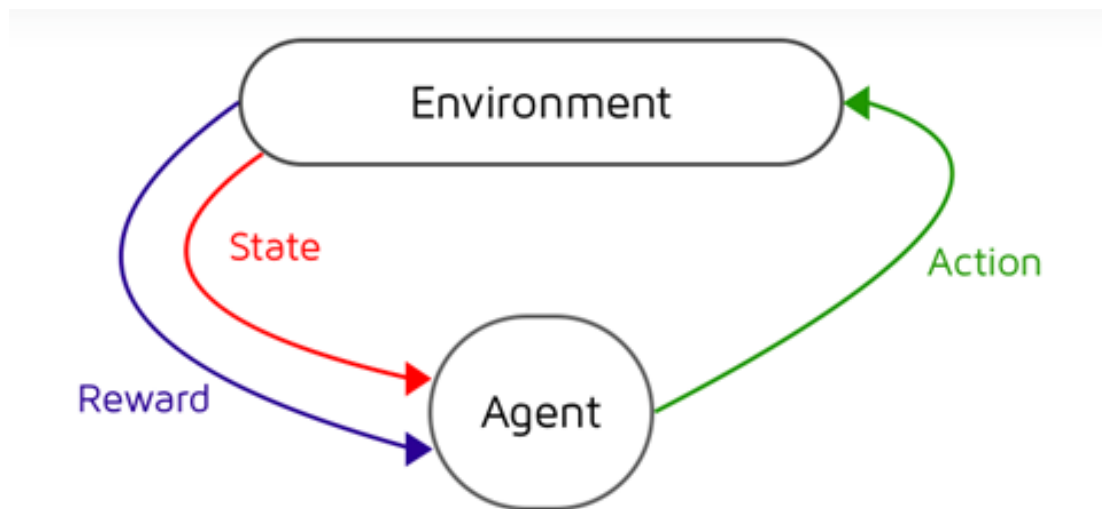
Our goal is to find an investment strategy on a single exchange-traded fund (SPY) that would maximize our cumulative wealth. We benchmark our results against the tried-and-true “buy-and-hold” and MACD techniques.

Reinforcement Learning:

You let the dog do whatever it wants and then whenever it behaves well you go, “Good dog” and when it misbehaves you go “Bad Dog!” and then over time dog learns to do more of the good dog things and fewer of the bad dog things.

Markov Decision Process (MDP):

It is a mathematical framework used for modelling decision-making problems where the outcomes are partly random and partly controllable.



Q-learning:

The 'Q' in Q-learning stands for quality. Quality in this case represents how useful a given action is in gaining some future reward. Q-learning is an off-policy reinforcement learning algorithm that seeks to find the best action to take given the current state. It's considered off-policy because the q-learning function learns from actions that are outside the current policy, like taking random actions, and therefore a policy isn't needed. More specifically, q-learning seeks to learn a policy that maximizes the total reward.

Hill Climbing Algorithm:

Hill climbing algorithm is a local search algorithm that continuously moves in the direction of increasing elevation/value to find the peak of the mountain or the best solution to the problem. It terminates when it reaches a peak value where no neighbour has a higher value. In this algorithm, we don't need to maintain and handle the search tree or graph as it only keeps a single current state.

Recurrent Reinforcement Learning:

Recurrent reinforcement learning (RRL) was first introduced for training neural network trading systems in 1996. "Recurrent" means that previous output is fed into the model as a part of the input. The RRL technique is a successful machine learning technique for building financial trading systems. The distinction of (Deep) Recurrent RL, is that the function mapping the agent's observations to its output action is a Recurrent Neural Network. A Recurrent Neural Network is a type of neural network that processes each observation sequentially, in the same way for each time step.

Points to consider:

1. Gamification of trading - Historical & Current Prices, Technical Data, Buy/Sell/Do Nothing, Profit & Loss.
2. Train the system - Each entry and exit is an individual game, Run through the price series sequentially and randomly.
3. Reward function design - Pure Profit & Loss on exit, otherwise zero, PnL from the start of trade to every time step t.
4. Features to use - Open, High, Low, Close, Volume (OHLCV), Technical indicators, Time of day, Day of Week, Time of year, Different time granularity.
5. Test the system - Sine waves, Trend curves, Random walks, Adding Noise to clean test curves.
6. Types of Deep Learning Algorithm.

Steps:

1. Import necessary libraries.
2. Load Dataset.
3. Exploratory Data Analysis.
4. Prepare Data.
5. Train Data.
6. Test Data.
7. Evaluate Model.
8. Deploy Model.