Google Stock Market Price Prediction using Reinforcement Learning Technique

P. S. Ezekiel[1], O.E. Taylor[2] & C.G. Igiri[3]

[1,2,3]Department of Computer Science, Rivers State University, Port Harcourt, Nigeria.

Abstract- The financial market has become an essential market in today's economy. It is a market where different commodities are being exchange or sold at a certain price, though prices of commodities in the stock market are not stable. They are being influenced by some external factors like, politics, natural calamities, investor's sentiment, exchange rate e.t.c.This system propose an agent using Reinforcement Learning Technique to predict google stock market data. The agent was trained using three Reinforcement Learning Algorithms. The algorithms used in this system are Deep Q-Learning Network, Double Q-Learning Network and Dueling Double Q-Learning Network. We first create an environment of which the agent can learn from. The environment consist of state, action and reward. The state is the representation of the environment; the action is the movement of the agent within the state or the choice/decision made by the agent at a given state. The action taken by the agent are stay, buy or sell. Our experimental results shows that Double Q-Learning Network outperforms the Deep Q-Learning Network and Dueling Double Q-Learning Network by receiving the highest reward in each state that the agent tries to take the best decision that is to say, buying the movement of the market is an upward trend, selling when the movement of the market is a downward trend and holding back when the movement of the market is unstable. The Dueling Q-Learning Network had the best reward of about 29.4 on 35 epoch, Double Q-Learning had reward of about 35.4 on 45 epoch, and the Deep Q-Learning Network had reward of about -81.8 on the first epoch.

Keywords: Reinforcement Learning, Deep Q-learning Network, Double Q-Learning Network, Dueling Q-Learning, Google Stock Market data

1.0 Introduction

The financial market has become an essential market in today's economy. It is a market where different commodities are being exchange or sold at a certain price, though prices of commodities in the stock market are not stable. They are being influenced by some external factors like, politics, natural calamities, investor's sentiment, exchange rate e.t.c. The commodities that can be exchange in the stock market are exchange of currency pairs (EUR/USD, USD/JPY, AUD/USD, e.t.c), gold, silver, crypto currency and oil. The stock market has become more essential in recent times with many persons trying to invest in the market. Financial Market has a critical function in the monetary advancement of any Country. Henceforth Nigeria and other non-industrial nation's development profoundly relies upon Stock and the Stock Market Performance. The rise in the stock market, then the growth of the country's economy would be high. Stock market is an establishment in the venture securities exchange that monitors assets for monetary affairs. It retains reserve funds and gives liquidity to ventures, diminishes speculation hazards, offers straightforwardness for speculations and supports business. For a better economic development, a long=term investment is required to be committed. The stock trade gives long-term capital to significant areas of the economy including organizations and the legislature. Stock trade indexes are regularly utilized as a pointer of financial mend[1]. Financial market forecast is the way toward attempting to decide the future worth of any stock. Web-based media offers a vigorous outlet for individual's considerations and emotions. Investigation of web-based media is emphatically identified with assumption examination as this can be utilized to remove feelings and sentiments from messages like tweets. Data mining systems like Natural Language Processing, Random Forest and Neural Network is utilized for investigating interpersonal organization content [2]. For a long time, conventional measurable forecast strategies.

Looking for an effective model to anticipate the costs of the financial markets is a functioning exploration topic today [3] notwithstanding the way that many research articles have been published online over a long period of time.Amidst financial markets forecast, stock value expectation is considered as one of the most difficult errands [4]. Among the cutting edge procedures, AI strategies are the most broadly picked methods in recent years, given the fast improvement of the AI community. The other reason is that the customary factual learning calculations cannot adapt to the non-fixed and non-linearity of the securities exchanges [5].Reinforcement learning is a Machine Learning procedure, which depends on thinking about actions for amplifying the reward in a specific circumstance. It is utilized to locate themost ideal conduct or way the agent must consider at any

time. Reinforcement Learning is not quite the same as the supervised learning and Un supervised learning methods as there is no answer-based decision making here, agent chooses what to do in other to complete a given task without preparing a dataset, the reinforcement-learning agent can learn from experience. This paper presents a Reinforce Learning Technique in predicting google stock market price. The algorithms used in this paper are Q-learning, Double Q-learning and a Deep Neural Network algorithm.

