AAPL stock trading using A2C

Joel Poah Zhi Kang  
Diploma in Applied Artificial Intelligence and Analytics  
School of Computing   
Singapore Polytechnic

Joel.21@ichat.sp.edu.sg

*Abstract – Stock trading is one of the most fascinating things to learn about and a topic starter when stock trading is mentioned. Reinforcement learning in deep learning methods has allowed stock traders to utilize robo investor.*

Keywords—Deep learning, Artificial Intelligence, Reinforcement Learning(key words)

# Introduction

Stock trading— An ear-catching phrase that sparks conversations when mentioned about it. Stock traders are always looking out for the next best opportunity to make as much money as possible through the stock market .Stock trading platforms has also noticed this and implemented robot investors that utilise artificial intelligence and deep learning methods to show indicators or even help users trade.

There are many tempting AI platforms out there which offer services for custom indicators that helps to trade. Today I will be uncovering how robot investors decide whether to buy or hold a stock and whether it is worth investing in AI robot investors service and at the same time discuss how reinforcement learning methods are justified.

Reinforcement learning is a deep learning method where there is an Environment where an agent can move around in. The agent can perform actions in various states it finds itself in. The agent will then try to approximate and find the best move forward that will give the most reward or return(long term reward) . However stock trading is not just simply a game or an environment where an agent can play in without consequences. Bad decisions made my the agent can cost the user who invest in the AI a lot of money .

# Related Works

Reinforcement learning methods have been around for a decade and researchers like Jiang, Z., Xu, D., Liang, J., 2017. have tapped into this algorithm for stock trading and on top of that introduced the recurrent learning feature (Jiang et al., 2017) with the Ensemble of Interdependent Instance Evolutions (EIIE). They combined multiple independent models to enhance performance of portfolio management where the model contains small recurrent subnet and takes in price inputs of individual assets before submitting a voting score to the SoftMax function (which determines the potential growth of an asset).[1]

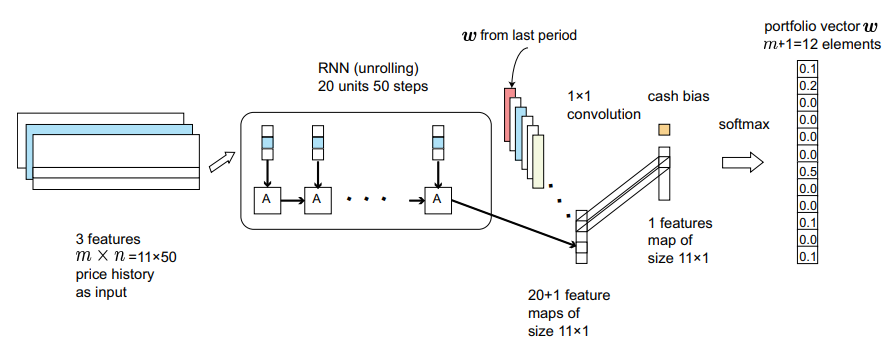


Fig. 1. RNN implementation of EIIE (Jiang et al., 2017)

# Methodology

In this paper, the following Python libraries will be used:

**Open AI Gym Environment** – An open source environment that include multiple games and real life environments that can be used to train or test deep learning models and their effectiveness in a given environment

**Matplotlib** - A visualization library for visualizing the action made by the agent in a given state and also to show the price of the stock .This can help us determine if the agent is learning to make the right move .

**TensorFlow** – A Deep learning framework which is often seen as the most popular deep learning framework, which is very powerful & easy-to-use and has excellent community support.

**Stable-Baselines3** – Stable Baselines offers high-quality implementations of Reinforcement learning algorithms built on top of TensorFlow where we can train and optimize it for any given environment or Q-learning problems.

## Data

Data is retrieved from marketwatch.com where they provide any known stock data out there and its free to download in a csv format of up to 1 year per csv file.

## Exploratory Data Analysis

Exploring the data is important to draw useful insights that can help improve the model or better feature engineer it. I first used pandas to import the csv file that was downloaded from the marketwatch.com. The Data was downloaded on 2 February 2023 which contains Data from 2 February 2022 to date.

