

CPS824 Project Proposal

Jeff Keith
Computer Science Dept
Ryerson University
Toronto, Canada
jkeith@ryerson.ca

Ian MacPherson
Computer Science Dept
Ryerson University
Toronto, Canada
ian.macpherson@ryerson.ca

Bretton Tan
Computer Science Dept
Ryerson University
Toronto, Canada
bretton.tan@ryerson

Paul Messina
Computer Science Dept
Ryerson University
Toronto, Canada
paul.messina@ryerson.ca

Abstract—In this project we aim to build a stock trading algorithm using reinforcement learning. The dataset and API provided is from the *gym-anytrading* github [1] which compiles OpenAI Gym environments for reinforcement-based trading algorithms. The performance of the agent will be measured against a gain (%) of return over time.

I. MOTIVATION

Reinforcement learning has been shown to perform quite well for a range of tasks such as playing video games [2], flying autonomous helicopters [3], and playing the stock market [4] to name a few. More interestingly, the methods used for the latter application have been mostly shrouded in secrecy by large quantitative analysis hedge funds - known more colloquially as ‘quants’. Some of these funds include D.E. Shaw, Two Sigma, and Renaissance Technologies, which manage a combined ~\$274B USD [5, 6, 7].

The goal of the following project is to learn the core concepts of algorithmic trading using reinforcement learning and gain insight into how such algorithms may be improved to perform better. Future work may include testing reinforcement learning methods against deep learning methods and/or deep reinforcement learning methods for comparison. We will begin with a small goal of building an algorithm to recommend stocks to the investor that gain a percentage above the initial investment over a defined time interval.

II. METHODS

We will use a reinforcement learning algorithm paired with the *gym-anytrading* environment for training and evaluation. The agent will be given a positive reward if a stock chosen increases in value, and a negative reward if the stock chosen decreases in value over a predetermined period of time. The intended outcome is to have the agent provide stock recommendations for an investor to decide on whether or not to purchase the stock.

The stock recommendation may act as a starting point, so that a human may still be able to factor in their own analysis and research. We believe this is a good starting point as to not give the algorithm free-reign over buying and selling when the only tests that have been conducted are in a lab setting. This will allow for more data to be collected to enhance the performance, and strengthen the confidence in the algorithm’s

decision making capabilities.

III. INTENDED EXPERIMENTS

The intended experiments that are planned to be conducted are listed below in no particular order:

- Once the algorithm is complete, test the performance over predefined intervals of time (1-month, 3-months, 1-year from historical market data) and note the results. Once this experiment is complete, we intend to determine if there is an optimal timeline on the recommendation of the stocks. (i.e. is there a strong correlation between the initial recommendation and long-term performance of the stock, or will performance only hold for the short term).
- Compare the results of the reinforcement learning agent to randomly sampled stock picks. This random sample may act as a baseline for the reinforcement learning agent to beat. This additionally will serve to prove that the reinforcement agent is learning, and not simply recommending stocks at random. As mentioned before, the key metric will be % return over time.
- Once tested and trained on a historical dataset, test the algorithm on real-time data to determine whether lab conditions can be replicated in practice.
- Change/tune the value of gamma, to try and find the most optimal gamma value for this application
- Perform test to determine which value function estimation algorithm, Monte Carlo or Temporal Difference, optimizes the bias-variance trade off most effectively

IV. PLANNING AND MILESTONES

Our milestones will consist of the following:

- 1) February 18, 2021: Complete research
 - a) Research best RL methods to use
- 2) February 26, 2021: Working Prototype
 - a) have a working prototype completed so that we can begin the experimentation phase
- 3) March 4, 2021: Complete experiments
 - a) Test the time interval which best predicts stock performance

- b) Test against random stock picks
- 4) March 11, 2021: Begin building final algorithm
 - a) integrate lessons learned from experiments into a final recommendation system
- 5) March 18, 2021: Final Testing
 - a) Final algorithm evaluation
 - b) Complete stock recommendation
- 6) March 25, 2021: Begin writing report
- 7) April 1, 2021: Complete video presentation
 - a) Create powerpoint presentation
 - b) Report submission

REFERENCES

- [1] GitHub. 2021. AminHP/gym-anytrading. [online] Available at: <<https://github.com/AminHP/gym-anytrading>> [Accessed 10 February 2021].
- [2] Volodymyr Mnih and Koray Kavukcuoglu and David Silver and Alex Graves and Ioannis Antonoglou and Daan Wierstra and Martin A. Riedmiller, . "Playing Atari with Deep Reinforcement Learning". CoRR abs/1312.5602. (2013).
- [3] Pieter Abbeel and Adam Coates and Morgan Quigley and Andrew Y. Ng, . "An Application of Reinforcement Learning to Aerobatic Helicopter Flight." Advances in Neural Information Processing Systems 19, Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 4-7, 2006. MIT Press,
- [4] Terry Lingze Meng and Matloob Khushi, . "Reinforcement Learning in Financial Markets". Data 4. 3(2019): 110.
- [5] D. E. Shaw group. 2021. D. E. Shaw group. [online] Available at: <<https://www.deshaw.com/what-we-do/investment-approach>> [Accessed 10 February 2021].
- [6] Two Sigma. 2021. Two Sigma. [online] Available at: <<https://www.twosigma.com/>> [Accessed 10 February 2021].
- [7] Whalewisdom.com. 2021. RENAISSANCE TECHNOLOGIES LLC Top 13F Holdings. [online] Available at: <<https://whalewisdom.com/filer/renaissance-technologies-llc>> [Accessed 10 February 2021].