2.　　　Related Work

Adaptive stock trading strategies with deep reinforcement learning methods [6] proposed an adaptive stock trading strategy using reinforcement learning technique. They made use of Gated Recurrent Unit (GRU) on time-series nature of financial market data to get useful financial features, which can speak to the inherent qualities of the financial market for versatile trading choices. Moreover, with the customized plan of action and spaces, they proposed two trading techniques, which are GDQN (Gated Deep Q-getting trading pattern), and GDPG (Gated Deterministic Policy Gradient trading pattern). For verification of the robustness and adequacy of Gated Deep Q-getting trading pattern and Gated Deterministic Policy Gradient trading pattern, an assessment was carried out on both trading techniques in the moving and in the unpredictable financial market from various nations. Trial results show that the proposed GDQN and GDPG beat the Turtle trading pattern as well as accomplish more steady returns than a state of the art direct reinforcement learning technique. Test results exhibit that the GDPG with an actor critic system is steadier than the GDQN with a critic only system in the ever-developing financial market.

Deep Long-Short Term Memory with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems [7] proposed a supervised deep learning algorithm and a deep reinforcement learning algorithm for predicting a short-trend currency pairs in the forex market in other to escalate the profit on investing in the High Frequency Trading technique. With a normal exactness of about 85%, the proposed technique can anticipate the medium-short term pattern of a currency cross dependent on the historical pattern of this and by methods for correlation of information with other money crosses utilizing strategies known in the stock field with the term exchange. The final fragment of their proposed pipeline incorporates a matrix-exchanging engine dependent on the previously mentioned forecasting pattern that will perform high recurrence activities to boost profit and limit drawdown. The exchanging framework has been approved more than most stock market trading framework used on the EUR/USD cross currency verifying the performance regarding returns of the Investment (98.23%). Notwithstanding a decreased drawdown (15.97 %) which confirms its financial supportability.

Reinforcement Learning in Financial Markets [8] carried out systematic review on recent articles on stock market forecasting using reinforcement learning techniques.All articles reviewed had some unreasonable presumptions such as no exchange costs, no liquidity issues and no offer or ask spread issues. Exchange costs had significant impacts on the profitability of the reinforcement learning technique contrasted with the tested baseline technique. In spite of indicating measurably significant profitability when using reinforcement learning in comparison with baseline models in numerous investigations, some demonstrated no important degree of profitability, specifically with huge changes in the value design between the framework preparing and testing information. Besides, hardly any presentation correlations between reinforcement learning and other refined machine/deep learning models were given. The effect of exchange costs, including the bid/ask spread on profitability has additionally been surveyed. In conclusion, reinforcement learning in stock/forex market is still in its initial turn of events and further exploration is expected to make it a solid strategy in this space.

Stock Price Prediction using Reinforcement Learning and Feature Extraction [9] used sentiment analysis from social media and reinforcement learning in forecasting the worth of the stock market. They predicted the movement of the stock market by analyzing the economical data, which has to do with the use of real time data and historical data. Q-learning algorithm was used over a comparatively detectable Markov Decision Process made up of any amount of Stocks extracted as a State and giving three actions which are Buy/Sell and Hold. Making use of the tweet analysis and social media comments concerning the stock market, a profitable data can be acquired which can be helpful in deciding the general public analysis of the stock market. Making use of these two distinctive methods, a profitable framework can be developed in forecasting the stock market more accurately.

Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning [10] carried out a survey on stock market exchange using reinforcement learning and three proposed techniques to escalate total gain and reflect real stock market conditions while overcoming the restrictions of the stock market data. Firstly, an automated exchange system was designed using a combination of a reinforcement learning technique and a deep learning algorithm connected to a deep q-network to forecast the amount of shares to exchange in the stock market. Secondly, a survey was carried out on different trading patterns using Q-values in analyzing which trading pattern is more suitable in terms of profit wise in an unstable market. Finally, a transfer learning method was used in preventing incompleteness from deficient stock market data. In verifying their proposed methods experimentally, four stock indices were used, which are S&P500, KOSPI, HIS and EuroStoxx50. An increased profit by 4 times in the S&P500, 5 times in the KOSPI, 12 times in the HSIand 6 times in the EuroStxx50 was realized. In an unstable market, keeping back the decisions to buy/sell increases the total profit by 18% in the S&P500, 24% in KOPSI, 49% in EuroStoxx50.

