International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 01 | January - 2023 Impact Factor: 7.185 ISSN: 2582-3930

OPTIMIZING STOCK TRADING STRATEGY WITH REINFORCEMENT LEARNING

1st Ankita More
Computer Department
Pune Institute Of Computer Technology
Pune,India

2nd Chetan More
Computer Department
Pune Institute Of Computer Technology
Pune,India

3rd Mehek Sirwani
Computer Department
Pune Institute Of Computer Technology
Pune,India

4th Guide-Yogesh Handge
Computer Department
Pune Institute Of Computer Technology
Pune,India

***Abstract*—We propose to determine a single exchange-traded fund (SPY) investing strategy That will maximize our total wealth. We compare our findings to the tried-and-true "buy-and-hold" and Moving average convergence divergence (MACD) strategies. Here, we Intend to base an evaluation on every essential standard that is considered when establishing the pricing. As mentioned, we intend to forecast future price Fluctuations for a specific stock. Prices from previous days, as well as financial Media stories connected to the firm of interest, are used to create these forecasts. Reinforcement learning is all about taking the right steps to maximize your reward In a given situation. It is used by a variety of software and computers to determine The best feasible action or path in a given situation. The Reinforcement Machine Learning model is employed in this work to forecast the closure price using past Data. We develop a predictor for multiple firms using these trained models that Forecasts the every day close stock prices. By providing inputs such as open, high, And low prices, stock volume, and the latest events about each firm, this predictor May be used to determine the price at which the stock value will close for a certain Day. The goal of this project is to learn and get hands-on experience in Data Analytics and Machine Learning.**

***Index Terms*—Algorithmic trading, deep learning, state representation learning, imitation learning,reinforcement learning.**

I. INTRODUCTION

A stock or share is known as a financial token or instrument that signifies the ownership of a company in some proportion. For example, Netflix Company had 100 shares and you bought one share. It means that you would own one 100th of Netflix. In reality, companies like Netflix have millions of shares but statement effectively sums up the argument: When you hold shares of a corporation, you effectively own a chunk of it if the company increases your stock price also increases. In the 1600s the European countries were colonizing and they are looking for Pollachius products which they couldn't have had found in the European countries. So, they began to import and export from the colonies or started trading gold, porcelain, spices, and silks around the globe. They have three options for trading via roadways, railways, and waterways. But running this massive operation wasn't cheaper. While shipping was a good and a cheaper option, they're lifeless and they were affected by pirates and natural calamities like storm cyclones, etc. This led to losses for the merchants, so they decided to start asking for money from investors or the public. This led to the very first stock exchange of the world is Amsterdam Stock Exchange, based in Amsterdam, Netherland. The Amsterdam Stock Exchange in 1602 was the first stock exchange led by the Dutch East India Company is said to be valued at around 7.8 trillion today also known under the initials VOC (Vereenigde Oostindische Compagnie). In order to fund their expensive voyages, the company turned to private citizens or individuals who could invest money to support the trip in exchange for a share of the ship profits. This practice allowed the company to afford even grand voyages, increasing profits for both themselves and their savvy investors. Selling these shares in coffee houses and shipping ports across the continent, the Dutch East India Company unknowingly invented the world's first stock market. Since then, companies have been collecting funds from willing investors to support all kinds of businesses. Today the stock market has schools, careers, and even whole television channels dedicated to understand them. But the modern stock market is significantly more complex than its original incarnation. Later on, in London stock exchange was started and named as London Stock Exchange (LSE) in 1801. Later in the United States of America (USA), New York Stock Exchange (NYSE) was started in 1817. In India, between1861 to 1865 the American civil war took place that resulted in an acute global shortage of cotton. All the Langsam mills had to turn to other sources like India, as cotton was exported from India in very large quantities. For example, Surat cotton price went from a price of 3 pence per pound to 20 pence per pound; that's more than six times. Between1861 to 1865, the export of cotton resulted in unusual changes in the form of bullion to 850 million rupees, much of which landed in

