Stock market Prediction using Reinforcement Learning Technique

Shymalagowri S[1], Hema Priya N[1], Rathika PD[2]

1 Department of Information Technology, PSG College of Technology, Coimbatore, India

2 Department of Robotics and Automation Engineering, PSG College of Technology, Coimbatore, India

Abstract

Several recent studies have attempted to construct efficient mechanical trading systems utilizing machine learning for stock price prediction and portfolio management. However, these systems have a drawback in that they are primarily focused on supervised learning, which is insufficient for situations involving long-term goals and delayed rewards.Financial markets, which include the stock market, are vital for the growth of capitalism since they are marketplaces where trading takes place. The stock market is essentially the ownership of a portion of publicly traded corporations in the form of shares, which may be sold to claim ownership. To accurately estimate a company's stock price, one must devote time to it in order to make the optimal trade. Our results show that the Double Q-Learning Network performs better than the Deep Q-Learning Network and Dueling Double Q-Learning Network by receiving the highest reward in each state where the agent tries to make the best decision, i.e., buying when the market is moving upwards, selling when the market is moving downwards, and holding back when the market is moving unstable.

Keywords - Stock price prediction, Reinforcement learning, Q learning, Deep Q-learning Network, Double Q-Learning Network, Dueling Q-Learning.

I. INTRODUCTION

In the present economy, the financial sector has grown crucial. It is a market where various commodities are exchanged or sold at a set price, albeit commodity values in the stock market are not always consistent. They are influenced by external factors like politics, natural disasters, investor attitude, and exchange rates, among others. The stock market has become increasingly popular in recent years, with many people attempting to invest in it. Any country's monetary advancement is dependent on the financial market. Nigeria's and other non-industrial countries' development will be heavily reliant on stock and stock market performance in the future. If the stock market rises, the country's economy will increase at a rapid pace. The stock market is a part of the venture securities exchange that keeps track of assets for monetary purposes. It preserves reserve capital and provides liquidity to enterprises, reduces speculation risks, provides transparency for speculators, and promotes business. A long-term investment is essential to achieve improved economic development. The stock market provides long-term money to important sectors of the economy, such as corporations and government. A financial market forecast is a method of attempting to predict a stock's future value. Individuals' thoughts and emotions can be freely expressed through web-based media. Assumption examination is strongly associated with web-based media investigation because it may be used to remove feelings and sentiments from messages such as tweets. Despite the fact that many study publications have been published online over a long period of time, looking for an effective model to predict the costs of financial markets is still an active research topic today. One of the most difficult errands in financial markets forecasting is stock value expectation. Given the rapid development of the AI community, AI techniques have become the most widely chosen ways among cutting-edge processes in recent years.

Reinforcement learning is a sort of Machine Learning that involves considering behaviours in order to enhance the reward in a particular situation. It is used to determine the most appropriate behavior or approach that the agent should take at any given time. In reinforcement Learning there is no actual response to decisions, which distinguishes learning from both supervised and unsupervised learning systems. Instead, the agent chooses what to do in order to complete a task without having to prepare a dataset; and the reinforcement-learning agent can learn from experience. This project uses a Reinforce Learning Technique to predict the price of Google stock. Algorithms like Q-learning, Double Q-learning, and a Deep Neural Network are employed in this project.

II. LITERATURE REVIEW

Investors are much bothered about the dynamic variations in stock price and since then so many new and unique methodologies are adopted. Most of the recent methodologies aim at foreseeing the outcomes and make the model learn, through various Machine Learning techniques. A few of them are identified and analyzed here in this article.

Prasanthan et.al. [13] suggested a predictive analytics methodology for all purchasing information can be automatically saved and to highlight important reviews about the business, intelligent business reports are generated for future business. [14] The simplest reinforcement learning algorithm i.e. the Q-Learning algorithm is considered to be one of the most suitable algorithms. [15] analyzed CNN based Deep Q learning to train an agent based on fused images of stock data and Google trends. Deep Reinforement Learning (DRL) is deployed by [16] Zhang et.al., tested on dynamic datasets, proved the methodology successfully.

III. PROPOSED METHODOLOGY

In order to predict Google Stock Market Data, the suggested system employs a Reinforcement Learning Technique. This is an agent-based learning system in which the agent acts in an environment with the purpose of maximising the record. Reinforcement learning does not utilize labelled data. With little historical data, reinforcement learning performs better. It employs the value function and calculates it in accordance with the action's policy. Reinforcement learning may be applied to stock price prediction since it follows the same principles of requiring less historical data and functioning in an agent-based system to predict higher returns depending on the present environment.

The Reinforcement Learning Agent interacts with the environment, which is made up of a set of state, reward, and actions, to determine which choice to make. Every position in the trading environment is represented by a state, which also describes the current condition of the environment. At each state, the action is the choice or decision made by the agent. The agent's options are to stay, buy, or sell. The agent's efficiency is determined by its ability to make good decisions. While making a buy or sell decision could result in a loss in an unstable state (condition), the agent may opt to stay (the agent is either buying or selling). In an upward trend, the agent can opt to buy, and in a downward trend, sell. Loss or profit is the agent's reward.