Adaptive stock trading with dynamic asset allocation using reinforcement learning [11] developed a new stock trading technique that integrates asset allocations dynamically using a reinforcement learning technique. Meta Policy (MP) is the assets allocation-trading pattern designed to make use of a short-term data from both stock market direction and the proportion of the stock market financed over the asset. Localized forex traders are constructed with strategy-based on numerous predictors, which they used in deciding the purchase of currency per direction. A compact environment is designed for the learning agent using Meta Policy in reinforcement learning. A testing was carried out on a Korean stock data that shows that Meta Policy surpass other fixed asset-allocation trading pattern and reduces the risk intrinsic in localized forex traders.

Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction [12] developed a hybrid framework using a genetic algorithm and a long short-term memory algorithm. In estimating the time window size and architectural components of the long short-term memory network, an experiment and misconception based on heuristics were frequently used. They partly investigates the attributes of the financial market data by proposing a systematic approach that will regulate the time window size and architecture for the long short-term memory network utilizing genetic algorithm. A Korea Stock Market Price Index (KOSPI) data was used in testing their proposed hybrid framework. Their test result shows that the hybrid model of long short-term memory network and genetic algorithm outplays the benchmark model.

An innovative neural network approach for stock market prediction [13] developed a framework using an innovative neural network algorithm in predicting the financial market. To analyze the impact of the market attributes of the stock market price, conventional neural network algorithms may incorrectly forecast the financial market, considering the inceptive weight of the arbitrary selection issue. They trained two models using long short-term memory neural network with fixed layer and a long short-term neural network algorithm with encoder to forecast the financial market. Their testing results demonstrates the long short-term memory with a fixed layer (57.2% accuracy) is better than the long short-term memory with encoder (56.9% accuracy).
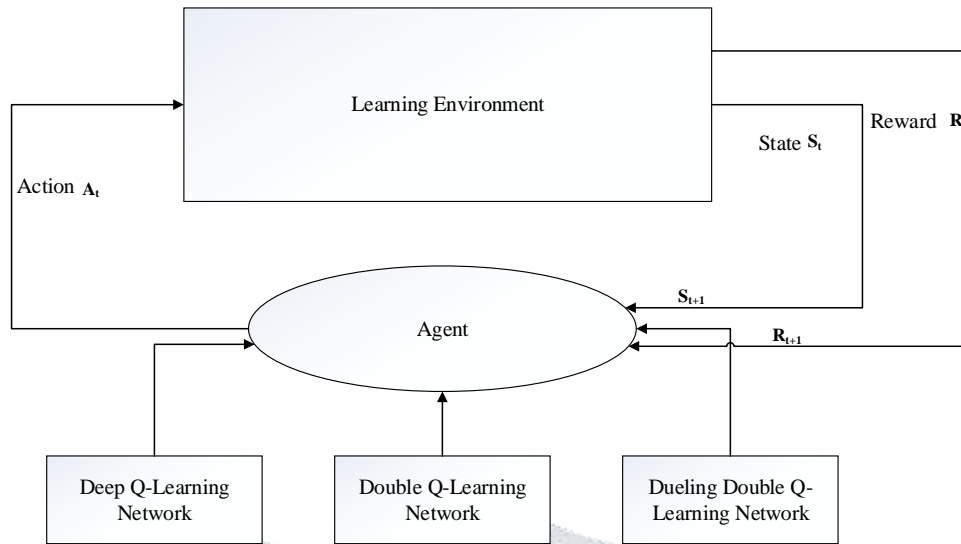
3. Methodology

Figure 1: Architecture Diagram of the Proposed System

The proposed system uses a Reinforcement Learning Technique in predicting Google Stock Market Data. The system made use of three Reinforcement Learning Algorithms in training the agent. The Algorithms used are Deep Q-Learning, Double Q-Learning and Dueling Double Q-Learning. The architectural components of the proposed system can be explained as follows:

Learning Environment: The environment is made up of a set of state, reward and actions of which the Reinforcement Learning Agent interacts with in other to decide on what choice to make. The states comprises of the every positions in the trading environment, it also describes the current condition of the environment. The action is the choice or decision taken by the agent at each state. The actions taken by the agent are stay, buy or sell. The ability of the agent to make right choices determines the efficiency of the agent. In an unstable state (condition), the agent may decides to stay (the agent is either buying or selling) because making a buy or sell choices may lead to loss. The agent may decide to buy in in an upward trend and or sell in a download trend. The reward of the agent is either loss or profits.

