Application Of Deep Reinforcement Learning in Stock trading Strategies

Finance and Analytics club

Overview

- Necessity for stock trading strategies
- Assignment - 1
- Reinforcement learning
- Q-value
- Q-learning
- Exploration vs Exploitation
- Deep Q-learning

Necessity for stock trading strategies

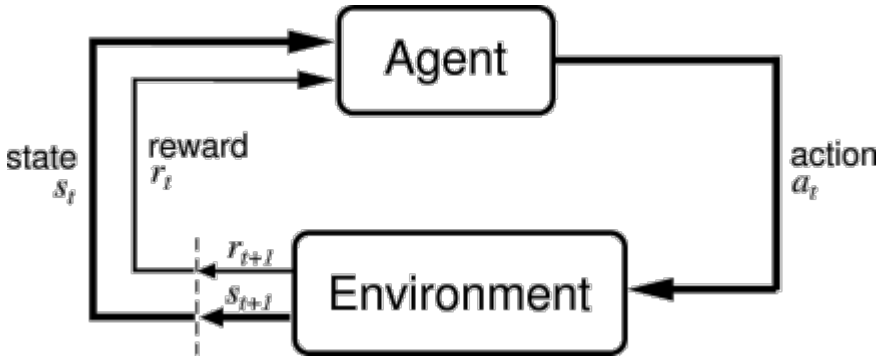
- Suggests weather to Buy, Sell or remain Neutral
- Patterns are detected based on previous data to help make better decisions
- Shows a clear path amid huge inflow of information
- Reduce the emotional interference caused by humans

Assignment - 1

- Task:
  - To take 10 years weekly data for any 3 firms and predict their Adjusted Closing price using at least 5 different algorithms
- Performance by participants:
  - Linear Regression, Lasso and Ridge Regression, Elastic Net Regression, Polynomial Regression, SGD Regressor, SVM Regression, KNN Regression, Moving Averages Method, Neural Networks, etc. are used by various participants.

Reinforcement learning

- One of the 3 ML paradigms
- Agent:
  - Learner and decision maker
- Environment:
  - Everything outside agent with which agent interacts
- Agent interacts with environment to find itself in new scenario
- Goal is to maximize a reward function over time



Q-value

- It is a measure of overall expected reward over a period of time

$$Q(s,a) = \mathbf{E}\left[\sum_{n=0}^{N} \gamma^n r_n\right]$$

Where,

s is the state of system

a is the action taken

N is the number of states from state s till the terminal state

γ is the discount factor

r^0 is the immediate reward received after performing action a in state s.

Q-learning

- Premise of this method is to choose the best action in every state
- Bellman equation for Q-Learning is

$$Q(s,a) = r + \gamma \max_{a'} Q(s', a')$$

- Q values of all state action pairs are determined by using the above equation
- Limitations:
  - When environment becomes complex, it will be computationally expensive to calculate all the state action pairs. (Chess game for example)
- Deep Q-learning solves this problem by using Deep neural networks to approximate the Q values

Exploration vs Exploitation

- Exploration is taking action whose outcome is not known yet to increase the knowledge base of agent
- Exploitation is taking known actions in the most optimal way to get the most reward from existing knowledge
- Pure exploration or pure exploitation are both bad. We need a proper mix
- Epsilon greedy policy suggest probability to choose weather to explore or to exploit in the next step.

Deep Q-learning

- This addresses the limitation of Q-learning.
- Aim of DQN is to train a deep neural network to learn the Q-value function and predict Q values of each state action pair.
- Bellman equation for DQN is

$$\text{Target } Q = r + \gamma \max_{a'} Q(s', a'; \theta')$$

- The cost function to train neural network is

$$L(\theta) = E[(\text{Target } Q - Q(s_t, a_t; \theta))^2]$$

Thank you!