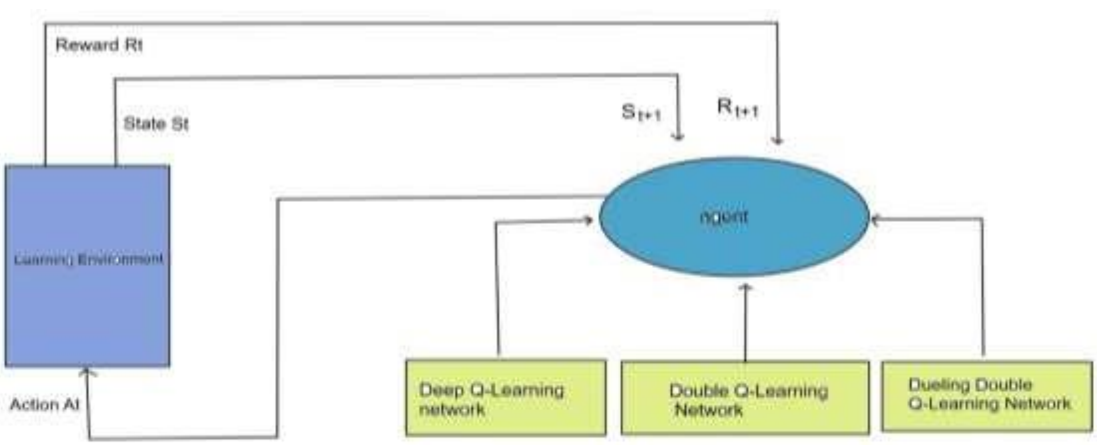

Figure 1 - Architecture Diagram of the Proposed System

The three Reinforcement Learning Algorithms that will be used to train the Reinforcement Learning Agent are Deep Q-Learning, Double Q-Learning, and Dueling Double Q-Learning as shown in Figure 1. In order to make the best decisions, the agent learns and interacts with the learning environment. The agent's creation begins with the declaration of some variables, each with a set of parameters. Episode count, memory size, batch size, epsilon, epsilon start, epsilon end, epsilon_decay, gamma, total rewards, and total profits are the variables used.All of them are threshold constants that govern how the financial market buys, holds, and sells.

A. Deep Q-Learning Network:

Q-Learning is a value-built Reinforcement Learning Algorithm that uses the Q Function to discover the optimum action/choices. This aids the Reinforcement Learning agent in performing the appropriate actions or making the appropriate choices in a given state. To estimate the Q-Values Function, Deep Q Learning employs a neural network. The State will be taken as an input and the Q-Value of all feasible actions/choices will be brought about the output.

$$Q(s, a) = r(s, a) + \gamma \max_{a} Q(s', a) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \ldots\ldots\ldots\ldots (1)$$

$$Q(s, a) \rightarrow \gamma Q(s', a) + \gamma^2 Q(s'', a) \ldots \ldots \ldots \gamma^n Q(s''^{\ldots n}, a) \ldots\ldots\ldots\ldots (2)$$

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_{a} Q(S_{t+1}, a) - Q(S_t, A_t)] \ldots\ldots (3)$$

Figure 2 - Deep Q-Learning Network Formula

Where,
Q - Q learning factor
s and **a** -actions carried out by the agent on a particular state
γ **-** gamma
a - the agent's rate of learning in the environment.
t - time taken by the agent in completing one action in a state

