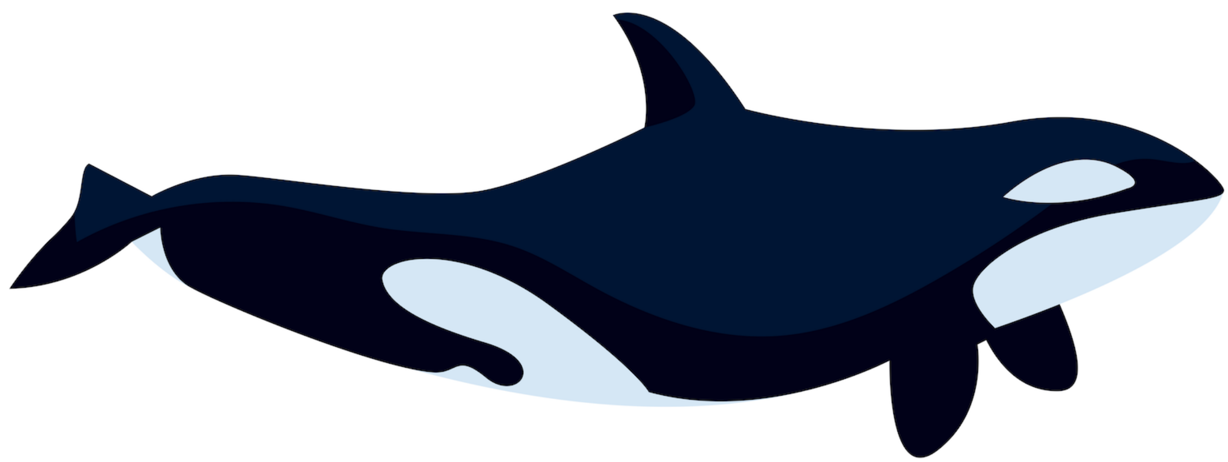
The Wall Street Orca

By Daniel Vishna, Ghaleb Butto, Elias Wakile

The Wall Street Orca



Investment problems made easy

Abstract

Individuals are increasingly interested in investing in stock market as a side hustle to increase their income. However, many are inexperienced investors who are unsure how and where to invest, how long to hold or even sell stocks for optimal returns. Respetively, the number of traders in the stock market increased, making the market became highly dynamic, requiring a smaller reaction time from the investor. In the recent years, to be able to keep up with the market fluidity, wallstreet sharks but also among average investors began using artificial intelligence systems for their ability to outperform investors by their capacity to make rational decisions, unveiling hidden market trends, analyze investment risks, and forecast future profitability [Elaji 2021]. For this reason, we propose a HFT (High Frequency Trading) capable online Deep Q-Learning agent that when provided with a continuous input of stock analytics, unremittingly trades in the stock market. To guarantee the success of the trades we chose to engineer our Learning agent to trade the toppers of the S&P 500 due to their being highly traded. We tested the results of our learning agent by comparing its' profits to those of an extrapolation-based trading system as it mimics the trading strategy of an inexperienced investor. In our article, we provide an introduction to the world of stock trading. Afterwards, we gave account on the functionning of our model. Followingly, we presented the results from the testing of the algorithm. Finally, we presented the potential work that could be further made to increase the model's success.

1 Introduction To The Problem

1.1 What is the stock market?

A stock market is the aggregation of buyers and sellers of stocks which represent ownership in a company or corporation, which are claim on the buiseness' assets and earnings. The stock market is the most important way for a company to raise money - by selling shares of the company's ownership, an investor obtains a potential profit if the share value will increase, where in that case he will be able to sell the share for a higher price, hence making a profit. The ensemble of shares owned by a trader is called the trader's portfolio. The portfolio's value is derived by the value of the stocks it currently contains.

1.2 What does all the parameters of stocks mean?

Most website provide investors with 5 basic parameters of shares: volume, opening price, closing price, high price and low price.

- The volume: it is the volume of a stock is the number of its' shares traded over a specific period of time.
- The opening price: the price of a security at the opening of the period of time it is listed on an exchange.
- The closing price: the price of the last traded price in a security in a period of time.
- The high price: the highest price of a stock traded in the period of time.
- The low price: is the lowest traded price of a stock at a period of time.