Agent: The Reinforcement Learning Agent will be trained using three Reinforcement Learning Algorithms, which are Deep Q-Learning, Double Q-Learning and Dueling Double Q-Learning. The agent learn and interacts with the learning environment in other to make optimal choices. The creation of the agent begins with some declaration of some variables with some parameters. The Variables used are epoch_num, max_step, memeory_size, batch_size, epsilon, epsilon_decrease, epsilon_min, start_reduce_epsilon, update_q_freq, gamma and show_log_freq, total_rewards and total profits. All this are threshold constants that enforces the whole buy, hold and sell in the financial market.

Deep Q-Learning Network: Q-Learning is a value-built Reinforcement Learning Algorithm, which is used in finding the best action/choices using Q Function. This helps the Reinforcement Learning agent in taking the right action/choices in a given state. The Deep Q Learning make use of a neural network to estimate the Q-Values Function. The State will be taken as an input and the Q-Value of all feasible actions/choices will be bring about the output.

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

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

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

Where:

Q is the Q learning factor

s and a are actions carried out by the agent on a particular state

γ is the gamma

α is the rate at which the agent learns in the environment

t is the time taken by the agent in completing one action in a state
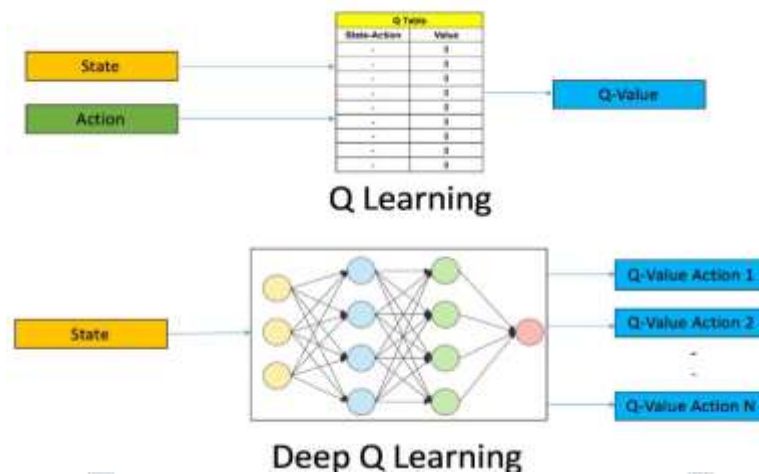

Figure 2: Deep Q-Learning Architecture

Double Q-Learning Network: Double Q-Learning Network entails using two different Q-Value estimators (Q^A and Q^B), each of which is used to update one another. Making use of this individual estimators, we can unprejudiced Q-Value approximate of the actions/choices selected using the contrasting estimator. Consequently, maximization bias can be keep away from by ignoring update from prejudice estimator. This expression can be represented mathematically as

$$Max_aQ_t(s_{t+1},a)\dots6$$

$$E\{Max_a\ Qt(s_{t+1},a)\}\dots.7$$

$$Max_aE\{Q_t(s_{t+1},a)\}\dots8$$

$$Q^A(s',a*) = Max_a\ Q^A(s',a)\dots9$$

$$Q^B(s',a*) = Max_a\ Q^B(s',a)\dots10$$

Equation 6 is an approximate for equation 7, which in turn estimates equation 8. Therefore equation one is an unprejudiced sample 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.

Algorithm of the Double Q-Learning Network

Step1. Initialize Q^A,Q^B,s

Step2: repeat

Step3: Choose a, based on $Q^A(s,\cdot)$ and $Q^B(s,\cdot)$, observe r, s′

Step4: Choose (e.g. random) either UPDATE(A) or UPDATE(B)

Step5: if UPDATE(A) then

Step6: Define a* = $argmax_aQ^A(s',a)$

Step7: $Q^A(s,a) \leftarrow Q^A(s,a) + \alpha(s,a)\ (r + \gamma Q^B(s',a*) - Q^A(s,a))$

Step8: else if UPDATE(B) then

Step9: Define b* = $argmax_aQB(s',a)$

Step10: $Q^B(s,a) \leftarrow Q^B(s,a) + \alpha(s,a)(r + \gamma Q^A(s',b*) - Q^B(s,a))$

