Part 1: Research Paper Review

"A Comprehensive Review of the research paper Application of Deep Reinforcement Learning to Algorithmic Trading"

Zia Ullah Khan

Registration number: 2101142

CF969-7-SP-Big Data for Computational Finance, University of Essex, Colchester Campus.

Keywords: Artificial intelligence, deep reinforcement learning, algorithmic trading, trading policy.

1. General description of the article

This review article entitled "An Application of Deep Reinforcement Learning to Algorithmic Trading" has demonstrated a novel approach for determination of optimum trading point during business endeavors in the stock market using deep reinforcement learning (DRL) agent. The Sharpe ratio indicator is maximized by the new DRL approach, by stimulated the popular DQN algorithm, which indicated performance on a wide range of stock market. Further, the training data if generated from a historical stock market via artificial trajectories for limited set and used reinforcement learning (RL) agent. Thereby, a novel assessment methodology for performance of trading strategies is established.

2. Background

Artificial Intelligence owing to their key element of deep learning techniques specifically deep neural network (DNN) has gained massive successes in the past few years for application in image classification, speech recognition and natural language processing. In parallel the DRL techniques, which can solve complex sequential decision-making problems are used to innovates and improve activities in finance sectors such as the trading, investment, risk management, portfolio management, fraud detection and financial advising. Several industries utilize algorithmic approach for quantitively trading assessment using computers, set of mathematical rules, to make complex algorithmic decisions based on historical data. The approach followed in this paper is DQN algorithm. The review article is structured as an (i) introduction to algorithmic trading (AT), (ii) formalizing AT and relating it with reinforcement learning approach, (iii) explanation of Trading Deep Q-Network algorithm (TDQN) trading strategy design based on DRL notions, and (iv) methodology to assess the performance of different trading strategies.

The Fintech industry used assessment tools, but their research results are not publicly available due to huge amount of stake[1], examples are the trend following and mean reversion strategies Chan (2009)[2], Chan (2013)[3]. Also, there is no well-established framework, which can fairly compare trading strategies. For discovering new investment policies, Moody and Saffell (2001)[4] introduced a recurrent RL algorithm, without the need to build forecasting models. Dempster and Leemans (2006)[5], used adaptive RL to trade in foreign exchange markets and the trading costs, which are not well defined. One of the major drawbacks in scientific community is the lack of scientific approach with the excellent communication results but lack of objective criticism. Ioannidis (2005)[6], has mentioned that most of the findings of the published contents which are sensitive are false, when the subject directly in touch with trading algorithm, especially in financial sciences. Therefore, this paper works on presentation of an unbiased scientific evaluation of the deep reinforcement-learning algorithm.

3. Fundamentals of Algorithmics described in this article

3.1.Algorithmic trading problem formalization

Algorithmic trading or qualitative trading, automatically making trading decisions based on a set of mathematical rules computed by a machine, which is very beneficial to market for the significant improvement in liquidity, Hendershott et al. (2011)[7].The DRL algorithms used in this paper is applicable to multiple market with extension to other market planned in the future. such as, the FOREX trading and the rise of new cryptocurrencies like Bitcoin offer new offers new interesting possibilities to apply algorithmic trading strategies. A discretization operation for continuous timeline is performed for the algorithmic trading problem, in which timeline is high number of discrete trading time t for constant duration Δt such as $t + \Delta t$ ($t - \Delta t$), this Δt is closely linked to the trading frequency by the trading agent, which is chosen as small as possible because the trading frequency $\Delta t = 1$ is limited. Subsequently, a trading strategy is performed which is a programmed policy can be stochastic or deterministic and outputs a trading action according to the available information to the agent at specific time.