1.3 How does the stock market work?

A trader makes an auction to buy or sell a certain amount of shares for a given price hoping for someone to make a bidding coinciding with his auction. In that case, a trade is made. Stocks with larger volumes have a higher chance for auctions to coincide with biddings. When there are more buyers than sellers, buyers could be ready to increase their offers to purchase the stock. As a result, sellers will increase their asking prices, driving up the cost and vice-versa.

1.4 What is happening in the stock market?

Many researches findings suggest that the price of a stock does not solely derive from the real value of the company [Morck, Shleifer, Vishny 1990]. To an inexperienced investor, this fact might be quite confusing, as we have previously defined a stock to be a partial ownership of a company. Thereupon, inexperienced investors often loose their money in their first investment. Thence, we all know a Cassandra who warns everyone that investing in the stock market will result in losing that money. Thus, how does the stock market really work ?

Being a market, the stock market mostly depends on supply and demand. Given that the supply is mostly fixed, we get that what defines the stock price is its' demand. There is a good evidence that the demand, and thus the stock price, is driven by inverstors' sentiments and beliefs which are not always rationally justifiable [Morck, Shleifer, Vishny 1990]. If one were to assume by contradiction that those are just noises that cancel out, he would reach to the false conclusion that wealth is redistributed on the stock market between savvy investors and reckless traders. Nonetheless, it is not the case [Morck, Shleifer, Vishny 1990]. Whereof, a prudent investor must vigilantly follow the market evolution and swiftly recognize opportunities. While this might be a nearly impossible and sysiphic task for the average human, it is quite an easy task for computers.

1.5 What is the problematic of the stock market?

For the previously mentioned reasons, the stock market nowadays requires from the investor a 24/7 attention which might be an impossible

task for a human but an easy task for a computer. Natheless, even though it is a felony, the market is often due to manipulation by malicious investors. A stock manipulator trades in front of other traders who are looking for information on the stock's real worth. These information seekers unmistakably increase market efficiency in a market devoid of manipulators by driving prices up to the level suggested by the informed party's information. More information seekers suggest more rivalry for shares, which might degrade market efficiency and make it simpler for market manipulators to enter the market. [Aggarwal, 2004]. Thence, making it a potential danger zone for the average person for whom investing is no more than a side hustle. Whereof, our relentless model, ready to always vigilant to trade in the stock market, permits the average investor to calmly invest in the stock market without attributing to it much attention.

2 The Model

2.1 What is Deep Q-Learning?

Reinforcement learning consists of an agent, a set of states, and a set of actions per state. The agent seeks to maximize its overall reward. This is accomplished by increasing the reward for reaching the present state by the greatest reward possible from future states. By performing an action, the state is changed, added to that, when an action is carried out in a particular state, the agent is rewarded (determined by the reward function). This essentially modifies the current action by the potential future reward. A model-free reinforcement learning technique called Q-learning is used to determine the worth of a given action in a given situation. The term "model-free" refers to the fact that it is not dependent on a model of the environment and that it is capable of handling stochastic transition and reward issues without the need for adjustments.

All states, actions, and rewards in the experience replay approach are influenced by earlier states. In other words, there are correlations between states, actions, and rewards. Due to those correlations, The approximation function is unable to robustly learn. Replaying an experience removes correlations by randomly extracting the learning data from the event and storing it in a cache [Jang 2019]. The target network and Q network are independently prepared using the target Q approach [Mnih, 2015]. By utilizing the target network to determine the target value and instructing the Q network to train based on the target value, correlations are decreased [Zhang, 2015].

The fundamental principle of deep Q-learning is that it combines Q-learning with an artificial neural network by using experiential replay [Zhang 2015]. When the agent interacts with the environment, specimens (s, a, r, and s') are generated. There are many conceivable environments and samples. Because of the correlation between the samples, if the agent learns from the samples produced in accordance with the scenario, the learning may proceed unusually. Deep Q-learning collects several samples to find a solution to this issue. Samples that have been placed in the memory are randomly organized and retrieved as frequently as feasible while CNN learns. However, employing an excessive amount of memory might slow down learning [Jang 2019].

2.2 How is our model implemented?

We implemented our Deep Q-Learning model in an object oriented manner to guarantee the understandability and extendability of the code to permit us to further develop the code in the future without having to completely change our code. First, we built an architecture of classes for which each class has a well-defined job. Hence, the claases:

- NeuralNetwork: The class implementing the Deep Q-Learning.
- PortFolio: The class implementing the PortFolio of our model as an investor

- Stock: The class describing the assets held in a given stock.
- StockData: The class saving the data for a given stock
- StockMarket: The class saving the data of all the stock market.