Step11: end if
Step12: s ← s′
Step13: until end

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

$$V(s) - \text{the value of being in the state s } \dots 11$$

$$A(s,a) - \text{the advantage of taking actions a in the state s } \dots 12$$

$$Q(s, a^*) = V(s) \dots 13 - \text{under deterministic policy}$$

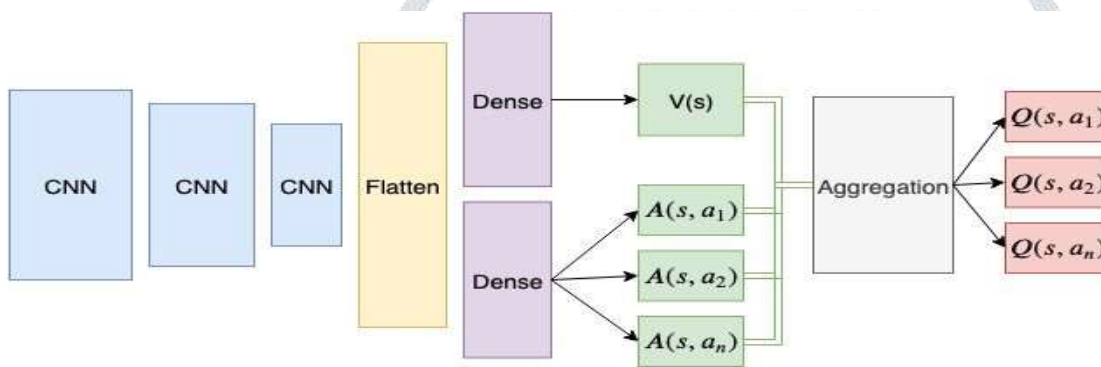$$Q(s, a^*) = V(s) + A(s, a) \dots 14 - \text{under optimum policy}$$

Figure 3: Architecture Diagram of the Dueling Double Q-Learning Network

4. Result and Discussion

This system propose an agent using Reinforcement Learning Technique to predict google stock market data. The agent was trained using three Reinforcement Learning Algorithms. The algorithms used in this system are Deep Q-Learning Network, Double Q-Learning Network and Dueling Double Q-Learning Network. We first create an environment of which the agent can learn from. The environment consist of state, action and reward. The state is the representation of the environment; the action is the movement of the agent within the state or the choice/decision made by the agent at a given state. The action taken by the agent are stay, buy or sell. The reward is the outcome of the agent after a successful action. For every decision taken by the agent, a reward is being given. The rewards increase from minus (-) sign to plus (+) sign when it takes the best decision. The following are some variables and parameters in the learning environment. They are profits, action, rewards, time and history. We initialized reward=0, time=0 and we said if action==0, the agent should hold on, action==1, the agent should buy and action==2 the agent should sell. After performing either of this action, a reward is being given to the agent in other to tell it's performance. After receiving the reward, the agent starts again by moving to the next state in other to take a better action and to receive a better reward. After creating our learning environment, we trained our agent using three Reinforcement Learning algorithms (Deep Q-Learning Network, Double Q-Learning Network and Dueling Double Q-Learning Network) in other to have a better optimal result to help the agent make a better choice at each given state. For each of the Q- Learning Algorithm, we have the following variables and parameters. epoch_num=50, memory_size=200, batch_size=20 for Deep Q-Learning Network, batch_size=50 for both Double Q-Learning Network and Dueling Double Q-Learning Network. epsilion=1.0, epsilon_decrease = 1e-3, epsilon_min = 0.1, start_reduce_epsilon = 200, train_freq = 10, update_q_freq = 20, gamma = 0.97, show_log_freq = 5. All this are threshold that enforces the the agent in making the right decision like knowing when to hold, buy or sell. Figure 5 and 6 shows the total

