Project Proposal

Predicting Stock Price using Actor-critic methods & Time-series Analysis

Objective

A strong financial system is important for a society thrive. With COVID sending people out of job or making people work from their homes, there has been a surge in people looking for investing in stock markets as a secondary source of income. Adding to this there are many companies adopting ML methods to ease people into the world of stocks & bonds. For these companies, time-series analyzing models are crucial for their operations. So, the project is to focus on combining ML models for algorithmic trading (time-series analysis) & automating trading process (actor-critic method), which can predict the price of stocks based on real life data.

Related work

From assignment 2, we having working knowledge of the function approximation & the Deep Q-Networks (DQN), which replaces the conventional Q-table for higher degree functioning. Actor-critic method, being similar to DQN, involves the usage of 2 interacting neural nets (NN) which returns a Q-score representing the value of taking that action given the state.

Algorithm 1 Q Actor Critic

Initialize parameters s, θ, w and learning rates α_θ, α_w; sample $a \sim \pi_\theta(a|s)$.
for $t = 1 \ldots T$: **do**
    Sample reward $r_t \sim R(s, a)$ and next state $s' \sim P(s'|s, a)$
    Then sample the next action $a' \sim \pi_\theta(a'|s')$
    Update the policy parameters: $\theta \leftarrow \theta + \alpha_\theta Q_w(s, a)\nabla_\theta \log \pi_\theta(a|s)$; Compute the correction (TD error) for action-value at time t:
        $\delta_t = r_t + \gamma Q_w(s', a') - Q_w(s, a)$
    and use it to update the parameters of Q function:
        $w \leftarrow w + \alpha_w \delta_t \nabla_w Q_w(s, a)$
    Move to a $\leftarrow a'$ and s $\leftarrow s'$
end for

Figure: Q Actor Critic Algorithm

Technical Outline

The dataset chosen is the S&P 500 (Standard & Poor's). It is one of the commonly followed equity (capitalization-weighted) indices as it measures the performance of the listed top 500 companies in the Unites States.

The high-level steps of operation are as follow,

1. Preprocess - To format the data as per requirement

2. Actor & critic - To define the policy model & to compute approximation respectively

3. Agent - To train an agent to perform RL

References

1. Reinforcement Learning - An Introduction by Richard S. Sutton & Andrew G. Barto [Chapter]

2. Berkley's Deep RL Course - Actor-critic Algorithms

3. Understanding Actor-critic methods & A2C [Medium]