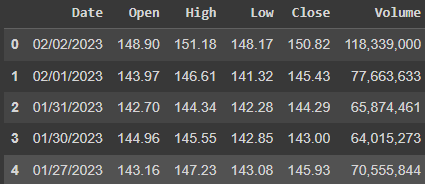


Fig. 2 Head of Pandas Data frame AAPL stock 2023

There were no missing rows of data with the open , high , low ,close and volume of the AAPL stock for each day. Open refers to the price in which a stock first trades upon the opening of a trading session while close is the stock price at the end of the day for the US market. I also noticed that the Open price may not be the same as the Close price of the previous day because of the market activity that is happening during the night of the US market. High is the highest price of a specific stock in the day which was traded. Volume refers to the number of shares that has been traded in the day and is a key indicator of the public’s interest in a particular stock .

### Data visualisation

There were so many columns when coming to decide which key to showcase the distribution of a typist, but I have decided that it is most appropriate and accurate to analyze the middle character of the password that was typed which is the ‘R’ since it probably can distinguish the typists the most because it is capitalized and also affected by the way the first few characters were typed. I will be analysing the columns ‘DD.five.Shift.r’ , ‘UD.five.Shift.r’ and ‘H.Shift.r’ out of the columns shown below.

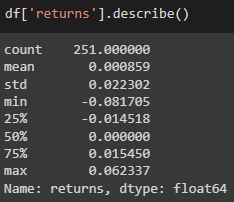
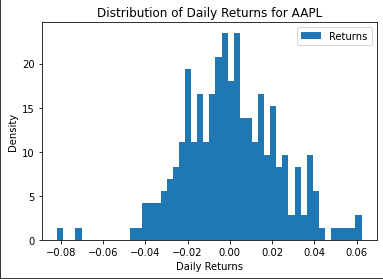


Fig. 4 Histogram of Daily return for AAPL stock in the past year

We see a normal distribution formed for the AAPL stock where the return is calculated by the taking the percentage change of the stock value after the day ends. A normal distribution is when the mean, mode and median of the sample are equal and its important because it helps us estimate for probability distribution of the returns of the AAPL stock (Malik, 2020)

Based on the normal distribution of returns around 68% of the values will be within 1 standard deviation aka -2% to + 2% , 95% within 2 standard deviations or -4.4% to + 4.4% and 99% of values will fall within 3 standard deviations or -6.6% to 6.6% in a single day.

### Bollinger Bands Analysis

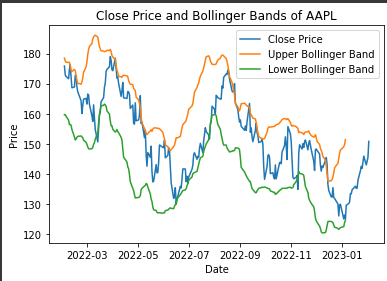


Fig.4 Bollinger Bands of AAPL

Analyzing the AAPL stock with Bollinger bands utilizing a moving window of 20 days and 2 standard deviations away from the mean. Based on the Bollinger bands of the AAPL stock it seems like the AAPL stock is on a uptrend for the first 2 months of 2023 and may continue to be bullish till April but this is not financial advise .

## Feature Engineering / Preproccessing

### Data Wrangling

The Data column is not in the exact date type for the Open AI gym environment. Initially it was in the date format ‘02/02/23’while the environment gym-anytrading requires a datetime dtype in order for the agent to be trained properly. We can do so just by using the pandas built in function to\_datetime(). [5] On top of that the Volume column is an object type which the agent may not properly benefit from it if its not recognized as a proper number. Thus we will first have to remove the ‘,’ found in the Volume object before converting to an integer using the .astype() function.

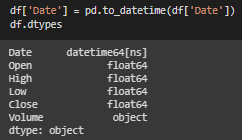


Fig.5 Converting to Date type in Python

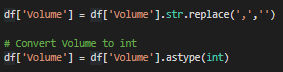


Fig.6 Converting to integer in Python

### Splitting train and test data

The Data is ranged from 2 Feb 2022 to 2 Feb 2023. I will use the first 230 days as training and test on the last 20 days to see if the RL model will be able to trade based on the past 230 days of data.

# A2C for Stock Trading

## Neural Network

Neural networks are a subset of machine learning, a combination of modern computer science and cognitive psychology. Data processing and analysis are based on the same processes that neurons in the brain process information. (J, Jeremy,2017) Your neurons require a certain amount of activation energy to fire and send information out of the neuron to axons, synapses, other neurons, etc. Data is fed into fully connected neural networks and matrix multiplication is carried out between the weights and the input. The weights are then improved through back propagation which reduces the validation loss of the function.