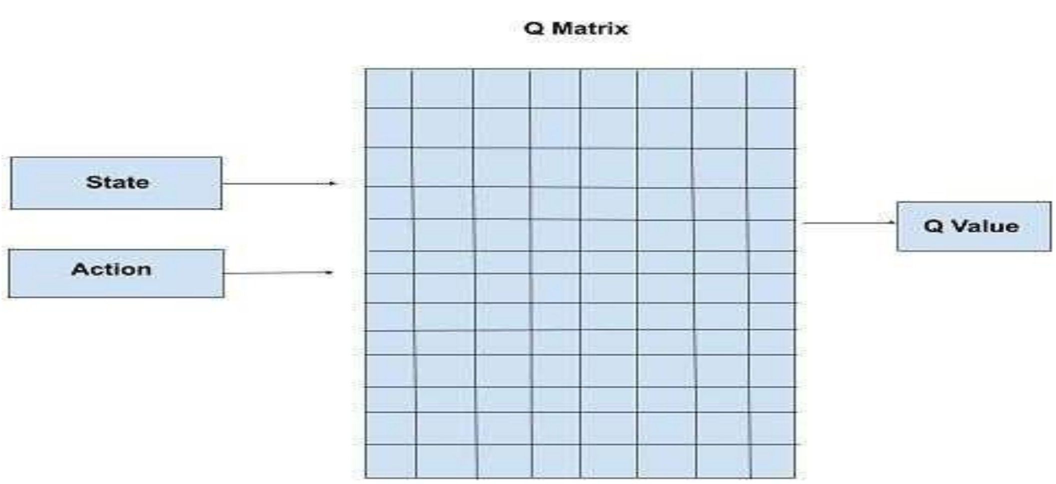

Figure 3 - Q Learning

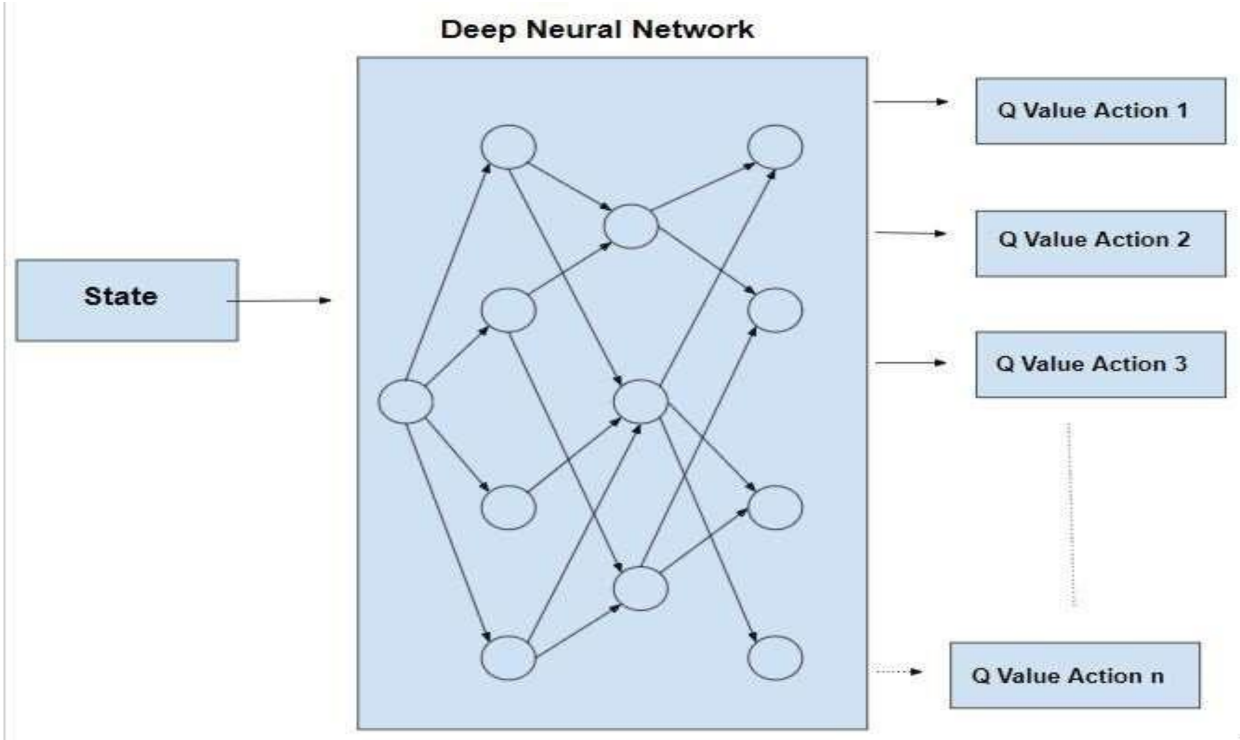

Figure 4 - Deep Q-Learning Architecture

B. Double Q-Learning Network:

The employment of two different Q-Value estimators (Q^A and Q^B) to update each other is referredto as a double Q-Learning Network. Making use of these individual estimators, we can unprejudiced Q-Value approximate the actions/choices selected using the contrasting estimator. Consequently, maximization bias can be kept away from by ignoring updates from prejudice estimators. Mathematically, this expression can be stated as

$$\text{Max}_a\, Q_t\,(s_{t+1},\, a)\ldots6$$

$$E\{\text{Max}_a\, Qt(s_{t+1},a)\}\ldots.7$$

$$\text{Max}_a E\{Q_t(s_{t+1},a)\}\ldots8$$

$$Q^A\,(s',\, a*\,) = \text{Max}_a\, Q^A(s',a)\ldots9$$

$$Q^B(s',a*\,) = \text{Max}_a\, Q^B(s',a)\ldots10$$

Figure 5 - Double Q-learning Formula

Equation 6 is an approximation for equation 7, which in turn estimates equation 8. Hence condition one is a fair-minded example drawn from 7. Instead of using equation 9 to update Q^A, as traditionally done in Q- Learning Network. Q^B which is equation 5 will be used to update Q^A.

C. Dueling Double Q-Learning Network:

Dueling Double Architecture is made up of two streams that constitute the value functions, while sharing a customary convolutional feature learning component. The both streams are incorporated through a distinct combined layer to yield an approximation of the state-action value function. Mathematical, Dueling Double Q-Learning Network can be expressed as:

V(s) – the value of being in the states..11
A(s,a) – the advantage of taking actions a in the states 12
$Q(s, a^*) = V(s)$ – under deterministic policy..13
$Q(s, a^*) = V(s) + A(s, a)$ – under optimum policy ...,. 14

