Review of Reinforcement Learning Applied in High Frequency Trading

1. **Introduction:**

High-frequency trading (HFT) is a particular style of algorithmic trading that transact large numbers of orders within seconds, which can provide liquidity by simultaneously quoting bid and ask price on an asset and reap profits off the spread. Currently, it is estimated that HFT occupies about 50% of equity trading volumes in the U.S. (see HFT definition in Nasdaq) *[1]*, with this figure continuing to rise. HFT is primarily characterized by the optimization of speed that is mere technological challenge, and the sophistication of strategies, namely, the algorithms that decide when to trade, at what price and what quantity, which is a much more significant challenge.

The non-stationary dynamics of financial markets can be taken as stochastic environments in which a sequence of dynamic decisions must be made through trial-and-error, it is intractable for mathematical modelling due to full complexity and low SNR (signal-to-noise ratio), especially for HFT. Given the characteristics, reinforcement learning (RL) is believed to be a very potential solution for HFT problems.

RL is a very general and popular paradigm for sequential decision-making problems, to be specific, it applies state-based models that take actions in an unknown and uncertain environment to maximize cumulative reward (or minimize related costs), while the best long-term strategy may involve short-term sacrifices, therefore, the agent in RL needs to find an optimal balance between exploitation (make the best decision given current acquired knowledge) and exploration (gather more information to make the best overall decision) *[2]*. Furthermore, when augmented with deep learning, deep reinforcement learning (DRL), it can scale to problems that were previously out of reach for RL such as searching in large state and action spaces because of the curve of dimensionality. At present, there are a number of successful applications of RL in Gaming such as AlphaGo *[3]*, AlphaStar *[4]*, etc., which show the effectiveness and promise of RL. However, the research of RL applied in HFT is extremely nascent, only a few academic works are conducted in this area.

2. **Background and Problem Statement:**

When conducting research about HFT, the time range during which certain volumes of orders are traded, is always discretized into smaller uniform time-step, for example, 2 minutes can be segmented as 120 time-steps with each 1 second.

Obviously, the trading modes can be categorized as 2 types:

(a) multiple-assets (MA) or asset portfolios (AP),

(b) single asset (SA).

In fact, MA/AP can be decomposed into several additive independent SAs after portfolio allocation. Most of published academic works about applying RL into HFT focus on trading single asset.

When trading single asset, there are 3 styles:

(a) Trading single security (or fixed sizes of securities) by multiple directions (buy, sell, and hold), namely, frequently change positions (buy, sell or hold) of the single security given the time range to achieve as high wealth as possible,

(b) Trading multiple securities by single direction (buy or sell), that is, for example, to buy (sell) exactly V shares of stock/futures given the time range, while minimize the expenditure (maximize the revenue),

(c) Trading multiple securities by multiple directions.

For simplification and convenience, current research generally concentrates on the methods (a) *[5][6][7][8][9]* and (b) *[10]*, which are much easier when compared with method (c).

Simultaneously, transaction costs are considered in some papers, while slippage (or implementation shortfall) and market impact ("the impact of its own trades on the market" *[9]* are often too hard to simulate except trading on real markets such as Nasdaq and NYSE.

Undoubtedly, RL as one of the main sub-domains of machine learning, when applying it into HFT, feature engineering (or feature selection) is necessary and significant. Early papers frequently use profit/wealth, Sharpe ratio and common technical indicators (or macrostructure data) *[5][6][7][11]*, in contrast, recent ones constantly introduce limit order book (or microstructure data) *[9][10][12]* to design states, actions and value functions (or rewards) of RL.

While there are various RL models and their variants, for instance, Q-learning and Recurrent Reinforcement Learning (RRL) are taken to train agents in some early research, and experiments show RRL, the RL variant that "requires the use of reinforcement versions of standard recurrent learning algorithms" because of temporal dependence on system state, performs better in trading than Q-learning, the classic RL model that "can suffer from the curse of dimensionality and is more difficult to use" *[7][13]*. Owing to the rise of deep learning, deep reinforcement learning (DRL) that is the combination of deep learning and reinforcement learning can be much more powerful than classic RL, and is applied into related trading research recently such as predicting the stock price movement *[12]*, and designing the trading agents *[9][14]*.

However, in several papers, apart from RL/DRL, other algorithms are used as a helper to design or train trading agents, for example, utilizing genetic algorithm to choose better features for training *[6]*, employing imitation learning techniques to balance between exploration and exploitation *[15]*, and applying LSTM, CNN or their variants to predict price trend *[16][17][18]*, etc.

Based on cost considerations, the majority of academic works are conducted on historical data of financial market such as foreign exchange, stock, futures and so on. But in a few research, realistic simulation of financial market data (limit order book) is developed *[14][19]*, which can provide more flexibility and may be a new trend in future research.

Noticeably, progress made so far with RL/DRL applied into HFT proves the prospects and feasibility, nevertheless, the statistical improvements in performance brought by RL/DRL do not demonstrate decent profitability level, especially the lack of experimental testing in real-time, online platforms *[20]*, which means there is a long way to go.