reward and the total loss made by the Deep Q-Learning Agent. Figure 7 shows the total profits made by the Deep Q-Learning Agent when tested (Total training rewards = -33, profits = 5189, test reward = 5, profits= 2051). In the figure below, our experimental results shows that the Double Q Learning algorithm made the best choices in terms of hold, buy or sell. Therefore receiving a better reward (Figure 11).
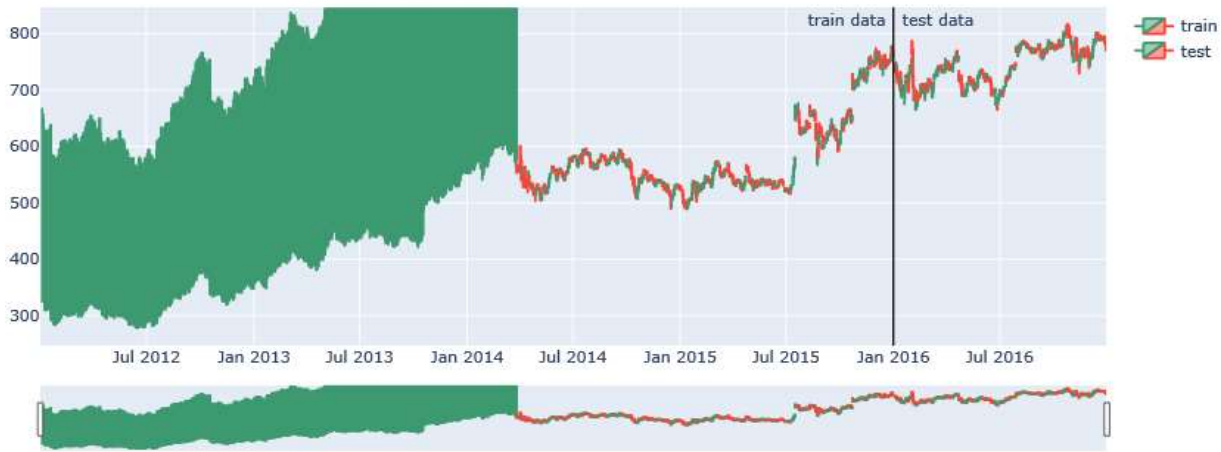

Figure 4: A Chart representation of Google Stock Market Data from July 2012- July 2016

We analyze the google stock market data using a chart (Japanese Candle stick to be precise). We divided the data into two parts; the first part that is from July 2012-July 2015 was used for training while the second part that is Jan 2016-July 2016 was used for testing.

5	0.0999999999999992	5025	-81.8	167635574.18832874	46.96183729171753
10	0.0999999999999992	10050	-244.6	900157.962435913	42.39922547340393
15	0.0999999999999992	15075	-285.4	261861.93145208358	42.155256032943726
20	0.0999999999999992	20100	-293.2	124006.48863092661	42.15951490402222
25	0.0999999999999992	25125	-199.6	66662.23405396938	43.802566676673889
30	0.0999999999999992	30150	-180.6	154383.84677546023	45.14868140220642
35	0.0999999999999992	35175	-138.8	27417.99582029581	45.576520681381226
40	0.0999999999999992	40200	-149.8	10372.84915638417	48.263673305511475
45	0.0999999999999992	45225	-93.2	19484.584003716707	52.81578612327576
50	0.0999999999999992	50250	-133.0	3429.161796119064	46.77142381668091

Figure 5: Showing the training process with result of the Deep Q-Network

The above training and result are as follows. The first column is the epoch, second column is the epsilon, third column is total_steps taking by the agent, fourth column is the log_reward (which I the sum of the total reward in the total steps taken by the agent, fifth column is the log_loss (total loss in the total steps) and the Sixth column is the elapsed_time (That is the elapsed time taken to move to the next step.
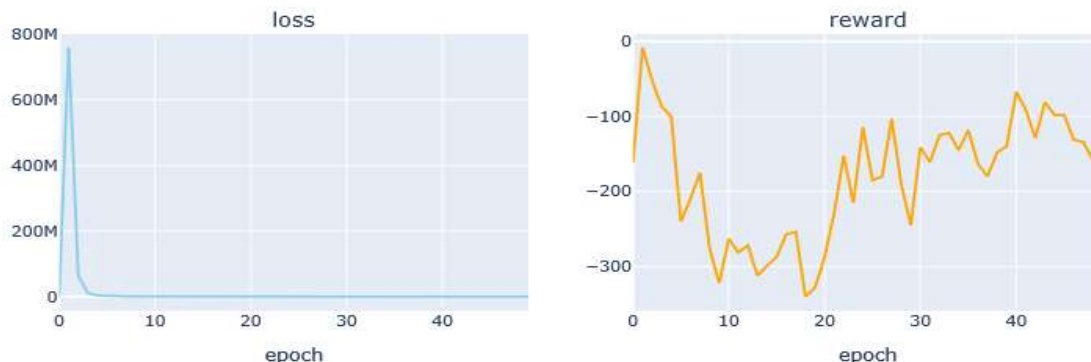

Figure 6 Showing graphical representation of loss and reward of the Deep Q-Learning Network