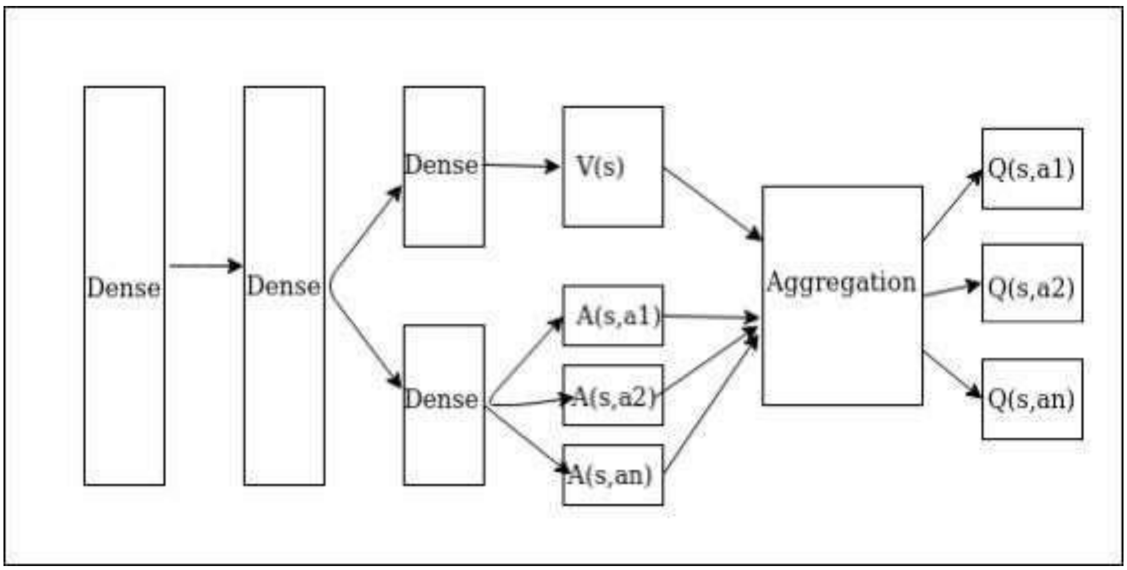

Figure 6 - Dueling Double Q-Learning Architecture

DATASET DESCRIPTION

Dataset : Google Stock

Description:

The dataset is collected from kaggle. It contains 7 columns such as Date, Open, High, Low, Close, Adj Close, Volume. It provides stock data from 2004 through 2020 organised by date (September). The features of the dataset provides details such as
Open - Opening price of the stock

High - Maximum price of the stock for the dayLow - Minimum price of the stock for the dayClose - Closing price of stock for the day

Adj Close - Data is adjusted using appropriate split and dividend multipliers for the closing price for the day.

Volume - Volume is the physical number of shares traded of that stock on a particular day.

