Webinar: "Applying Deep Reinforcement Learning to Trading" Q&A

Q: How does the Learner figure out the Expected Value? For instance, if the probability of UP movement is 25%, doesn't the Learner need to know the expected movement ($10 vs. $100) change?

A: It learns the expected value of taking an action A in a particular state S by testing different actions and recording the results over time. It does not explicitly model the probability of any particular outcome.

Q: Have RNNs (recurrent neural networks, DNNs with memory, LSTM, etc) been used to train, or it's more easy or practical to layout all the temporal inputs to a traditional DNN?

A: CNN, RNN, LSTM are all candidates for use in learning the models. We have so far used 20 days of daily returns as input to the model.

Q: Can you recommend some good books on Deep learning, Reinforcement Learning and how to apply ML to trading?

A: Here you go

Q: What do you mean by stock regime (regime change)?

A: The macro economic environment. For example we are now in a regime of increasing interest rates and modest inflation. 5 years ago interest rates and inflation were both near zero.

Q:How would you compare Deep reinforcement learning with learning with Genetic programming (Genetic algorithm)? Could you somehow integrate them?

A: Answered in the video.

Q: How do you incorporate unrealized P&L into your data? Is it just the rolling return from a start of your data?

A: It is the unrealized return (cumulative return) of your current position. If you entered the position 5 days ago, it is the cumulative return of that position over the last 5 days.

Q: Do you know any results regarding using a policy gradient method vs. a DQN?

A: Not yet.

Q: I'm curious why you train a NN for each action when using RL. Why not just have one NN for all the actions?

A: Answered in the video.

Q: Can your system be used in short time frames like 30 mins ?

A: Yes, if you have intraday data.

Q: Have you systematically evaluated different lookback/forward return horizons?

A: Not yet looked at different lookbacks. The forward return happens automatically though.

Q: Have you tried one model for all stocks?

A: Yes.

Q: What is the benefit of four separate DL models instead of a single model that outputs the same four outputs?

A: Answered in the video.

Q: Did you try training the model to receive a negative reward if the stock went up, but was not holding it?

A: That is a feasible reward function. We have not used it yet though.

Q: Is the code open source?

A: No

Q: How can we reproduce these results (github, model details, etc)?

A: It is code developed for research that we haven't released yet. Stay tuned for papers.

Q: How do you decide the hyperparameter range when you training the model?

A: For each parameter we usually choose a reasonable range of values, then split each up into 3 specific choices.

Q: What type of models did you find successful? (CNN vs RNN) and (On-policy vs off-policy)?

A: We've been using 20 days of return values as input to a NN. In other work we have found CNNs and RNNs to be useful, but not yet in the context of RL.

Q: Did you always find using separate networks for each action more successful than using a single network with some sigmoid over actions?

A: Yes.

Q: How do you quantize the action space? You said it was (BUY SELL HOLD), but are there discrete values the agent is allowed to buy or sell in?

A: It depends on the specific test, but overall, Long, Short, Cash seems to work well.

Q: How was "Unrealized P&L" calculated?

A: ((price[i]/price[j]) -1) Where i = current day, j = entry point day.

Q: Is the Q-Learning box the Deep Learning network?

A: Yes.

Q: How do you determine the x window size, x1, x2, x3, x4 etc? Is larger window better?

A: We experiment with many different values and use the best ones.

Q: What are the best input indicators which have worked so far?

A: RSI, ADX, CCI.

Q: Why not use Policy Gradient?

A: Any algorithm that solves the RL problem is fine.

Q: Is forward step fixed? like 1 week / 1days to compute reward and update Q.

A: One day.

Q: I tried a similar strategy using ADX, RSI and stochastic oscillator as the inputs. It works very well from the short side but unable to pick up good long positions. Can you highlight the possible reasons for this?

A: That is interesting. It is usually easier to find long strategies and very hard to find short strategies. I don't have an explanation for you.

Q: What is done when training does not converge Q? Results - was the look back too

short?

A: Q is guaranteed to converge on in sample data. Could be. It is worth experimenting with different lookback values.

Q: Why do you repeat the training in the same length time series . How is it going to help?
A: Q will not converge to an optimal policy after one pass on the data.

Q: What technology/frameworks are used for Reinforcement Learning?

A: We use a custom implementation of Q Learning on top of Keras and TensorFlow.

Q: How can interested individuals begin building experience/skill sets in the field of predictive analytics in finance?

A: Answered in video. I recommend taking a course, mine is available at Udacity.

Q: Is the previous daily return on a stock/index acceptable as an input to the state?

A: Yes.

Q: It seems you used a fully connected neural networks, have you tried learning the Q table with recurrent neural networks?

A: Not yet, but that is definitely something we intend to try.

Q: Have you looked at other reinforcement-learning models for this challenge? Why Q-Learning specifically?

A: Some folks recently completed work with video games and DL and found great success using Q Learning. That inspired our approach. Other approaches may also work.

Q: How do you guys use QL in your platform? And how we can utilize it as users?

A: So far it is stand alone code. It is not yet integrated into our platform. If you are interested in using that way, please contact us as that will help accelerate the project.

Q: Have you tried using other metrics like sharpe ratio and drawdown to compute reward function?

A: Not yet, but those are good metrics.

Q: Do you use Options or Hierarchical Q learning to improve generalization? Prediction?

A: Not yet, but those are worth exploring.

Q: Have you experimented with attentional networks to improve performance?

A: Not yet.

Q: When training your models do you take into account commission and slippage costs? if so how has that performed in live environment?

A: Yes we have tested with transaction costs. We have not tried it live though.

Q: Have you looked at including momentum indicators?

A: Yes, they work well.

Q: I have had improved results with training simple ensembles that integrate 2 different cost methods and then trade by combining them. Have you tried this?

A: That sounds like a good idea. No, we have not tried it yet.

Q: Have you applied Monte Carlo methods for search?

A: Not yet.

Q: Does Lucena partner with other companies in the ML space or have process to do so?

A: We primarily partner with data providers and data consumers such as hedge funds. If you are interested in a partnership, contact us and we can talk.

Q: You use the Q-table, is that use the Neural network to construct that table?

A: Yes.

Q: Where's the deep network in your DLR for trading example?

A: It represents the Q table.

Q: To understand clearly we use last 20 days trailing data for training Q policy?

A: Yes, that is the "state" input.

Q: How much time does it take to train the Q policy?

A: Only a few minutes.

Q: A big prerequisite for RL systems is the want of a lot of data. What methods do you use to generate synthetic data that is similar in distributional moments to market data (not Normal!).

A: We use only real historical data.

Q: Do reinforcement methods require the data you provide to be stationary?

A: To support the theoretical guarantees, yes. But the data is never stationary in practice.