Optimizing Trading Ensembles

Raul Girbal [*1] **Avoy Datta** [*1] **Kyle D' Souza** [*1]

1. Project Idea

Ever since the first stock exchange was established in 1602 in Amsterdam, financiers, analysts, and mathematicians have been trying to come up with a model that would "understand the stock market." In other words, predict the future price of a basket of stocks with a degree of accuracy that was significantly better than a human and, equally important, of a random algorithm. With the advent of the electronic markets and microsecond trading, profitable, low-variance automated trading strategies have become ever more important for portfolio managers across the investor spectrum.

With incredibly complex markets, featuring millions of interacting signals and actors, we find that algorithms with sufficient informational depth and the ability to handle changing properties of stochasticity are well suited for tackling a portfolio allocation problem. Deep Reinforcement Learning techniques that leverage differing Actor-Critic Methods to create an ensemble of trading strategies have shown promise in previous work (Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy), and we want to improve the ensemble framework by providing improved sub-algorithms, better ensemble management processes, and higher quality indicators and data. We would like to test our improved ensemble framework on both cryptocurrency markets, standard FX markets, and the traditional US equity markets.

2. Why is this a decision making problem?

As we've mentioned above, automate stock trading strategies are vital; however, the automated nature of these strategies and recommendations creates multiple decision layers, particularly in dynamic, ever-changing markets. There are a few clear decisions that any ensemble trading mechanism has to make. First, what should be the correct strategy weights between certain technical, fundamental, and sentiment triggers? Secondly, should one buy, sell, or hold? In the world of options, these questions only get more complicated. Lastly, a final decision framework can come from a risk/reward ratio. Given a certain risk ratio, how do we either maximize return or minimize risk? For all of these questions, the work done in this class on Reinforcement Learning and Markov Decision Processes have been shown to be effective, given the right scope and scalability.

3. Sources of Uncertainty

The financial markets are full of uncertainty. We are leaning towards exploring optimizing trading ensembles in equity markets; however, we feel as if there are more significant gains to be made with regard to sentiment in cryptocurrency markets or foreign exchange environments. Risk and Reward are difficult to measure: and even then, market swings are hard to trace down often to one trigger and events can tend to result in uncommon results depending on what the corresponding retail or algorithmic reaction is. There is undoubtedly valuation uncertainty, growth vs. value uncertainty, company or item-specific uncertainty, and macroeconomic global uncertainty. Using an effective ensemble model could hopefully amplify the weight for strategies when they are useful and suppress them when not useful, while not muddling any of the individual signals in the process.

4. Prelim. Algorithmic Ideas

The ensemble idea relies on the realization that while reinforcement learning is a great tool to navigate markets, different RL techniques will function best in different markets. Having a ensemble of them to pick and choose given the market conditions is the approach we'll be taking to optimize returns, and we'll be improving the method which we use to pick between the algorithmic choices.

We will be using an Actor-Critic approach for our reinforcement learning algorithms. This focus's on combining an agent that determines the state and act directly (without calculating long value of said action) with a critic that learns a value function and judges the actors choice. The co-interaction of the actor-critic approach should provide more effective learning as well as faster convergence. Some Actor-Critic based algorithms we're thinking of leveraging

[*]Equal contribution [1]Stanford University, Stanford, CA 94305. Correspondence to: Kyle DSouza <kvdsouza@stanford.edu>, Avoy Datta <avoy.datta@stanford.edu>, Raul Overdjiik Girbal <rgirbal@stanford.edu>.

CS 238 Final Project, Stanford University, Fall 2020.

are:

1. Proximal Policy Optimization
2. Adaptive Fuzzy Actor-Critic Learning
3. Advantage Actor Critic

We expect to add and remove algorithms from this list as we run tests and build our ensemble strategy, however these will give us a solid baseline from which to start from (1, 2).

5. Data

Data available in Yahoo Finance and Wharton Financial Databases (through the GSB). The database is available to Stanford students through libguides.stanford.edu.

6. Related Work

Options Trading Ensembles: Our project was inspired by prior work by Yang et al. (2020). The authors acknowledge that a *single* profitable strategy in a dynamic market is often impossible to design. The solution? Use an Ensemble of *statistically independent* reinforcement learning strategies all trained with the objective of maximizing investment returns. The authors look at three actor-critic-based algorithms:

- Proximal Policy Optimization (Schulman et al., 2017)
- Advantage Actor Critic (Babaeizadeh et al., 2017)
- Deep Deterministic Policy Gradient (Silver et al., 2014)

The authors first justify the use of DRL in portfolio optimization, citing aspects such as *experience replay* and the *exploitation-exploration* tradeoffs of Deep Reinforcement Learning algorithms. The authors define their **action space** as the number of shares they can buy and sell, and use a range of technical features as features to the neural network.

References

Babaeizadeh, M., Frosio, I., Tyree, S., Clemons, J., and Kautz, J. Reinforcement learning through asynchronous advantage actor-critic on a gpu. In *ICLR*, 2017.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347, 2017. URL http://arxiv.org/abs/1707.06347.

Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., and Riedmiller, M. Deterministic policy gradient algorithms. volume 32 of *Proceedings of Machine Learning Research*, pp. 387–395, Bejing, China, 22–24 Jun 2014. PMLR. URL http://proceedings.mlr.press/v32/silver14.html.

Yang, Hongyang and Liu, Xiao-Yang and Zhong, Shan and Walid, Anwar, Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy (September 11, 2020). Available at SSRN: https://ssrn.com/abstract=