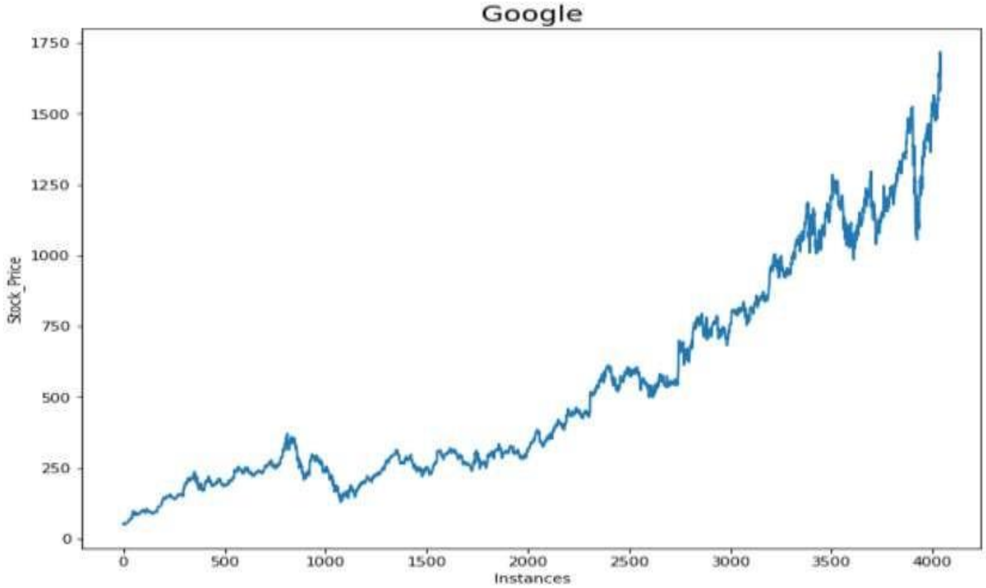

Figure 7 - Google Stock Price

VI. RESULTS

A. Deep Q-learning Network:

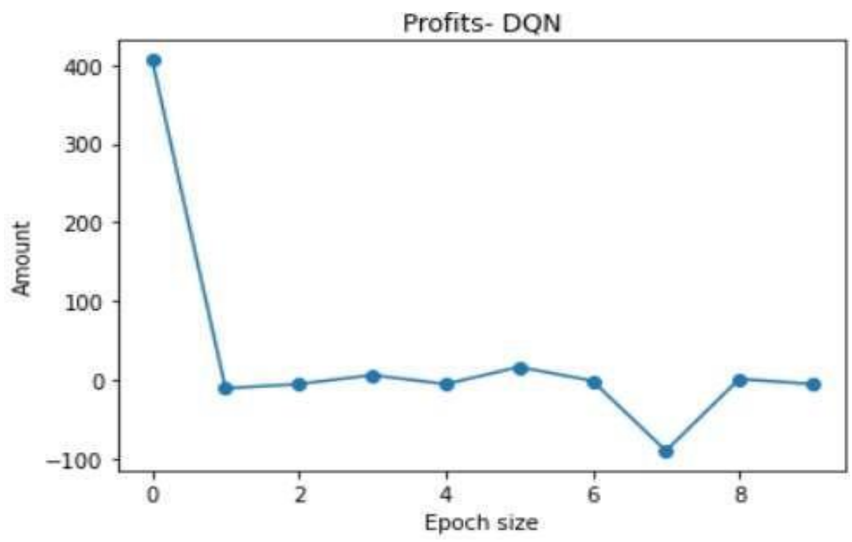

Figure 8 - Profits of Deep Q-Learning Network

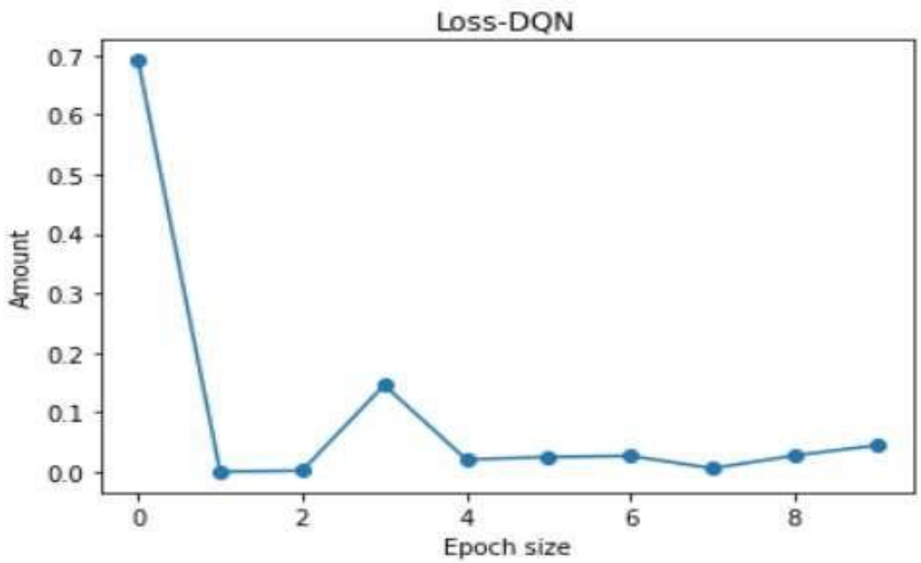

Figure 9 - Loss of Deep Q-Learning Network

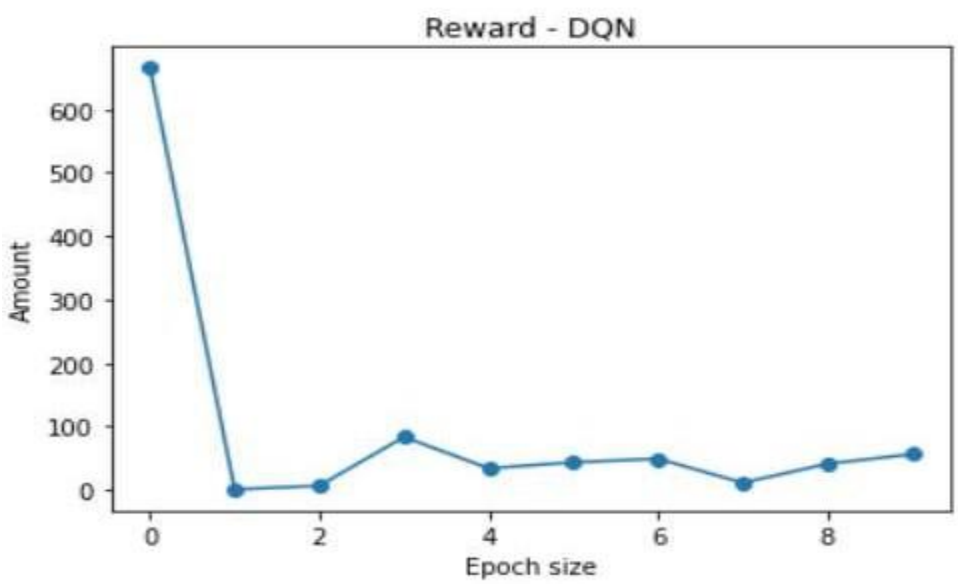

Figure 10 - Rewards of Deep Q-Learning Network

Inference:

In figure(8,9,10) the x-axis is taken as the size of epoch and y axis is taken as the amount.

Figure 8 shows the profits of Deep Q-learning network.The profit gets reduced drastically from the 0th epoch and only a slight variation was seen till 6th epoch.It got reduced in 7th epoch but raised after 8th epoch and the profit is maintained after 8th epoch with a little fluctuation.

Figure 9 shows the graphical representation of loss of Deep Q-Learning network.The loss gets increased in epoch 3 but then it was decreased and low loss is maintained till the 10th epoch with less fluctuation.

Figure 10 shows the plotting of rewards of the Deep Q-Learning network.A reward in Reinforcement Learning provides feedback to the machine. Whenever the machine makes a correct decision reward is provided.In the proposed model ,for every episode reward is taken as the sum of profits of each epoch,if profit is achieved else reward is taken as zero if loss is achieved.As profit increases reward increases.As profit gets reduced in epoch 7,reward also gets reduced in epoch 7.Similar to profit,reward gets increased in epoch 8 and it is maintained the same for all the further epochs.

The DQN model has the best profit of 16.80 and reward of 43.73.

