

Monday, 19 of December 2022

RL (ML) in Financial Ranch.

Agent is busy all the time.

Tony (YIT19488399)

UNIVERSITY OF ROEHAMPTON, LONDON, UK

2022.23 MACHINE LEARNING - CMP020X303A

MODULE TUTOR: DR. YUANLIN GU

#####



FIG. 1

#####

Who has the power? Banks, insurances, governments, buffet of lawyers, established families and salaried workers compete in the business of making money. For sure many other people engage in the “environment”,

This academic essay is about applying Reinforcement Learning technology to financial markets that have place virtually using Internet and network communication. Particularly, it is focused on CME group, who started to think about electronic trading platforms in 1987; in 1992 the first electronic trading platforms was developed for future contracts [1], and in 2006 – 2007 electronic trading start to booming in agricultural futures market in Chicago (USA) [2]. The initial idea was to improve local agriculture and livestock and today is possible to trade futures about cryptocurrencies like Bitcoin and Ether.

Reinforcement learning (RL) is the most recent machine learning (ML) technique that started as a concept in 1960 in engineering literature with Richard Bellman (26 August 1920 to 19 March 1984, USA) using theory from Hamilton and Jacobi, both lived at the beginning of the nineteenth century in Europe. From 1980 RL experience a revival and today is common to find academic and industry publication and papers each month and /or week [3]. The famous Bellman equation follows, and it has many different “views”. Three of these mathematical expressions are:

Fist “view” (Fig. 2):

$$Q^*(s, a) = r + \gamma \max_{a'} (Q^*(s', a'))$$

The value of taking action (a) in state (s)
Reward
Discount factor
The maximum possible value of the next state (s') given the best action (a')

Second “view” (Fig. 3):

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha \left(R_{t+1} + \gamma \max_a Q_t(s_{t+1}, a) - Q_t(s_t, a_t) \right)$$

New State
Old State
Reward

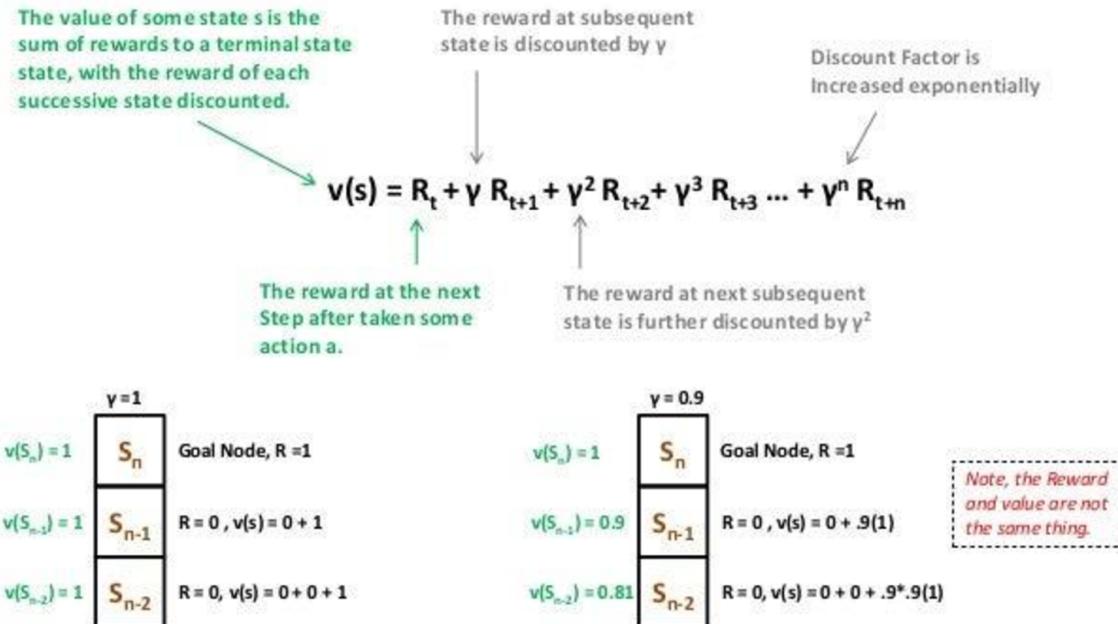
Learning Rate
Discount Factor

~~~~~

Third “view” (Fig. 4):

# Bellman Equation

- Principle of the Bellman Equation