Each NeuralNetwork works on a specific share, we chose to implement our model in the following manner as to permit our model to learn the behavior of our stock at a given instance and to ensure that the situation in another stock will not interfere with the situation in our stock. A state is composed of the features of a stock we are going to give account of in the next subsection. Additionally, we implemented our own reward function for the model which we will be followingly discribing and explaining in the last sub-section of this section.

2.3 What is a state ?

A state is composed of 7 parameters of a stock which are:

- MACD
- ADX
- CCI
- RSI
- $EPDD/t/t-1$
- $EPDD/t-1/t-2$
- $EPDD/t-2/t-3$

Followingly, We will give account on those parameters:

2.3.1 MACD - Moving Average Convergence Divergence

This parameter is the respective momentum of a trend in the stock's price . Designed to unveil changes in the strength, direction, momentum and duration of a trend in a stock's price. It is computed by subtracting the 26-period from the 12-period exponential moving averages. The exponential moving average (EMA) is defined to be $EMA_{Today} = ClosingPrice_{Today} \times \frac{1}{1+periods} + EMA_{Yesterday} \times \frac{periods}{1+periods}$. Let us notice that it is possible for this parameter to notice a false reversal signal in the case of a triange pattern of in a momentum slowdown by giving a significantly lower value [Chong 2008]. Nevertheless, in our model we will exploit this bug as a useful feature signaling our model to sell stocks before the reaching of the stock's peak of ascent, since at the peak of ascent selling might be difficult and even almost impossible. This feature is quite important considering the fact our model is a high frequency trade model.

2.3.2 ADX - Average Directional Index

The average directional index is mainly used to track important trend movements. It ranges from 0 to 100 while the first quarter of range indicates a weak trend and the fourth quarter indicates an extremely strong trend. Normally, quarter breakout is a result of buyers-sellers disagreement on price, indicating an imbalance of demand and supply [Gurrib, 2018]. ADX can provide us signals for our strategies indicating for when to pullback from a stock by exploiting an imbalance. Its formula is $ADX = \frac{13 \times PriorDX + CurrentDX}{14}$ for $DX = 100 \times \frac{|(+DI(-(-DI)|}{|(+DI)+(-DI)|}$ for $\pm DI = 100 \times \frac{Smoothed\pm DM}{AverageTrueRange}$ for $+DM = CurrentHigh - PreviousHigh$ and $-DM = PreviousLow - CurrentLow$ with $Smoothed = CurrentDM - \sum_{i\in[14]} \frac{PrevDM_i}{14}$.

2.3.3 CCI - Commodity Channel Index

This parameter measures the distance between the current price and the historical average price. It indicates us when a stock is oversold or overbought. As an unbounded oscillator, the CCI can move endlessly higher or lower. Because of this, overbought and oversold levels are often identified for each specific asset by examining historically extreme CCI levels when the price reversed from. A CCI close to 0 indicates the stock price is close to its historical average price. Its formula is $CCI = \frac{TypicalPrice - MovingAverage}{0.015 Variance(TypicalPrice,, MovingAverage)}$ for $TypicalPrice = \frac{High+Low+close}{3}$ and $MovingAverage = \sum_{periods} \frac{TypicalPrice}{periods}$. [Shah, 2019]

2.3.4 RSI - Relative Strength Index

The RSI describes the speed and magnitude of a security's stock price. When the RSI is above 70 (a sell signal), a stock is deemed overvalued, and when it is below 30, it is deemed undervalued (a buy signal). The RSI offers a positive indication when it is above 50, and a bullish signal when it is below 50 for the security. Its formula is $RSI = 100 - \frac{100}{1+RS}$ for $RS = \frac{AverageGain}{AverageLoss}$ for the average gain and average loss being the average gain or loss in the last 14 periods of time respectively. [Chong, 2008]

2.3.5 Our own designed features

We designed for our net a feature we call the exponential proportional daily derivative of timestamp x and timestamp y ($EPDD/x/y$) for $x < y$ (by means that x is earlier than y) that is computed in the following manner: $\frac{e^y}{e^y+e^x}$. By using this parameter, we are capable of explaining to the computer the difference in prices from timestamp x to timestamp y in a parameter normalized from 0 to 1 making profits much more "intelligible" for our neural network. One can possibly express this parameter as $sigmoid(y-x)$. Given that this parameter is differentiable in $\mathbb{R}$ it will not pose much problem to the neural network computing the Q-values to compute the weight of the features in its' backtracking derivative. Note t the current timestamp and $t-1$ the previous timestamp and so on... We included as features the $EPDD/t/t-1$, $EPDD/t-1/t-2$ and finally $EPDD/t-2/t-3$.