B. Double Deep Q-learning Network:

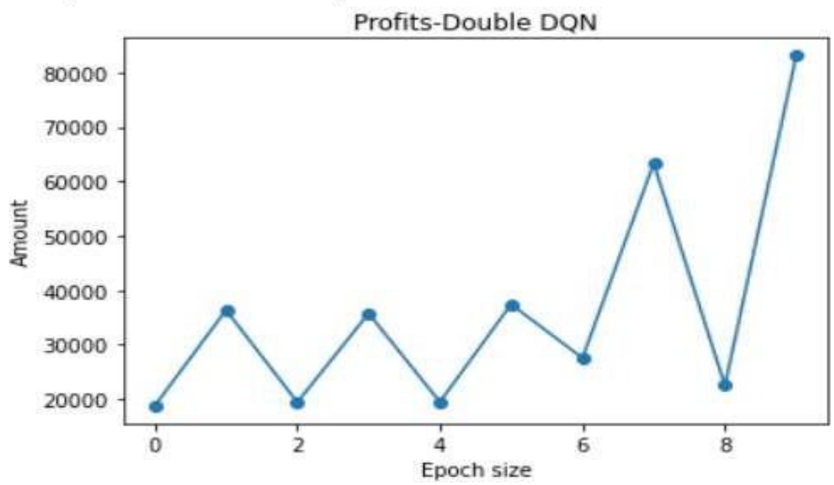

Figure 11 - Profits of Double Deep Q-Learning Network

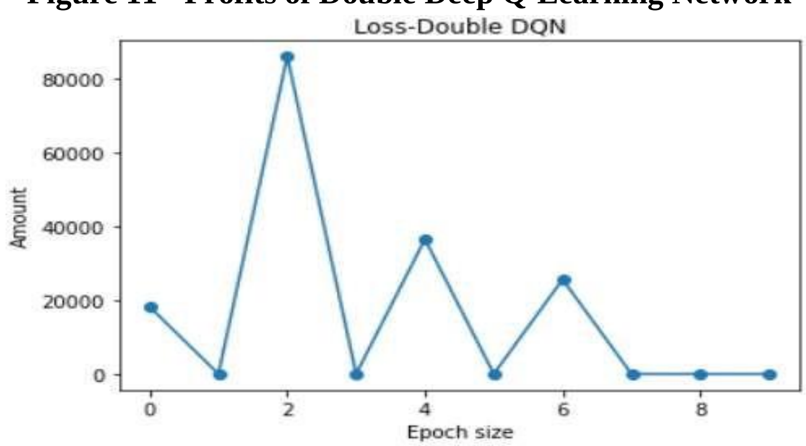

Figure 12 - Loss of Double Deep Q-Learning Network

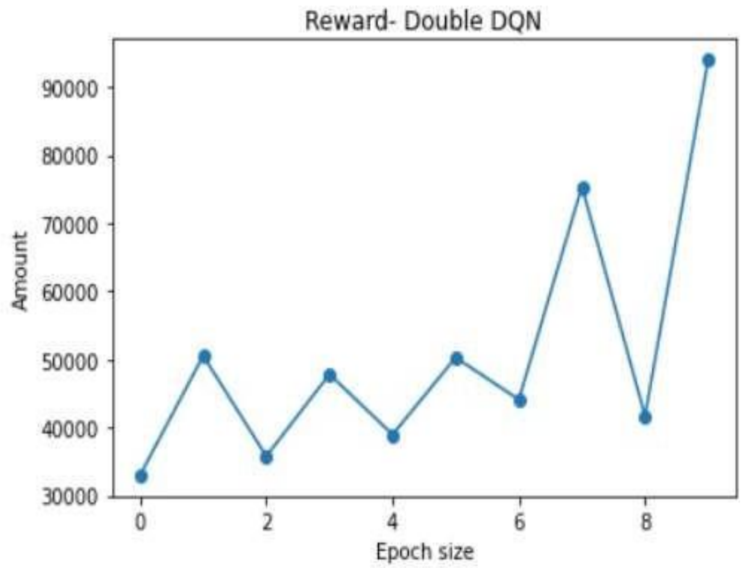

Figure 13 - Rewards of the Double Deep Q-Learning Network

Inference:

In figure(11,12,13)-the x-axis is taken as epoch size and y-axis is taken as amount.

In figure 11,one can clearly see that profit is increasing from 0 gradually to reach the highest value.It is clear that the profit is at its peak at the 10th epoch.It has a value of 83165.05 at 10th epoch.

Figure 12 shows the graphical representation of loss for Double DQN.At epoch 2 loss reaches its maximum value but it decreases after each epoch and saturates to almost 0 from epoch 8.

Figure 13 shows the rewards for the Double DQN model.Reward gradually increases from 0 at epoch 0 and fluctuates at each epochs by dropping down and increasing again.Finally, as profit was at its peak at epoch 10 reward also reaches its peak at epoch 10.It has the best value of 94023.80 at epoch 10.

