ML ASSIGNMENT 3

Task:-

The task here is to implement Q learning algorithm in the stock trading environment. The main goal here is to train our Q learning agent to maximize the profit , that is to increase the revenue from what we had initially. Here I have implemented Q learning algorithm from scratch and trained it on stock trading environment to give maximum output.

Dataset:-

Here I have used stock market dataset of NVIDIA company. The dataset is of last 5 years. The dataset has 1258 entries. The features include information such as the price at which the stock opened, the intraday high and low, the price at which the stock closed, the adjusted closing price and the volume of shares traded for the day.

ENVIRONMENT:-

Environment is the Agent's world in which it lives and interacts. The agent can interact with the environment by performing some action but cannot influence the rules or dynamics of the environment by those actions. When an agent performs an action in the environment, the environment returns a new state of the environment making the agent move to this new state. The environment also sends a Reward to the agent which is a scalar value that acts as feedback for the agent whether its action was good or bad.

This environment is based on stock trading. It has three basic functions that are reset, step and render .

reset: Resets the environment and returns a random initial state.

step(action): Step the environment by one timestep. This method implements what happens when the agent takes the action to Buy/Sell/Hold.. Returns

observation: Observations of the environment. Returns: observation Integer in the range of 0 to 3 representing the four possible observations that

the agent can receive

reward: If your action was beneficial or not

done: Indicates if we have successfully picked up and dropped off a passenger, also called one episode

info: Additional info such as performance and latency for debugging purposes

render: Renders one frame of the environment (helpful in visualizing the environment). Here it plots total money earned or lost by the agent.

Here our environment was based on Stock market trading. In which agent can perform activities like selling, holding and buying stocks.

Agent:-

Here the agent is the persons who manipulates the stocks in the environment. He can perform any action in the environment but that won't change the environment. Here our agent can sell, hold and buy stocks. I have implemented agent from Q learning algorithm. Here we are training the agent in the environment on the given stocks data of NVDIA company. The agent interacts with the environment and obtains the use full parameters for stocks trading.

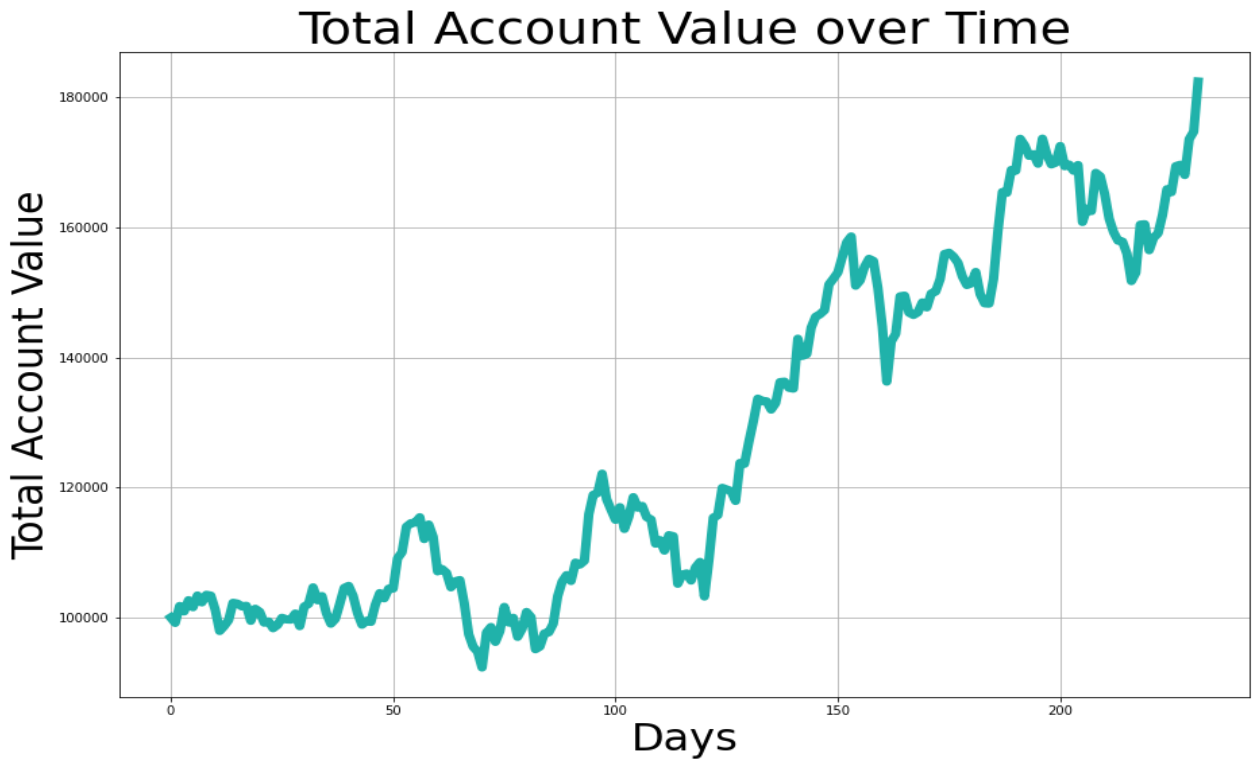
Here first we are training the Q learning in the environment. Firstly we are set the current state by giving it the value from the reset variable from the environment. Then a random action is chosen based on the value of the epsilon. That action is then given to the step function in the environment then the step function returns observation, rewards and done parameters. Then after that we update the Q table based on the rewards that we get and the action taken.