Bombay which naturally brought massive investment opportunities. Such opportunities rang up spontaneously and brought a new generation of entrepreneurship being in the right place at the right time played an important role. Mr. Premchand Roychand is formally known as cotton King who by virtue of being the first Indian broker to read and write in English, found himself suited to lead the equity markets in Mumbai at Dalal Street, Pin - 400001. He helped to float several companies with equity investment with increasing investment. There was an increase in activity and the number of brokers was increased to 200 - 250. When things cooled down at the end of the American civil war, official trading in the stock market begin in India. Twenty-five brokers came together to form "The Native Share and Stockbrokers Association" by contributing just Rs.1 from each one. Before the establishment of the Stock market in India, people used to sit under Banyan trees and trade physical shares. Later, this stock exchange Mumbai was renamed as Bombay Stock Exchange (BSE) in the year 2002. The first fully automated electronic exchange with a nationwide presence was incorporated as National Stock Exchange (NSE) in 1992. As the stock market helps companies to raise capital, generates personal wealth for an individual investor, serves as an indicator of the state of the economy, and helps to increase investment. So, it got more importance in this research

II. Motivation

Recently Trading is being done enormously by individuals, small and large stakeholders. It is a platform for expanding Wealth and Capital dynamically without employing new infrastructure, personnel, or software systems. It has made the life of individual stakeholders and the work of business corporations so much easier by providing them with stocks services at exceptionally low costs of platform charges. However, capital remains vulnerable in stock trading. Lack of protection of wealth is the only major concern that hinders increased use of Stock Trading. Thus there is need of some algorithm that can reduce wealth losses. Thus, to take care of security of both wealth and Capital we use a technique called machine learning.

III. Literature Survey

The Following shows the literature survey by comparing techniques propose in various references:

A. Reinforcement learning

Summary:It is a machine learning technique that focuses on training an algorithm following the cut-and-try approach. The algorithm (agent) evaluates a current situation (state), takes an action, and receives feedback (reward) from the environment after each act.

Limitation:Too much reinforcement learning can lead to an overload of states, which can diminish the results. Reinforcement learning is not preferable to use for solving simple problems. Reinforcement learning needs a lot of data and a lot of computation. It is data-hungry

B. Markov Decision Process

Summary:Markov decision process (MDP) is a discrete-time stochastic control process. It provides a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker

Limitation:If the time interval is too short, then Markov models are inappropriate because the individual displacements are not random, but rather are deterministically related in time. This example suggests that Markov models are generally inappropriate over sufficiently short time intervals.

C. Q-Learning

Summary:Q-learning is a model-free, off-policy reinforcement learning that will find the best course of action, given the current state of the agent. Depending on where the agent is in the environment, it will decide the next action to be taken

Limitation:A major limitation of Q-learning is that it is only works in environments with discrete and finite state and action spaces.

IV. Problem Statement And Scope

A. Problem Statement

The accuracy of the existing stock market prediction models is low because only a small dataset is used for training, the results will be less accurate. There is still a need to continually explore more new features that are more predictable. Even though multiple algorithms exist, there is no real-life implementation of these ideas for the beneficial of people. Efficient algorithms should be made available with easy accessibility and interface.

B. Scope

The purpose of this study is to understand the term Stock Trading using machine learning which is an algorithm that safeguards wealth. In future stock trading can become more robust using machine learning algorithms. We propose a system which uses multiple algorithms and models to provide more profit in Stock Market.

V. Algorithm Utilized

A. Reinforcement Learning

Machine learning models are taught to make a series of judgments via Reinforcement learning. In an unpredictable and potentially complex environment, The agent must learn to attain a goal. Artificial intelligence is put in a game-like Environment in reinforcement learning. To find a solution to the problem, the Computer uses trial and error. Artificial intelligence is given either rewards or Penalties for the acts it takes to get it to accomplish what the programmer Desires. Its purpose is to increase the total prize as much as possible. Even though the designer establishes the reward policy–that is, the game's Rules–he provides the model with no clues or ideas for how to solve the game. Starting with completely random trials and progressing to sophisticated tactics and Superhuman

skills, it is up to the model to find out how to do the task in order to Maximize the reward.

In reinforcement learning, developers use a framework that rewards desired actions While penalizing negative ones. To motivate the agent, this strategy assigns Positive values to desired acts and negative values to undesirable behaviours. To Obtain an ideal solution, the agent is programmed to seek long-term and greatest Overall return.

These long-term objectives keep the agent from stagnating on smaller objectives. The agent eventually learns to ignore the unpleasant and focus on the good. This Technique of learning has been used in artificial intelligence (AI) to drive Unsupervised machine learning using incentives and punishments.