C. Dueling Double Q-learning Network :

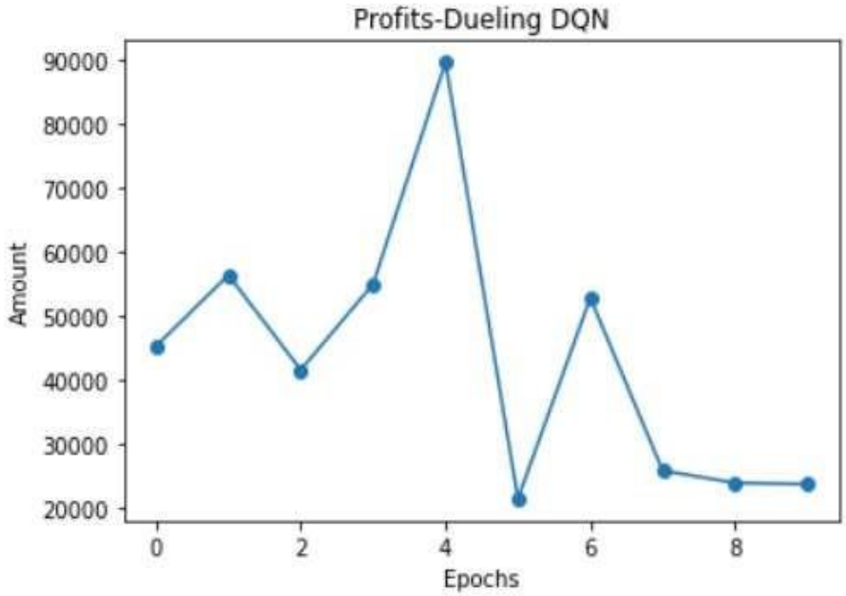

Figure 14 - Profits of Dueling Double Q-Learning Network

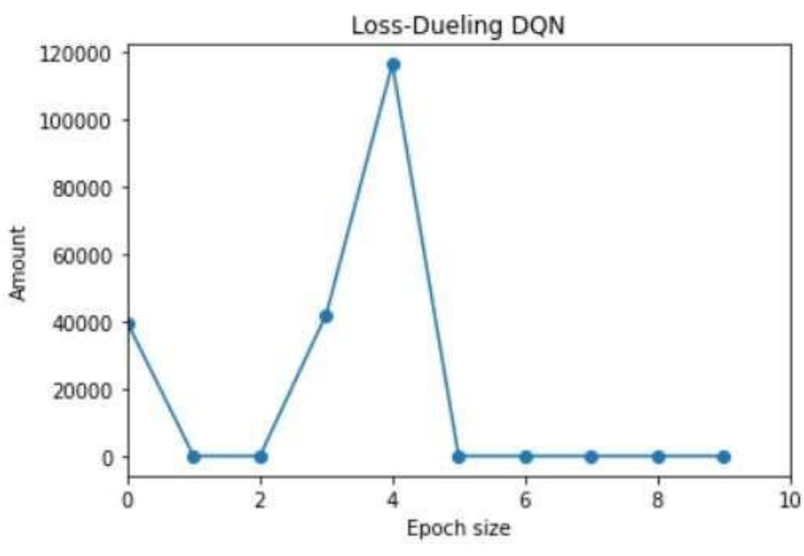

Figure 15 - Loss of Dueling Double Q-Learning Network

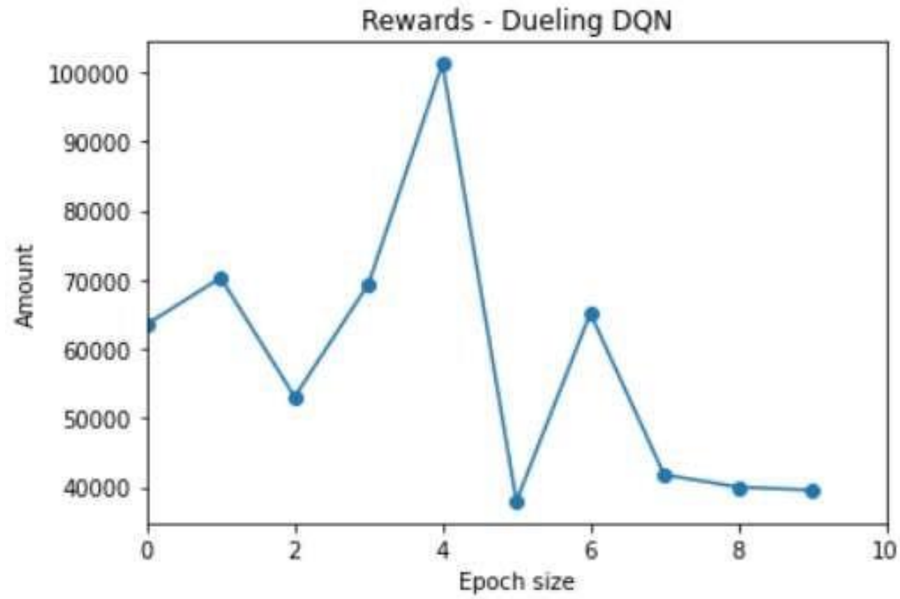

Figure 16 - Reward of Dueling Double Q-Learning Network

Inference:

In figure (14,15,16) - the x axis is taken as the number of epochs and y axis is taken as Amount.

Figure 14 shows the fluctuation in the profit of Dueling Double Q-learning Network. The graph shows a steep rise in profit at epoch 4 and there is a sudden fall after it. Also, the profit is very low at epoch 5. After epoch 5, there is a slight rise and fall in profits and then the profit is maintained between the amount of 20000 and 30000.

Figure 15 shows the graphical representation of the loss of Dueling Double Q-learning Network. Initially there is a slight fall in the losses and after some epochs the losses are increasing. There is a steep downward trend followed by a huge amount of losses at epoch 4. After epoch 5, the losses are maintained as zero with no fluctuations. Thus, it infers that only profit is gained after epoch 5 with no loss incurred.