I have updated Q table with the formula-

Qn(s,a)=(1-alpha) Qn-1(s,a)+ alpha[r + gamma max(Qn-1(s',a'))]

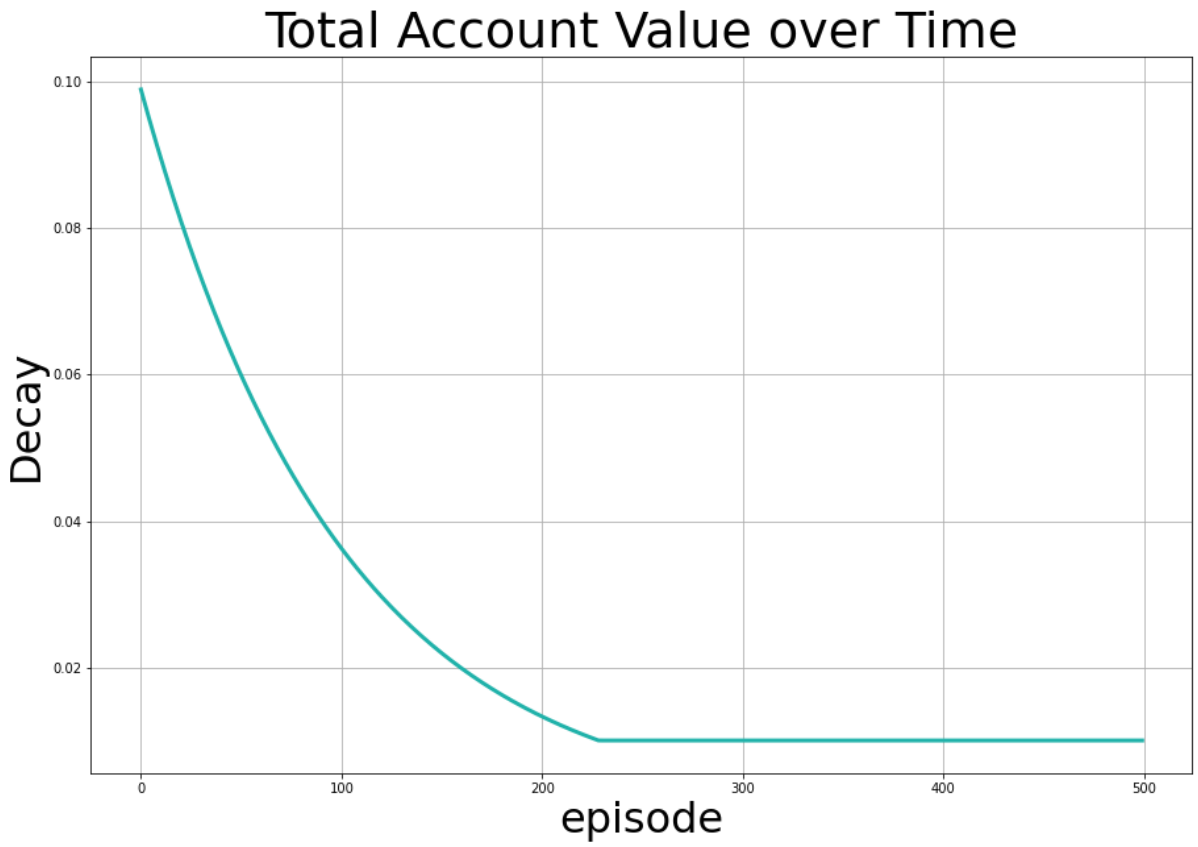
Then reward per episode is calculated . Then epsilon decay is calculated which decides the action of next state.

Here we have also implemented the evaluate function. It shows us the performance of the agent, how it has performed in the given environment. It tells us how much money our agent has collected or lost by selling, buying and holding the stocks. It plots the graph how much the agent has earned.



The the plot function shows us the value of how much there is epsilon decay and rewards that are received by the agent per episode.

Plot of epsilon decay-



Plot of Rewards per episode-