2.4 What is the action universe?

We kept a relatively simple action space, the numbers from -5 to 5, the number denotes how many shares should the model buy/sell in the current moment (negative number for sale of stocks, 0 for keeping, positive number for buying). We chose a relatively narrow range of actions to guarantee the success of our trades (buys/sells) as in larger number of stocks one might not find bidders to bid his auction. Furthermore, a smaller actions universe

2.5 What is the reward function?

We chose a relatively simple reward function, in case of choosing to buy or sell a share, the given reward is 0. In the case of selling a share, the reward is the profit made by the selling of the share, which is the revenue from the selling the number of shares we chose to sell minus the amount of money we invested in the number of the shares we sold.

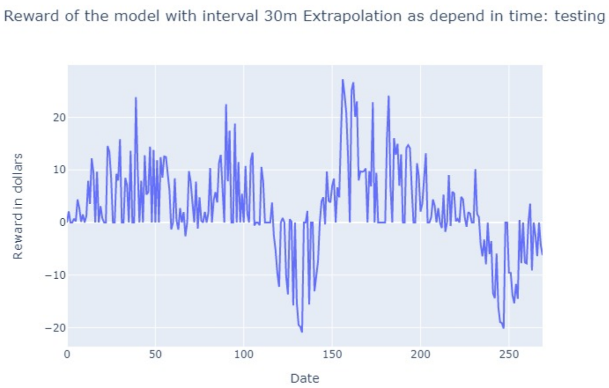
3 Results

3.1 The Baseline - Extrapolation

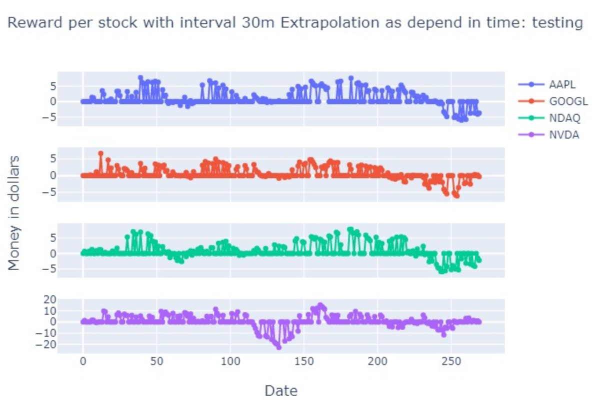
We kept a relatively simple baseline - by extrapolation. It's heuristic is relatively simple, if the share's closing price of the timestamp is greater than the opening's price (hence if the price's trend is going upwards - it's price increases) then buy a share (as long as he still has money left), if the closing price is smaller than the opening price (the stock's price drops - it's trendline is descending) sells a share, and finally if the price stays constant keep your shares. We implemented a chrometophobic model (chrimata in greek means money, phobia means fear, hence fear of loss of money) in order to "simulate" the behavior of an inexperienced investor but also to have an idea of the situation of the stockmarket and such guaranteeing that our model does better than simply burrying your money in a share and waiting for it to make more money.

3.2 The Baseline results

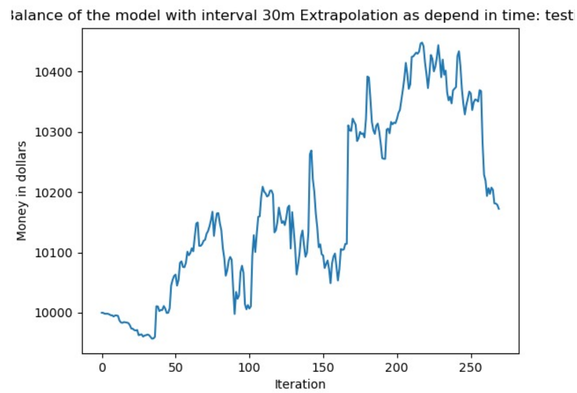
Let us present the results of our baseline over the last 31 days (21 July to August 21) on 30 minutes periods (each number designes the next period over these days, a day is not a full day but a day of trade which starts at 9:30 am and ends at 4 pm) :



As to refute a possible claim that we pruposely failed the model, we broke down the stocks we chose to invest during the testing of the baseline. We chose a diverse profile of stocks, there are shares like google and apple which were mostly profitable over the period of time, Nividia which was mostly unprofitable and the Nasdaq index which will permit us to have an objective perspective on the stock market during the timespan. Let us observe that the baseline model does not fit itself to market loss and just keep investing and possibly hurt its own profits.