~~~~~

Basically, RL is founded on the Bellman equation and it consist of optimization applied to dynamic programming [4]. Bellman equation looks for maximizing a long-term total reward, but each decision called action is based on current state/action and future state/action occurring in shorter time. Exploration and exploitation need to have an equilibrium and greedy behavior is not the best. Randomness is used in each decision by the agent, and this is called “epsilon-greedy” algorithm, learning and action selection. Rewards, punishments, and no reward/punishment are counted for each state/action. What makes this calculation great is the discount factor, because it is possible to find the best solution and/or path. Gamma (discount factor) weight the present reward with future rewards and its value is between 0 to 1, lower values for short term rewards preference and higher values for long term rewards preferences. Because computer is fast enough to calculate all paths/solutions, we can look for the best one in a huge number of possible solutions, sometimes unlimited [5]. Deep RL is the next step to solve this situation with two hidden layers used by convenience [6]. Agent is the chunk of software responsible of finding best solutions from the available founded. To better understand the idea, an agent can be seen as a rat in a laboratory maze with cheese in few places and many obstacles and traps around with some free space to allow the agent (rat) to move.

For trading futures in Chicago electronically, RL is applied to the environment where the price is the last, top, and most clear reference. Futures about derivatives like oil, gold, finance products, coffee or currencies futures can be found [7]. It is impossible to catch every detail of the environment in this extensive market size inside the “ranch” (stakeholders directly involved in the business and/or negotiations), and outside. Factors like weather can affect the quantity and quality of oranges from Florida [8], the location of final delivery and the document paper signed can affect the product because can be involved three different currencies. These are some examples easy to understand how big and extensive the environment is. One more important factor to consider in the environment is the presence of speculation and stakeholder that only are in these environment for make money/profits [9]. These people and firms accept risks in order to make gains. One more example of these “players” (speculators) is the property market where prices are inflated and you can have a look in the current time in London, UK, and generally in all the UK. Reward is money and with money life is much easier in the present capitalized and globalized world.

Focusing on electronic trading, previous ML techniques were applied like random forest, naïve Bayesian classifier, support vector machine, k-nearest neighbour and RNN-LSTM [10], all supervised ML techniques except LSTM who is unsupervised, but RNN is supervised deep learning. All techniques were good in the beginning and useful for the mastermind and father, but today trend is using deep reinforcement learning (DRL) were RL plus neural network (NN) are used. Deep means where two or more layers of NN are implemented in the model. Very important aspect of future trading is the time series condition of the market. One more ingredient to include is the “cloud”. Today is very popular “be on the cloud”. This means all your business, information, behaviour, and everything is managed and processed by big technological companies. In the Q3 2022, the worldwide infrastructure of the cloud was dominated by AWS (Amazon), Azure (Microsoft), and Google Cloud (google). These three are US corporations and I believe the future is dominate the market...

In conclusion, DRL and RL can be applied anywhere. From self- driving cars to vacuum cleaners. It started with playing games like StarCraft [11] and other much simpler video games, continue with autopilots or real life robots working in factories or transporting equipment and utensils. Those are few examples that come to my mind, but I am 100% sure that many other fabrications and investments are in progress. Like student of computer science, I am not sure about the legality, lawfulness, and ethics of this technology. It is very difficult to track and collect all not lawful situations that are happening in China from factories like Apple who distribute smartphones iPhone, for example. I mean that is possible to develop an RL robot whit evil purpose in China where resources are available for US customer, but if the US customer do this RL evil robot in the US, he or she will get legal problems and maybe it will not work because of the ethics, believes and standards of the territory.

Finally, change the reward after “bad” action is performed is the way of teaching the agent to be evil and maybe destructive. The policy becomes dangerous for the humanity! Please do not do it!!

Fig. 5:



References:

- [1] 'CME group'. [Online]. Available: https://en.wikipedia.org/wiki/CME_Group
- [2] 'University of Lincoln, electronic trading history.' [Online]. Available: <https://agecon.unl.edu/cornhusker-economics/2015/how-technology-has-changed-futures-markets-part-one>
- [3] 'RL history', [Online]. Available: <http://incompleteideas.net/book/ebook/node12.html>
- [4] 'Summary of the Bellman equation'. [Online]. Available: https://en.wikipedia.org/wiki/Bellman_equation
- [5] 'summary of the bellman equation'. <https://www.geeksforgeeks.org/bellman-equation/>
- [6] 'deep reinforcement learning number of layers'. <https://www.linkedin.com/pulse/choosing-number-hidden-layers-neurons-neural-networks-sachdev/>
- [7] 'cme products'. [Online]. Available: <https://www.cmegroup.com/markets/products.html#sortAsc&sortField>
- [8] 'Orange juice price'. [Online]. Available: <https://www.boerse-frankfurt.de/commodity/orangensaftpreis>
- [9] 'The Pros and Cons of Speculation in Commodity Futures'. [Online]. Available: <https://www.thebalancecommodity.com/the-pros-and-cons-of-speculation-in-commodity-futures-808944>
- [10] 'The best ML algorithms for stock price prediction'. [Online]. Available: <https://www.itransition.com/machine-learning/stock-prediction>
- [11] *StarCraft II*. [Online Video]. Available: <https://starcraft2.com/en-gb/>