## Reinforcement Learning

Reinforcement learning is an area of Machine Learning. The goal for the machine is to choose the most suitable action to maximize the reward or return in a particular system (“Reinforcement learning - GeeksforGeeks,” 2018). However, in an environment that is very big and its impossible to memorize the best move for every state action pair. That is where Neural Network come into play to act as function approximators to approximate the mapping of state to actions

## How Advantage Actor Critic(A2C) works

A2C combines 2 types of reinforcement learning algorithms (Policy Based and Value Based) together. Policy Based agents directly learns a probability distribution of input states to output actions while Value Based algorithms learn to choose actions based on the Q-value(Wang, 2021)

# Results and discussions

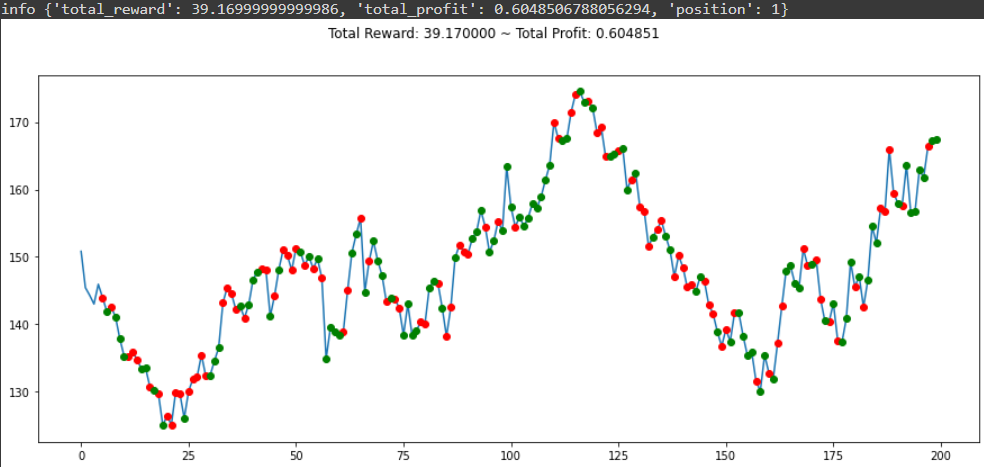


Fig.7 Random action agent performance

Here’s a dummy model which takes in a random action given an environment . There are 2 actions the agent can make and since its random , the model just decides to randomly buy and sell the AAPL stock. Profit is calculated by giving the agent an environment .In order to get a positive profit , Total Profit > 1.



Fig. 8. Trained Final A2C model on AAPL Stock

Above shows a trained final model using stable baselines3 A2C model where it only bought on very rare occasions(green) and took very little risk by selling 1-2 days later after each buy. The Model managed to earn a 2% profit from AAPL stock in 20 days.

# Conclusion

The model was able to break even and even made a 2% profit. However stock trading with A2C should be financial advise and it does not guarantee a return in investment every time.

Analyzing stock prices based on its opening and closing price isn’t the only data that can be fed into a predictor but I think there are many other factors that will affect the price of a stock such as media and demand of the product made by the company at a point in time. Nonetheless, all relevant information such as the opening and closing price of the stock are still relevant information based on the philosophy of technical traders (Charles et al., 2006; Lo et al., 2000), to be reflected in the prices of the assets, which are publicly available to the agent[?]

##### Acknowledgment

Thank you lecturers for the opportunity to learn about reinforcement learning and to explore ways to implement reinforcement learning to real life context like trading stocks.

##### References

1. D Charles, II Kirkpatrick, and Julie R Dahlquist. Technical analysis: The complete resource for financial market technician. ISBN-13, pages 978–0137059447, 2006.
2. Jiang, Z., Xu, D., Liang, J., 2017. A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem [WWW Document]. arXiv.org. URL https://arxiv.org/abs/1706.10059v2 (accessed 2.3.23).
3. Malik, F., 2020. Ever Wondered Why Normal Distribution Is So Important? [WWW Document]. Medium. URL https://medium.com/fintechexplained/ever-wondered-why-normal-distribution-is-so-important-110a482abee3 (accessed 2.4.23).
4. Neural networks: representation. [WWW Document], 2017. . Jeremy Jordan. URL https://www.jeremyjordan.me/intro-to-neural-networks/ (accessed 2.4.23).
5. Reinforcement learning - GeeksforGeeks [WWW Document], 2018. . GeeksforGeeks. URL https://www.geeksforgeeks.org/what-is-reinforcement-learning/ (accessed 2.4.23).
6. Wang, M., 2021. Advantage Actor Critic Tutorial: minA2C [WWW Document]. Medium. URL https://towardsdatascience.com/advantage-actor-critic-tutorial-mina2c-7a3249962fc8 (accessed 2.4.23).