Trading is a never-ending process with no conclusion in sight. Because we don't Have comprehensive information about the traders in the market, trading is likewise A stochastic Markov Decision Process. We employ model-free reinforcement Learning, also known as Q-Learning because we don't know the reward function or Transition probability.

B. Markov Decision Process (MDP)

It is a mathematical framework used for modelling decision-making problems Where the outcomes are partly random and partly controllable. A model of anticipating outcomes is the Markov decision process. The model, like a Markov chain, tries to predict a result based solely on the present state's Information. The Markov decision process, on the other hand, takes into account The features of actions and motivations. The decision-maker may choose an action Accessible in the current state at each phase of the process, causing the model to Advance to the next step and rewarding the decision-maker. At each step during the process, the decision-maker may choose to take any action Available in the current state, resulting in the model moving to the next step and Offering the decision-maker a reward.

An optimization problem may be given to a machine learning algorithm. The Programme will use reinforcement learning to try to optimize the behaviors made Inside a given environment in order to maximize the potential reward.

Reinforcement learning employs Markov decision processes to establish an ideal Balance of exploration and exploitation, whereas supervised learning approaches Need accurate input/output pairings to construct a model. When the probability and Rewards of a result are undefined or unknown, machine learning may employ Reinforcement learning through the Markov decision process.

C. Q-Learning

The 'Q in Q-learning stands for quality. Quality in this case represents how useful a Given action is in gaining some future reward. Q-learning is an off-policy Reinforcement learning algorithm that seeks to find the best action to take given The current state. It's considered off-policy because the q-learning function learns From actions that are outside the current policy, like taking random actions, and therefore a policy isn't needed. More specifically, q-learning seeks to learn a policy That maximizes the total reward. An agent interacts with the environment in one of two ways – exploit or explore And a Q-table is produced with the dimensions. An exploit option implies that all Options are evaluated and the one with the greatest environmental benefit is Chosen. An explore option is one that considers a random action without regard for The maximum future payoff.

VI. METHODOLOGY

A. Import necessary Libraries

Import The Required Libraries For Building model and Analysis Purpose. By importing libraries analytical Calculations becomes frugal and efficient.

B. Load Dataset

Data on stock market values from 2013 to 2018 is included in this dataset. This dataset contains the following information: date, open, high, low, and close values, as well as volume and stock names. Also, the dataset is in CSV format. The following is a quick description of each component in the dataset:

1. DATE - The day on which stock is exchanged.
2. OPEN - Any listed stock's daily opening price is the price at which it is initially traded.
3. HIGH - The maximum price at which a stock can be bought or sold during a trading day is known as the high.
4. LOW - The minimum price at which a stock can be bought or sold during a trading day is known as the low.
5. CLOSE - During a standard trading session, the last price at which a stock trades is referred to as the closing price.
6. VOLUME - The total number of shares or contracts exchanged in a securities or market within a certain time period.
7. NAME - The name of the stock.

C. Perform Exploratory Data Analysis

In every Data Analysis or Data Science project, exploratory data analysis, or EDA, is a critical stage. EDA is the investigation of a dataset to find patterns and anomalies as well as to generate hypotheses based on our knowledge of the information. During the course of a data science or machine learning project, EDA consumes around half of the time spent on data analysis, feature selection, feature engineering, and other processes. Because it is the most essential element or backbone of a data science project, where one has to perform several tasks like data cleaning, dealing with missing values, handling outliers, treating unbalanced datasets, handling categorical features, and many others. To automate our tasks in exploratory data analysis, we may utilise python packages like dtale, pandas profiling, sweetviz, and autoviz.

D. Create An Environment

In the reinforcement learning problem, the environment is a critical component. It's critical to have a solid grasp of the underlying world with which the RL agent will interact. This aids us in developing the best design and learning strategy for the agent. CREATE AN ENVIRONMENT In the

reinforcement learning problem, the environment is a critical component. It's critical to have a solid grasp of the underlying world with which the RL agent will interact. This aids us in developing the best design and learning strategy for the agent. The reinforcement learning task is intended to be a simple framing of the challenge of learning from interaction in order to accomplish a goal. The agent is the learner and decision-maker. The environment is the object it interacts with, and it consists of everything outside the agent. The agent chooses actions, and the environment reacts to those actions, presenting the agent with new scenarios. The environment also produces rewards, which are unique numerical values that the agent attempts to maximise over time.