Bibliography (academic paper):

- I. <https://ieeexplore.ieee.org/document/9877940>
- II. <https://doi.org/10.1007/s11063-022-11019-w>
- III. <https://doi.org/10.1155/2022/4698656>
- IV. https://datascience.columbia.edu/wp-content/uploads/2020/12/34_JPMorgan_Reinforcement-Learning-for-Trading.pdf
- V. <https://arxiv.org/abs/2208.07165>
- VI. <https://core.ac.uk/download/pdf/322327703.pdf>

Bibliography academic paper with Zotero:

[I]

'Application of Deep Reinforcement Learning in Stock Trading ... https://core.ac.uk › download › pdf PDF by Y Li · Cited by 57 — in the stock trading decisions and stock price prediction, the reliability and ... Keywords Reinforcement Learning · Financial Strategy · Deep Q Learning. 19 pages', [Online]. Available: <https://core.ac.uk/download/pdf/322327703.pdf>

[II]

'Deep Reinforcement Learning Approach for Trading Automation in The Stock Market', [Online]. Available: <https://arxiv.org/abs/2208.07165>

[III]

'Reinforcement Learning for Trading https://datascience.columbia.edu › 2020/12 › 34_J... PDF - How would you teach a 10 year old to trade stocks? - How can we transfer this logic to artificial intelligence? → Smart trading is nuanced but situational ... 18 pages', [Online]. Available: https://datascience.columbia.edu/wp-content/uploads/2020/12/34_JPMorgan_Reinforcement-Learning-for-Trading.pdf

[IV]

Y. Li, P. Liu, and Z. Wang, 'Stock Trading Strategies Based on Deep Reinforcement Learning', *Scientific Programming*, vol. 2022, p. 4698656, Mar. 2022, doi: [10.1155/2022/4698656](https://doi.org/10.1155/2022/4698656).

[V]

J. Wang, F. Jing, and M. He, 'Stock Trading Strategy of Reinforcement Learning Driven by Turning Point Classification', *Neural Processing Letters*, Oct. 2022, doi: [10.1007/s11063-022-11019-w](https://doi.org/10.1007/s11063-022-11019-w).

[VI]

T. Kabbani and E. Duman, "Deep Reinforcement Learning Approach for Trading Automation in the Stock Market," in IEEE Access, vol. 10, pp. 93564-93574, 2022, doi: 10.1109/ACCESS.2022.3203697., [Online]. Available: <https://ieeexplore.ieee.org/document/9877940>

Images references:

Fig. 1:

- Top Secret. No source can be provided.

Fig. 2:

- https://duckduckgo.com/?t=ffab&q=bellman+equation&iax=images&ia=images&iai=http%3A%2F%2Fmiro.medium.com%2Fmax%2F6562%2F1*BeOPi7fM8iLAagZbPPsSwQ.png

Fig. 3:

- https://duckduckgo.com/?t=ffab&q=bellman+equation&iax=images&ia=images&iai=http%3A%2F%2Fcdn-images-1.medium.com%2Fmax%2F800%2F1*iRHNLmHdthueArHD80CBLq.png
- <https://ailephant.com/overview-deep-reinforcement-learning/>

Fig. 4:

- <https://duckduckgo.com/?t=ffab&q=bellman+equation&iax=images&ia=images&iai=http%3A%2F%2Fimage.slidesharecdn.com%2Fai-bellmanintro-170909235039%2F95%2Fai-introduction-to-bellman-equations-6-638.jpg%3Fcb%3D1505001199>
- <https://www.slideshare.net/AndrewFerlitsch/ai-introduction-to-bellman-equations>

Fig. 5:

- <https://bloody-disgusting.com/movie/3581481/australian-version-new-dark-fate-trailer-drops-beautiful-f-bomb-includes-new-footage/>

#####END#####