Figure 7: Showing the training and testing data of the Deep Q-Learning Network

5	0.0999999999999992	5025	-104.8	1398.542899391055	27.24966859817505
10	0.0999999999999992	10050	-27.4	963.1990051284432	25.915453672409058
15	0.0999999999999992	15075	-21.2	99.64850127045065	25.95596694946289
20	0.0999999999999992	20100	-11.6	42.883525209408255	26.609519481658936
25	0.0999999999999992	25125	-12.6	30.493958994653077	28.867915153503418
30	0.0999999999999992	30150	-5.0	24.44919595096726	29.2444109916687
35	0.0999999999999992	35175	8.6	8.540228485967964	30.609916925430298
40	0.0999999999999992	40200	14.8	6.449918887659441	30.879780769348145
45	0.0999999999999992	45225	35.4	9.002095349796582	31.86078190803528
50	0.0999999999999992	50250	6.2	6.031558384248638	34.26875448226929

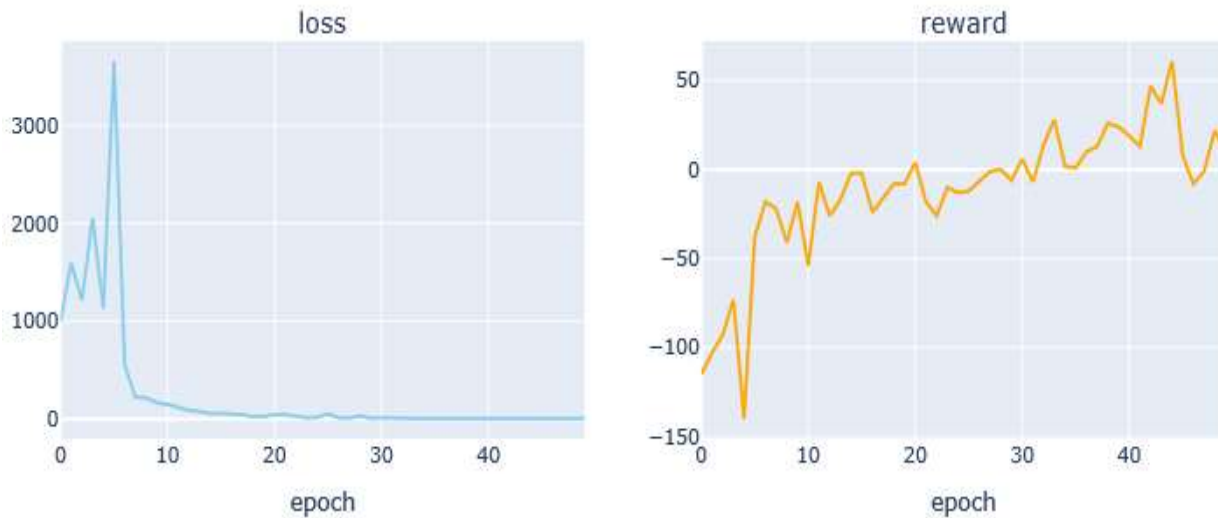Figure 8: Showing the training process with result of the Double Q-Network

Figure 9: Showing graphical representation of loss and reward of the Deep Q-Learning Network

Double DQN: train s-reward 0, profits 0, test s-reward 0, profits 0

Figure 10: Showing the training and testing data of the Double Q-Learning Network.

Here the agent choses to hold, therefore decides not to buy or sell, therefore getting a zero reward as well as a zero profit

5	0.0999999999999992	5025	-38.4	530.6218514889479	58.03651523590088
10	0.0999999999999992	10050	-1.0	24.748689428530632	58.721792459487915
15	0.0999999999999992	15075	-0.6	17.16154057532549	54.763424634933347
20	0.0999999999999992	20100	-5.6	4.80273076914018	55.00055932998657
25	0.0999999999999992	25125	5.6	26.430136572360063	57.45158767700195
30	0.0999999999999992	30150	-1.8	3.613614904589485	59.477909564971924
35	0.0999999999999992	35175	29.4	5.854322519496781	58.28664946556091
40	0.0999999999999992	40200	2.4	4.99849011699528566	55.98628258705139
45	0.0999999999999992	45225	-2.2	5.254588244436309	64.27919292449951
50	0.0999999999999992	50250	14.0	7.765418655442772	56.761634349823