3.2.Reinforcement learning problem formalization

RL techniques focuses on policies which maximizing an optimality standard and it depends on instant rewards over a time. Optimality criteria, which is the discounted sum of rewards over an unlimited time is the expected outcome, which ranges from gamma [0, 1], when value is 0 it only considers reward at present and ignore future rewards. RL agent is directly proportional to the discount factor when $y = 1$, RL agent considers rewards on equal basis. This parameter can be tuned as per circumstances. Major challenge in algorithmic trading is not observing the environment correctly. The information RL agent get contains current trading position, available cash, agent's shares quantity) and other information such as OHLCV (Open-High-Low-Close-Volume) [8]. Time information are date, weekday, month, week, year, time. Technical indicators of the stock market, are moving average convergence divergences (MACD), average directional index (ADX), and relative strength index (RSI) etc. Other macroeconomic factors such as interest rates or exchange rates are also useful in forecasting markets dynamics. Apart from that, other information can be extracted, which can be done on social media, newspapers and specific journals.

3.3.Deep reinforcement learning algorithm design

The DQN is a learning control procedure from high dimensional inputs based on the learning of calculation of the state-action value represented by DNN. The training part is based on artificial trajectories from a stock market historical OHLCV data on daily basis. The DQN algorithm is widely adopted for simulations. These are replacing the convolutional neural network (CNN) with a classical feed forward (DNN). Similarly double DQN, ADAM optimizer, Huber loss, Gradient clipping, Xavier initialization, batch normalization layers, regularization techniques such as Dropout, L2 regularization and early stopping. Data augmentation techniques are applied to data as the key challenge to AT is the limited data availability and quality of data. Therefore, augmentation techniques such as signal filtering, signal shifting, and artificial noise addition are applied so that to generate artificial data, which helps, improve the financial phenomena.

4. Evaluation

This research paper describes the working of TDQN, a deep reinforcement learning application to algorithmic trading problem. The strategy has achieved outstanding outcomes surpassing the benchmark trading strategies.

- Versatility and robustness to diverse trading costs
- Defining explicit rules appropriate according to financial markets.
- Performance could be improved more by reproducibility and generalization.

Furthermore, following strategies were adopted to assess the weaknesses and strengths of the TDQN algorithm.

- Buy and hold (B&H)
- Sell and hold(S&H)
- Trend following with moving averages (TF)
- Mean reversion with moving averages (MR)

4.1.Benchmark trading strategies

The paper took 30 stocks and all the hyperparameters have been kept constant for training of different algorithms. 8 years data have been taken to analyze the trends; validation set is also taken to tune hyperparameters of training dataset and test dataset.

a) Good results: Apple stock

TDQN algorithm has a significant variance. The algorithms is showing good results, almost 50 iterations of both training and test sets have been shown in the Fig. 1**.** The distribution of the daily returns is continuously fluctuating.

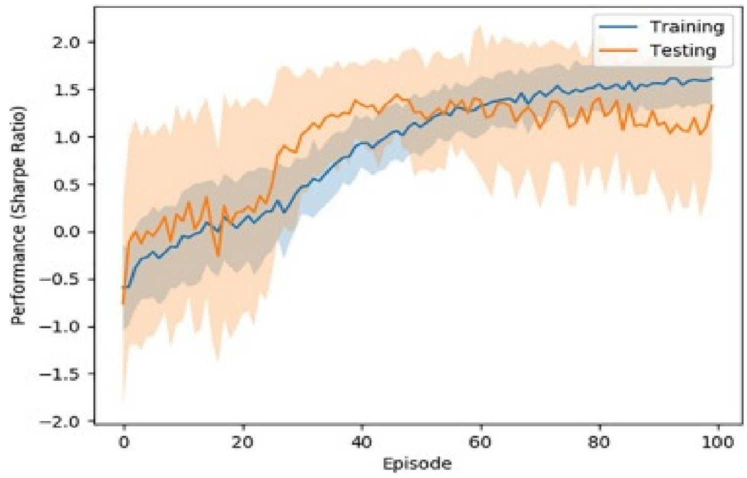

Fig. 1. TDQN algorithm expected performance for the Apple stock.

b) Mitigated results: TESLA stock

TESLA stock performance is analyzed, it has been observed that the tesla stock is quite difficult to trade in due to high volatility. Although TDQN algorithm has achieved high Sharpe ratio, but with high-risk trading strategy. Therefore, there is a high variance, and the graph is illustrating overfitting as depicted in Fig.2 , this is the key limitation of the algorithm.