3. Future development:

Basically, RL/DRL is a vibrant and important sub-field of machine learning that tries to computationally generate and efficiently refine models from statistics and algorithms, which is naturally a very candidate for application to problems arising in HFT *[10]*. Those applications are consistent with the tenets of technical analysis, but contradictory against efficient market hypothesis *[6]*, hence, more understanding should be made about whether generation of alpha is achievable from technical analysis. Meanwhile, RL/DRL is an open research direction, whose strengths and limitations are still being discussed and explored. Therefore, development and generalization of RL/DRL application across different financial markets are arduous. Based on the above review, the future directions of development will focus on:

1) Feature engineering:
   Try to filter the low SNR data or smoothen the noisy signals of HFT data, then both macrostructure and microstructure indicators can be combined with a balance to create features on the one hand, on the other hand, features should be kept informative and simplified to suit the computational capability. Actually, Japanese candlesticks is a meaningful trial to create feature for HFT use *[21]*, which can be a good sample.
2) Adaptive models:
   RL is a general framework with numeric algorithms, after combined with deep learning, more algorithms/models are introduced, which can be divided into 3 main groups: value-based such as Q-learning and deep Q-networks, policy-based such as Proximal Policy Optimization *[14]*, and actor-critic-based that integrates the advantages of value-based and policy-base. According to their characteristics, experiments can be conducted by adapting them or their variants into HFT.
3) Useful extra-helper:
   Some extra parts can be introduced as a useful helper, e.g., machine learning and deep learning techniques can be taken to predict price movements

which is an informative signal for feature engineering, or some adaptive systems that can manage risk, utility and automation degree above trading agents, and suchlike.

4) Datasets / environment:
There are 3 possible environments to test the trading agents trained using RL/DRL: firstly, the historical data of financial markets can be used to fundamentally evaluate trading agents, while the slippage (implementation shortfall) and market impact are always ignored in the environment formed by historical data. Secondly, realistic simulation of financial markets can be developed to assess more comprehensively the performance of the trading agents because it can reproduce many relatively stochastic and dynamic scenes. Thirdly, testing the trading agents in real financial markets that can judge accurately the performance of trained agents but it may cost much more.

5) Framework development:
Up to now, almost all of the academic research about this area is implemented in a quietly divergent way, which makes it not possible, at least not convenient, to compare the performance of trading agents created by different researchers. For that reason, the need to create a generic framework which not only makes it easy to develop trading agents but also provide comparative place, and FinRL *[22]* is a pioneer but not complete enough.

4. Reference:

[1] Breckenfelder, J. (2019). Competition among high-frequency traders, and market quality.

[2] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

[3] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. nature, 529(7587), 484-489.

[4] Vinyals, O., Babuschkin, I., Chung, J., Mathieu, M., Jaderberg, M., Czarnecki, W. M., ... & Silver, D. (2019). Alphastar: Mastering the real-time strategy game starcraft ii. DeepMind blog, 2.

[5] Moody, J. E., & Saffell, M. (1999). Reinforcement learning for trading. Advances in Neural Information Processing Systems, 917-923.

[6] Dempster, M. A. H., & Romahi, Y. S. (2002, August). Intraday FX trading: An evolutionary reinforcement learning approach. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 347-358). Springer, Berlin, Heidelberg.

[7] Gold, C. (2003, March). FX trading via recurrent reinforcement learning. In 2003 IEEE International Conference on Computational Intelligence for Financial Engineering, 2003. Proceedings. (pp. 363-370). IEEE.

[8] Dempster, M. A., & Leemans, V. (2006). An automated FX trading system using

adaptive reinforcement learning. Expert Systems with Applications, 30(3), 543-552.
[9] Briola, A., Turiel, J., Marcaccioli, R., & Aste, T. (2021). Deep Reinforcement Learning for Active High Frequency Trading. arXiv preprint arXiv:2101.07107.
[10] Kearns, M., & Nevmyvaka, Y. (2013). Machine learning for market microstructure and high frequency trading. High Frequency Trading: New Realities for Traders, Markets, and Regulators.
[11] Zhang, J., & Maringer, D. (2013, July). Indicator selection for daily equity trading with recurrent reinforcement learning. In Proceedings of the 15th annual conference companion on Genetic and evolutionary computation (pp. 1757-1758).
[12] Ganesh, P., & Rakheja, P. (2018). Deep reinforcement learning in high frequency trading. arXiv preprint arXiv:1809.01506.
[13] Moody, J., & Wu, L. (1997, March). Optimization of trading systems and portfolios. In Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr) (pp. 300-307). IEEE.
[14] Kumar, P. (2020, May). Deep reinforcement learning for market making. In Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (pp. 1892-1894).
[15] Liu, Y., Liu, Q., Zhao, H., Pan, Z., & Liu, C. (2020, April). Adaptive quantitative trading: An imitative deep reinforcement learning approach. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 02, pp. 2128-2135).
[16] Rundo, F. (2019). Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems. Applied Sciences, 9(20), 4460.
[17] Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2017, July). Forecasting stock prices from the limit order book using convolutional neural networks. In 2017 IEEE 19th Conference on Business Informatics (CBI) (Vol. 1, pp. 7-12). IEEE.
[18] Zhang, Z., Zohren, S., & Roberts, S. (2018). BDLOB: Bayesian deep convolutional neural networks for limit order books. arXiv preprint arXiv:1811.10041.
[19] Byrd, D., Hybinette, M., & Balch, T. H. (2019). Abides: Towards high-fidelity market simulation for ai research. arXiv preprint arXiv:1904.12066.
[20] Pricope, T. V. (2021). Deep Reinforcement Learning in Quantitative Algorithmic Trading: A Review. arXiv preprint arXiv:2106.00123.
[21] Gabrielsson, P., & Johansson, U. (2015, December). High-frequency equity index futures trading using recurrent reinforcement learning with candlesticks. In 2015 IEEE Symposium Series on Computational Intelligence (pp. 734-741). IEEE.
[22] Liu, X. Y., Yang, H., Gao, J., & Wang, C. D. (2021). FinRL: Deep reinforcement learning framework to automate trading in quantitative finance. arXiv preprint arXiv:2111.09395.