From Single to Multi-Agent AI Traders: A Follow-Up to 'Learning Not to Spoof' with Deep Q-Networks and Transformers

William Warlick, 2024

The nascent rise of Artificial intelligence is not only transforming day to day life, but rapidly changing the landscape of AI trading in financial markets Although 70% of share volume in U.S. equities is already estimated to be traded algorithmically, AI heralds a change from explicitly programmed strategies to dynamically learned strategies powered by neural inference engines. With its population and profitable possibilities, care must be put into avoiding AI misuse. Professor David Byrd's paper *Learning Not to Spoof* demonstrated a case where an AI trader with the intuitive goal of profit maximization may results in illegal and manipulative *spoofing*, a form of technical market manipulation (Byrd 2022.) He introduced a method to detect and potentially remediate inadvertent spoofing behavior, employing the detector as a normative guide. This summer, I have extended that work to more advanced forms of reinforcement learning and introduced multiple learning agents.

Informed by a literature review and research into the stock market, I experimentally tested the use of AI as a high-frequency trader in a stock market simulation built by Professor Byrd called ABIDES (Agent-Based Interactive Discrete Event Simulation). Unlike traditional economic models relying solely on statistical equations, ABIDES comprises a variety of virtual trading agents that interact and adapt to each other's actions, providing a more realistic market simulation. Within ABIDES, I implemented an update to the AI trader initially used in Professor Byrd's paper. I updated from the more programmatically explicit discrete Tabular-Q trader to a Deep Q-Network (DQN) trader. A DQN, using the greater state space flexibility of neural networks, is able take a continuous stream of information -- such as stock prices -- without losing information in grouping nearby datapoints into a few boxes as the tabular trader must. This result is the DQN being more useful in the stock market with its better generalization of what it has learned to a variety.

As such, in experimentation I found the DQN to follow a similar pattern as the Tabular-Q trader in inadvertently learning spoofing when given ability to place and cancel limit orders. I also found that the DQN often learned to act profitably without spoofing. I estimate this is because the DQN is better able to infer the hidden states of the simulation such as patterns of the market maker. The graph on the right shows the DQN learning spoofing over 45 days of training and converging on a strategy that proves profitable for 5 days of testing.

I trained transformers from the TensorFlow library, the breakthrough neural network architecture behind LLMs like ChatGPT, for their understanding of sequential data (like the stock market) for spoofing detection. This detection approach trained more quickly than the prior approach using a LSTM from Professor Byrd's work. Nevertheless, the transformers achieved parity in the accuracy of spoofing detection at 99.9%. However, detection becomes substantially more computationally intensive if one *assumes* the possibility of multiple DQNs working together. I assume the case that a trading firm might route through different brokers or be so large that different internal groups have their own connections to the exchange. If trading at high frequency, a seemingly intelligent move might be to share communication between DQN traders about strategy or upcoming trades, ensuring they align in profit rather than work against each other.

When multiple DQN agents collaborate by sharing trade content before reaching the exchange, they can remain profitable while inadvertently making spoofing detection exceptionally computationally expensive. In its technical use, the increase in the number of connection points to the exchange is worse than the problem becoming exponentially more computationally expensive. The computational demand of running spoofing detection on all possible subsets (up to the size of most presumed connections to exchange by an entity with common profit interest) of the number of connection points to the exchange is algorithmically infeasible. As the number of possible connection points in a single organization increase, a combinatorial number of configurations of connection points must be tested for spoofing.

This research reveals that the uninformed use of AI can accidentally mask fraudulent behavior, necessitating further investigation into detection techniques and the institution of best practices around AI. Developing methods for reducing the size of the detection pool will be critical, and one logical approach might involve determining a threshold of the capital needed to cause spoofing.

Faculty Mentor: David Byrd

Funded by the J.P. Morgan AI Research Faculty Fellowship

Byrd, David. "Learning not to spoof." Proceedings of the Third ACM International Conference on AI in Finance, 2022, https://doi.org/10.1145/3533271.3561767.