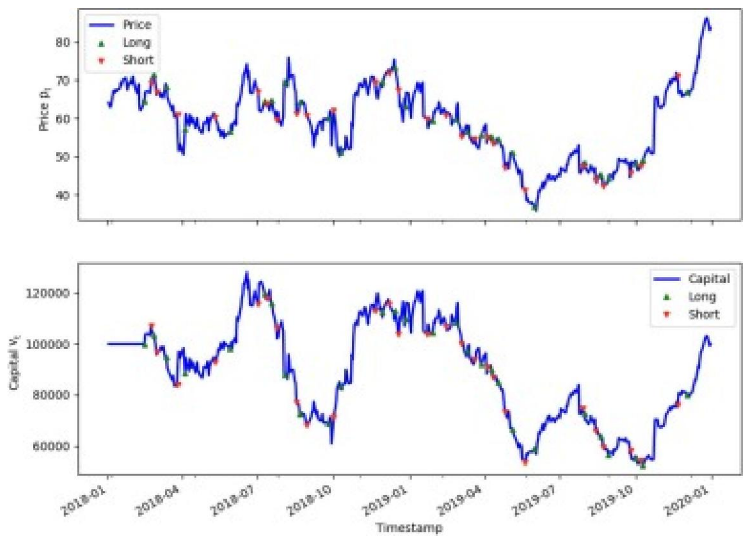

Fig. 2. TDQN algorithm execution for the Tesla stock (test set).

c) Global results

Performance of the TDQN is very alike to passive trading strategies such as B&H and S&H for different stocks. When there is high, uncertainty in active trading, DRL strategy tends towards a passive strategy. It can be observed that TDQN does not purely rely on one strategy at a time. Hence, one of the main benefits of DRL trading is its versatility and its ability to handle varying markets having different characteristics.

4.2. Limitations

In this research paper, an entirely different environment is studied with the algorithmic trading problem. Obviously, multiple challenges were faced during the research around the TDQN algorithm, the major ones being summarized hereafter.

- Poor observability of the trading environment, which limits the performance of the algorithm.
- The TDQN algorithm should have reflection of the past, which is one of core challenge as distribution.
- Overfitting should be properly handled to get reliable trading strategy. More research is required on this issue in DRL techniques to over fit.

5. Opportunities and Future outlook

Opportunities for improvement are listed below:

- To upgrade the DRL solution by applying LSTM layers in deep neural network
- Compare TDQN algorithm with Policy optimization DRL algorithms for instance proximal policy optimization.
- Formalization of the algorithmic trading problem into a reinforcement learning one by extending observation space.
- The gap between RL objective and the Sharpe ratio maximization can be narrowed down by advance applying RL reward engineering.
- The notion of risk distributions should be considered more than expected values in the TDQN algorithm.

6. References

[1] Y. Li, arXiv preprint arXiv:1701.07274, (2017).

[2] E.P. Chan, Quantitative trading: how to build your own algorithmic trading business, John Wiley & Sons2021.

[3] E. Chan, Algorithmic trading: winning strategies and their rationale, John Wiley & Sons2013.

[4] J. Moody, M. Saffell, IEEE transactions on neural Networks, 12 (2001) 875-889.

[5] M.A. Dempster, V. Leemans, Expert Systems with Applications, 30 (2006) 543-552.

[6] J.P. Ioannidis, PLoS medicine, 2 (2005) e124.

[7] T. Hendershott, C.M. Jones, A.J. Menkveld, The Journal of finance, 66 (2011) 1-33.

[8] T. Théate, D. Ernst, Expert Systems with Applications, 173 (2021) 114632.