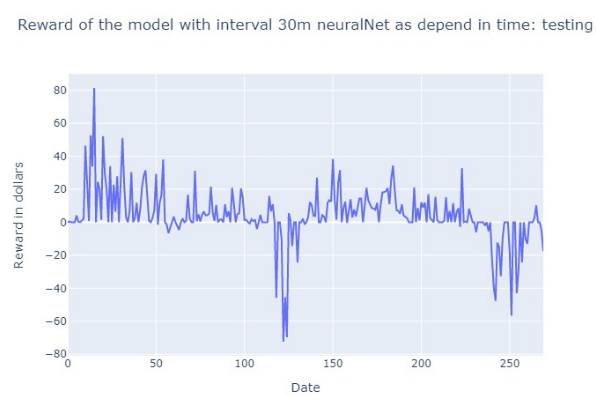
Let us observe the extrapolation makes a colossal amount of trades per period, which might be problematic as we do not always know if the trades will succeed in real life (especially in such a colossal amount as bidders are required for our trades). Now, let us present the real results we care about, the profit made by the model during this timespan.



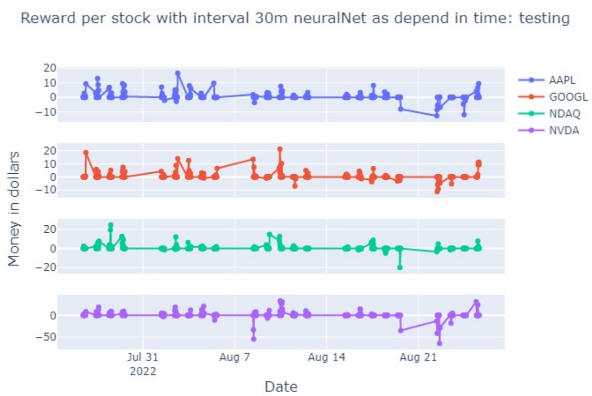
We obtained relatively good results of 3% profit during the last month which is relatively good given the fact our model is a baseline. Nonetheless, due to its high amount of trades per period we can hardly say the baseline model will really yield 3% profit in the real world.

3.3 The Model's results

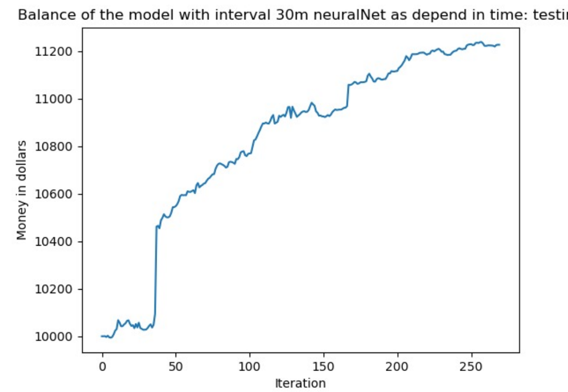
Let us present the results of our Deep Q-Learning over the last 31 days (21 July to August 21) on 30 minutes periods (each number designes the next period over these days, a day is not a full day but a day of trade which starts at 9:30 am and ends at 4 pm):



Let us observe that losing periods are considerably narrower than those of the extrapolation, as the model learns to match itself to the current market situation unlike the extrapolation model who simply keeps trading. As to refute a possible claim that we pruposely made the model win, we have used the same stocks and the same timespan we used as to test the extrapolation baseline.