E. Prepare Data

Process the data using the Pre-processing Techniques for further usage and Analysis Purpose. Remove The null values and Replace it with appropriate values.

F. Train Data

The data and Dataset used for the Training of Model is Called as The Train Data.

G. Test Data

The data and Dataset used for the Testing of Model is Called as The Test Data.

H. Evaluate Model

Model evaluation is a step in the model creation process that is often overlooked. This is the stage where the model's performance is determined. As a result, it's important to think about the model's results in terms of every conceivable assessment technique. Using various approaches can result in a variety of viewpoints.

I. Deploy Model

The main Task of project to Host the model to get the appropriate Results Based on Datasets and Databases. Deploying model means actual execution of the Model

VII. Limitation And Future Work

This research provides the first step towards users Capital in securing their wealth before giving it all away on losses. The research has explored the method of using different Machine Learning Model. More accuracy can also be achieved in the same but advanced manner.

We plan to apply our work to future markets where both long and short positions are always possible.

VIII. Output

Following are the Snippets of The Implementation of Model

Fig. 1. Figure of Evaluation of Model.

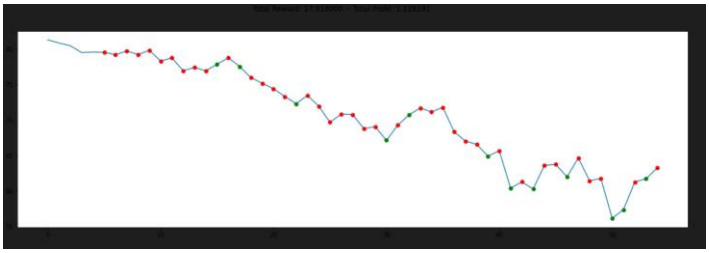

Fig. 2. Figure of Result of Model.

IX. Conclusion

Hence, we proposed a practical algorithmic trading method, Reinforcement-Trader, which achieves good profit using only long positions. We used reinforcement learning to learn a profitable trading policy from nonstationary and noisy stock data. We used dimensionality reduction and clustering to extract robust features.We also co-trained a regression model with a reinforcement learning model to provide accurate state information for decision-making. In the imitative reinforcement learning.

Points To Consider

[1] Gamification of trading - Historical and Current Prices, Technical Data, Buy/Sell/Do Nothing, Profit and Loss.

[2] Train the system - Each entry and exit is an individual game, Run through the price series sequentially and randomly.

[3]Reward function design - Pure Profit and Loss on exit, otherwise zero, Profit and Loss from the start of trade to every time step t.

[4] Features to use - Open, High, Low, Close, Volume , Technical. indicators, Time of day, Day of Week, Time of year, Diffrent time granulity.

[5] Test the system - Sine waves, Trend curves, Random walks, Adding Noise to clean test curves.

[6] Types of Deep Learning Algorithms.

Acknowledgment

It was indeed a great learning for me to write this report on" OPTIMIZING STOCK TRADING STRATEGY WITH

REINFORCEMENT LEARNING". I would like to thank our Seminar Coordinator Prof.D.D.Kadam, Head of Department Dr. G. Kale and Principal Dr.S.T.Gandhe for their encouragement and support. I would also like to thank my guide Prof.Handage, Department of Computer Engineering for his guidance and help. He has continuously motivated me throughout the seminar project and thus was of significant help to make it successful.

REFERENCES

[1] Y. Deng ,F. Bao ,Y. Kong ,Z. Ren and Q. Dai"Deep direct reinforcement learning for financial signal representation and trading" IEEE Trans. Neural Netw. Learn. Syst. vol. 28 no. 3 pp. 653-664 Mar. 2021.
[2] Y. Li W. Zheng and Z. Zheng "Deep robust reinforcement learning for practical algorithmic trading" IEEE Access vol. 7 pp. 108014-108022 2019.
[3] D. Fengqian and L. Chao "An adaptive financial trading system using deep reinforcement learning with candlestick decomposing features" IEEE Access vol. 8 pp. 63666-63678 2020.
[4] Deog.Yeong Park, "Practical Algorithmic Trading Using State Representation Learning and Imitative Reinforcement Learning," Ki.Hoon.Lee