Figure 16 shows the reward obtained by Dueling Double Q;learning Network. Whenever there is a profit incurred in figure 14, there is a reward obtained for the stock. As profit rises, reward increases. When there is a drop down in profit, reward also decreases. There is a maximum amount of reward (101202.72) at epoch 4 followed by a steep down trend and reward increases at epoch 6. Then, the reward is maintained lower after epoch 7.
Dueling Double Q-learning shows a higher profit of 41564.62 and reward of about 53076.22 with no loss incurred.

D. Tabulation:

MODELS	PROPOSED METHODOLOGY			BASE PAPER		
	PROFIT	LOSS	REWARD	PROFIT	LOSS	REWARD
DQN	16.80	0.024	43.73	8.92	0.013	24.54
DOUBLE DQN	83165.05	0.094	94023.80	66876.22	0.131	79995.61
DUELING DQN	41564.62	0.0	53076.22	33293.38	0.0	47045.31

Table 1 - Comparison of the metrics between base paper methodology and the proposed methodology for all three models.

Inference:

Table 1 shows the comparison of the amount of profit, loss and reward for the proposed methodology and the base paper. It also shows that the values of profit and reward in proposed methodology are higher than the base paper's due to the introduction of new threshold functions and modification of hyper parameters. Double DQN shows the best amount of profits with 83165.05and loss of 0.094 in proposed methodology whereas in base paper, the profits and losses are of 66876.22 and 0.131.

VII. CONCLUSION AND FUTURE ENHANCEMENTS

In a country's economy, the stock market is quite essential. The stock market is difficult to predict because of the constant fluctuation in stock values, which is dependent on various elements that build complicated patterns. In order to identify the trend in the stock market, reinforcement learning is used to predict the Google stock price prediction. The agent was trained using three reinforcement learning techniques such as Deep Q learning network, Double Deep Q-learning Network, Dueling Double Q-learning Network. The experiment results shows that Double Q-learning performs well when compared to the other Q-learning techniques implemented. There is a higher amount of profit (83165.05) and reward (94023.80) predicted by Double Q-learning technique. For future work, hybrid reinforcement learning models could be developed with new financial parameters and observe the performance of the model. It can further be improved by modifying the threshold functions introduced in this model so as to obtain better results.

REFERENCES

[1] P. S. Ezekiel, O.E. Taylor, C.G. Igiri "Google Stock Market Price Prediction using Reinforcement Learning Technique", 2021 Journal of Emerging Technologies and Innovative Research (JETIR), Vol. 8, No. 1, pp. 2349 - 5162, January 2021.

[2] Vijh, Mehar, Deeksha Chandola, Vinay Anand Tikkiwal, and Arun Kumar. "Stock closing price prediction using machine learning techniques." Procedia Computer Science Vol.167,pp. 599-606, 2020.

[3] Hu, Zexin, Yiqi Zhao, and Matloob Khushi. "A survey of forex and stock price predictionusing deep learning." Applied System Innovation, Vol. 4, No.1, pp. 9, 2021.

[4] S. Liu, F. Zhou "On Stock Prediction Based on KNN-ANN Algorithm", IEE fifth International Conference on Bio-Inspired Computing: Theories and Applications (BICTA), pp.310-312, 2010.

[5] Sharma, D. Bhuriya, U. Singh "Survey of Stock Market Prediction Using Machine Learning Approach", International Conference of Electronics, Communication and Aerospace Technology, pp.506-509, 2017.

[6] Henrique, B.M., Sobreiro, V.A., Kimura, H. "Literature review: Machine learning techniques applied to financial market prediction". Expert Syst. Appl. 124, 226–251, 2019.

[7] X. Wu, H. Chen, J. Wang, L. Troiano, V. Loia, H. Fujita "Adaptive stock trading strategieswith deep reinforcement learning methods" Journal of Information Sciences, Vol.538, pp.142-158,2020.

[8] T. L. Meng, M. Khush "Reinforcement Learning in Financial Markets" Multidisciplinary Digital Publishing Institute (MDPI): Data, pp.1-17, 2019.

[9] R. Sathya, P. Kulkarni, M. N. Khalil, S. C. Nigam "Stock Price Prediction using Reinforcement Learning and Feature Extraction" International Journal of Recent Technology and Engineering (IJRTE), vol.8, issue.6, pp.3324-3327, 2020.

[10] H. Chung, K. shin "Genetic Algorithm-Optimized Long Short-Term Memory Network forStock Market Prediction" Multidisciplinary Digital Publishing Institute (MDPI): Sustainability, pp.1-17, 2018.

[11] Lee, Jae Won. "Stock price prediction using reinforcement learning." ISIE 2001. 2001 IEEE International Symposium on Industrial Electronics Proceedings (Cat. No.01TH8570), Vol. 1, pp. 690-695, 2001

[12] Khanna A., Gupta D., Bhattacharyya S., Hassanien A.E., Anand S., Jaiswal A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1388., September 2021

[13] Prasanthan, S. T., D. Dharani, & N. Hemapriya. "A Systematic Review on Predictive Analytics for Stock Management System." Software Engineering and Technology [Online], 12.5 (2020): 85-91. Web. 23 Jul. 2022

[14] Sayak Paul, "An introduction to Q-Learning: Reinforcement Learning", May 15, 2019.

[15] Dasgupta, Agnibh. *Deep Q Learning Applied to Stock Trading*. Diss. Utah State University, 2020.

[16] Zhang, Junhao, and Yifei Lei. "Deep Reinforcement Learning for Stock Prediction." *Scientific Programming* 2022 (2022).