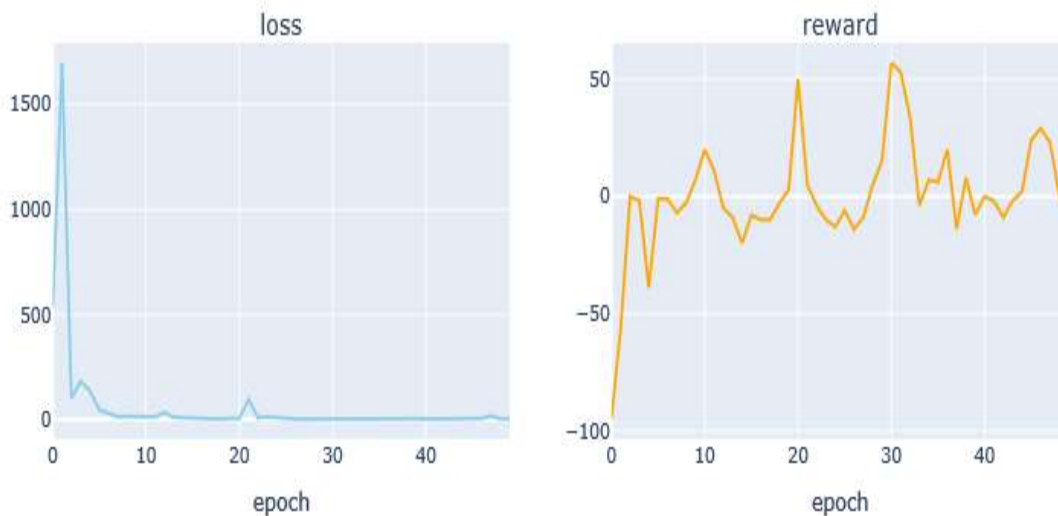Figure 11: Showing the training process with result of the Dueling Double Q-Network

Figure 12: Showing graphical representation of loss and reward of the Dueling Deep Q-Learning Network

Figure 13: Showing the training and testing data of the Dueling Double Q-Learning Network.

5. Conclusion

The stock market has become more essential in recent times with many persons trying to invest in the market. Financial Market has a critical function in the monetary advancement of any Country. Henceforth Nigeria and other non-industrial nation's development profoundly relies upon Stock and the Stock Market Performance. This paper present a Reinforcement Learning approach in predicting google stock market price. The agent was trained using three Reinforcement Learning Algorithms. The algorithms used are Deep Q-Learning Network, Double Q-Learning Network, and Dueling Double Q-Learning Network. Our experimental results shows that Double Q-Learning Network outperforms the Deep Q-Learning Network and Dueling Double Q-Learning Network byreceiving the highest reward in each state that the agent tries to take the best decision that is to say, buying the movement of the market is an upward trend, selling when the movement of the market is a downward trend and holding back when the movement of the market is unstable. The Dueling Q-Learning Network had the best reward of about 29.4 on 35 epoch, Double Q-Learning had reward of about 35.4 on 45 epoch, and the Deep Q-Learning Network had reward of about -81.8 on the first epoch. This paper can further be extended by building a trading bot using the mentioned three algorithm and test it on a demo account on meta4 or meta5 trading software. It can further by extended by changing some parameters in the mentioned threshold functions mentioned in the result and discussion so as obtain a more efficient result.

Reference

[1]. 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.

[2]. A. 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.

[3]. 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.

[4]. Zhang, X., Li, A., Pan, R."Stock trend prediction based on a new status box method and adaboost probabilistic support vector machine". Appl. Soft Comput. Vol.49, 385–398, 2016.

[5]. Hiransha, M., Gopalakrishnan, E.A., Menon, V.K., Soman, K. "Nse stock market prediction using deep-learning models", Procedia computer science 132, 1351–1362, 2018.

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

[7]. F. Rundo "Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems" Multidisciplinary Digital Publishing Institute (MDPI): Articles on applied science, pp.1-18, 2019.

[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]. G. Jeong, H. YoungKim "Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning" Expert System with Applications, vol. 117, pp.125-138, 2019.

[11]. O. Jangmin, J. Lee, J.W. Lee, B. Zhang "Adaptive stock trading with dynamic asset allocation using reinforcement learning" Information Sciences, vol. 176, pp. 2121-2147,2006.

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

[13]. X. Pang, W. Lin "An innovative neural network approach for stock market prediction" Journal of supercomputing, vol.76, issue.3, pp.2098-2118, 2020