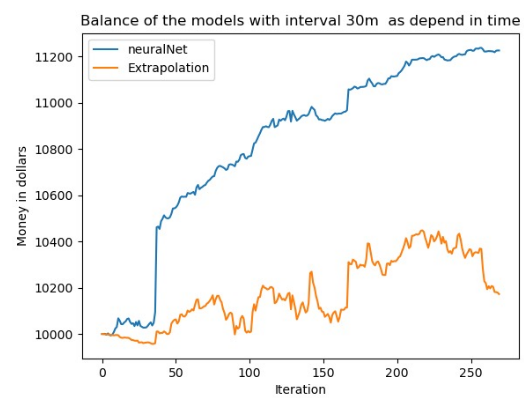


From this graph we can learn what have caused the losing period, descents in Nividia and apple. Because Nividia was doing averagely until it's descent, our model losed faith in the share quite fast, thus stopping the loss right away. Nonetheless, in the case of apple who yielded profit on most of the period, it took the model more time to understand the situation there, it is explainable by the fact it mostly yielded profit and it was aiming to give her another chance. This is a problem we encountered in the learning rate tradeoff, setting a low lr might caused us to slow down in fitting the model to the market, yet, it yielded a more robust model. On the other side, a too high one could follow market trends easily but behaved almost like the extrapolation. Hence, we chose our varying learning rate of adam which yielded optimal results for this problem (yet we still see it is not perfect). Let us observe that the graph is much less dense than the one of the baseline, denoting a considerably lower number of trade requests. Therefore, trades have a higher chance of taking place, making the model a more realistic one. Now, let us present the real results we care about, the profit made by the model during this timespan.



We obtained excellent results of 11.2% profit during the last month which is considerably better than the baseline model. Given the low number of trades it tries to make per period, we can see that this result is a more realistic one, having a relatively high chance to yield close results in the real world.

3.4 Result Analysis



As we have previously said, our Deep Q learning model is much more realistic than the baseline model as it requests a considerably lower amount of trades per period. Firstly, given the fact that our capital of trade is relatively low (10k$ on the S&P 500 is not even a grain of sand) and also given the fact that the market we invested in are of a volume order of 6 (millions of traded share per minute). Additionally, from the graph we

see that almost at every given period, our model yields more profit than the baseline model. A problem we would address in the future is the one of the learning rate which struggles to adjust itself to stocks that yielded profit for a period of time and then experience losses. Nonetheless, one can safely conclude that the Deep Q learning model did mostly good, yielding a respectable realistic profit of 11% for a month, which if taken into yearly account, almost triples the capital every year on average. This result is quite close to the extreme lower bound of profits made by other learning models of companies in the world of High Frequency Trading which range between 37% per quarter (due to a major loss of one of those companies in one of the stocks they invested in) to 800% per quarter [Baron, Matthew, Brogaard] (ours would scrape the 35%). Therefore, we conclude this project quite satisfied with our results for the moment, considering it was a success for us.

4 Bibliography

- Aggarwal, Rajesh K., and Guojun Wu. "Stock market manipulation-theory and evidence." In AFA 2004 San Diego Meetings. 2003.
- Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko. "The trading profits of high frequency traders." Unpublished manuscript.
- Chong, Terence Tai-Leung, and Wing-Kam Ng. "Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30." Applied Economics Letters 15, no. 14 (2008): 1111-1114.
- Gurrib, Ikhlaas. "Performance of the Average Directional Index as a market timing tool for the most actively traded USD based currency pairs." Banks and Bank Systems 13, no. 3 (2018): 58-70.
- Ellaji, C. H., P. Jayasri, C. Pradeepthi, and G. Sreehitha. "AI-based approaches for profitable investment and trading in stock market." Materials Today: Proceedings (2021).
- Jang, Beakcheol, Myeonghwi Kim, Gaspard Harerimana, and Jong Wook Kim. "Q-learning algorithms: A comprehensive classification and applications." IEEE access 7 (2019): 133653-133667.
- Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves et al. "Human-level control through deep reinforcement learning." nature 518, no. 7540 (2015): 529-533.
- Morck, Randall, Andrei Shleifer, Robert W. Vishny, Matthew Shapiro, and James M. Poterba. "The stock market and investment: is the market a sideshow?." Brookings papers on economic Activity 1990, no. 2 (1990): 157-215.
- Shah, Pinkal. "An Empirical Study on Options Trading Strategy Using ‘Commodity Channel Index’For NSE’s Nifty Options in India." In Proceedings of 10th International Conference on Digital Strategies for Organizational Success. 2019.
- Watkins, Christopher JCH, and Peter Dayan. "Q-learning." Machine learning 8, no. 3 (1992): 279-292.
- Zhang, Qingchen, Man Lin, Laurence T. Yang, Zhikui Chen, and Peng Li. "Energy-efficient scheduling for real-time systems based on deep Q-learning model." IEEE transactions on sustainable computing 4, no. 1 (2017): 132-141.