CS6700 : Reinforcement Learning
Project Proposal

Manan Tomar
ED14B023
manan.tomar@gmail.com

Rohan Saphal
ED14B035
rohansaphal@gmail.com

1 Introduction

The academic reinforcement learning [1] research community has largely stayed away from the financial markets. Possible reasons are, (1) the finance industry has a bad reputation, (2) the problem does not seem interesting from a research perspective, or (3) because data is difficult and expensive to obtain. In our project, we hope to explore training Reinforcement Learning agents to trade in the financial markets and how this can be an extremely interesting research problem.

Reinforcement Learning makes a right fit for applying to financial markets because :

- **End-to-End Optimization** : In the traditional strategy development approach there are several steps before we get to the metric we actually care about. For example, to find a strategy with a maximum drawdown, it is needed to train a supervised model, come up with a rule-based policy using the model, backtest the policy and optimize its hyper-parameters, and finally assess its performance through simulation. Reinforcement Learning allows for end-to-end optimization and maximizes (potentially delayed) rewards.
- **Learned Policies** : Instead of needing to hand-code a rule-based policy, Reinforcement Learning directly learns a policy. There's no need to specify rules and thresholds such as "buy when you are more than 75 percent sure that the market will move up". That's inherent in the RL policy, which optimizes for the metric. Also because the policy can be parameterized by a complex model, such as a deep neural network, it is possible to learn policies that are more complex and powerful than any rules a human trader could possibly come up with.
- **Learning in a Simulator** Since Reinforcement Learning agents are generally trained in a simulation, and the simulation can be as complex as desired, taking into account latencies, liquidity and fees, we always have the option to not test the policy in the real market until some performance level is maintained while testing in the simulator. The simulator can in turn be tweaked to match the market behavior more precisely.

2 Proposal

We mainly wish to benchmark the performance of state-of-the-art reinforcement learning methods ([2], [3] [4], [5], [6] to name a few) using an open source market simulator and analyze how well current methods fair in such complicated scenarios and under what conditions (for ex. in high frequency trading). The final takeaway from this project will be a comprehensive analysis on why or why not current algorithms fit the use case of designing trading strategies.

3 Simulator

For our experiments, we wish to use a recently released open source trading platform called **Sairen**. This simulator is compatible with OpenAI gym like environments and thus makes it a universal choice, allowing us to cover most of the prominent algorithms. Sairen provides access to real world live financial data to test machine learning algorithms on. It is based on intraday trading, done at the frequency of minutes or seconds instead of nanoseconds. The observations are the current market data, actions are whether to buy and sell, and the rewards are the profit or loss incurred.

References

[1] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*, vol. 135. MIT press Cambridge, 1998.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, p. 529, 2015.

[3] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[5] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International Conference on Machine Learning*, pp. 1889–1897, 2015.

[6] M. G. Bellemare, W. Dabney, and R. Munos, "A distributional perspective on reinforcement learning," *arXiv preprint arXiv:1707.06887*, 2017.