Algorithmically Trading like a Human with GPT-J

Joel Peckham
School of Computing
Southern Adventist University
Collegedale, TN 37363
Email: joelskyler@gmail.com

***Abstract*—Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.**

I. Introduction

Algorithmic trading has become the dominant way of buying and selling securities. In the U.S. stock market, algorithms account for 70-80% of trading volume [1]. The next generation of trading algorithms use neural networks to improve prediction accuracy. However, current opinion states neural networks can only increase efficiency by about 10% and that neural networks are not capable of inventing winning trading ideas [2]. We theorize that neural networks are capable of much more. By scaling the number of network parameters into the billions, training the network on a massive dataset which captures the complexities of human behavior, and then making inferences from current and by feeding the network with current news and opinons we believe that neural networks can model the world accurately enough to discover winning, novel trade ideas. Thankfully, the creators of GPT-J have already created such a neural network [3]. GPT-J is a transformer network [4] that has been trained on an 825GiB language modelling data set called The Pile [5]. We propose that with adequate fine-tuning and well engineered prompts, GPT-J can learn to trade like a human by reading the news, social media, and other sources such as SEC filings.

In this paper we will first implement a trading algorithm based on GPT-J, then we will evaluate the algorithm's prediction accuracy and compare it to the accuracy of similar algorithms. Finally, we will evaluate the real-world profitability of the algorithm using both backtesting and live market data.

A. Problem Statement

In summary, as algorihmic traders seek to employ increasingly competitive and complex trading algorithms, we intend to evaluate the viability of using GPT languange models to drive trading decisions. To achieve this goal we must first implement a trading algorithm based on GPT-J, then evaluate the accuracy of the algorithm's predictions, and finally evaluate the real-world profitability of the algorithm.

II. Background & Related Works

Many neural network structures and methods have been used to create trading algorithms. Some types include:

- Recurrent neural networks (RNNs) & Long short-term memory (LSTMs) [6][7].
  These are the most widely used network architectures for trading [8]. An LSTM trained on 900,000 sequences of length 30 days of Chinese stock market data yeilded an improvement of 12.9% in prediction accuracy over a random guess [9].
- Convolutional Neural Networks (CNNs) [8]
  CNNs can be used by converting time-series data into images [10]. Or they can be used to extract sentiment features from text [11].
- Deep reinforment learning (DRN).
  - Deep Q-learning [12] [13].
  - Deep robust reinforcement learning [14].
- Conventional deep learning [15].
- Transformer networks [16].

Most relevent to our work are methods that incorperate sentiment analysis of news sources. Mehta et al. (2021) evaluated a sentiment analysis methods and found that LSTMs could properly classify news tweets as indicative of positive or negative price movement with an accuracy of 92% correct [7]. Nan & Zaiane (2020) found that adding sentiment analysis to a Deep-Q learning algorithm could improve the sharpe ratio of the agent by a factor greater than 2 in their test cases. [13].

By our estimation, the vast majortity of previous works involving sentiment analysis used a pre-processing step to extract sentiment from the news, and then embedded those features into a time-series dataset. News sources were often limited to headlines, tweets and small snippits because of the memory limitations of RNN sentiment classifiers. With the indruduction of large transformer networks [4], capable of processing large amounts of text like OpenAi's GPT-3 [17] or Wang & Komatsuzaki's GPT-J [18], we believe a new class of trading network can be created. Our method will encode the current world state in a large text-input wich combines

sentiment, real-world facts, and stock price data into a single input. We believe that this new network can be trained to learn to trade like a human.

III. Execution Plan

A. Requirements & Goals

1) Functional requirements (user stories):

- As finance researchers, we want to quantify the ability of news releases, SEC filings, and social media posts to move stock prices.
- As AI researchers, we want to evaluate the viability of using GPT-J as a stock movement indicator so that we can understand the power of GPT-J to understand complex real-world interactions.
- As algorithmic traders, we want to evaluate the viability of using GPT-J as a stock movement indicator so that I can make more-informed trading decisions.

2) Non-functional requirements: Our overall flow of execution as outlined in figure 1 is as follows:

- We will evaluate the correlations between the following items:
  a) The release of SEC filings for company X and large movements in the stock price of company X.
  b) The release of news stories mentioning company X and large movements in the stock price of company X.
  c) The posting of tweets or other social media posts mentioning company X and large movements in the stock price of company X.
- We will evaluate the best ways of formatting prompts for GPT-J to increase output accuracy and consistency across varying inputs. Some options might include:
  a) Providing a form for GPT-J to fill out appended to the end of the input data.
  b) Asking GPT-J a direct question appended to the end of the input data.
  c) Appending a universe current stock prices to the beginning of the input.
  d) Appending multiple news stories from the past days and weeks at the beginning of the input.
- We will deploy the best model from my previous evaluations and test it on live stock market data. I will compare the model's performance against market indices like the S&P 500.

B. Testing & Evaulation Methods

1) Media Release Correlations: To measure the correlation between the release of a media item such as an SEC filing or news story, we will use the standard event study method as detailed in [19]. This method uses abormal returns in a given period to calculate the effect of a certian event on a stock's price. Abnormal return for a given day is defined as

$$AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t}) \tag{1}$$

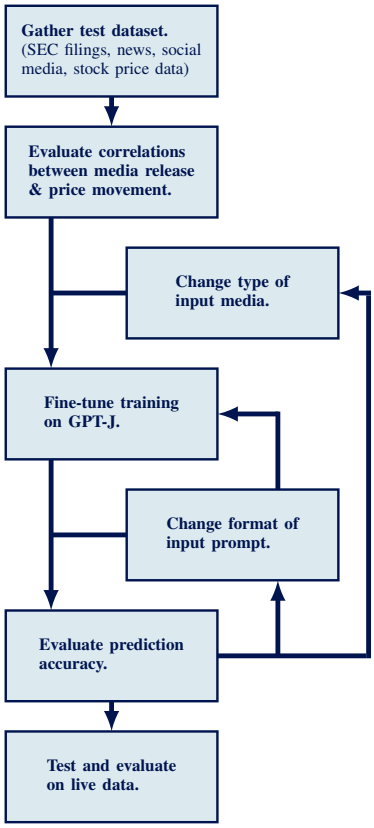

Fig. 1. Execution Flow

for firm i at time t where α_i and β_i represent the relationship between a given stock and it's refrence index. And where $R_{m,t}$ is the return of the actual reference market.

We will then take large sample of media release events of the same type and calculate the average abnormal return as follows:

$$AAR = \frac{1}{N}\sum_{i=1}^{N} AR_{i,t} \tag{2}$$

2) GPT-J Model prediction accuracy: Finally, we meausre the total impact of the event over a given period of time by using cumulate abnormal return:

$$CAR(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_{i,t} \tag{3}$$

where t_1 and t_2 are the start and end dates of the event window.

IV. Implementation

V. Conclusion

The conclusion goes here.

Acknowledgment

The authors would like to thank...

REFERENCES

[1] H. Samuelsson, "What percentage of trading is algorithmic?" 2021. [Online]. Available: https://therobusttrader.com/what-percentage-of-trading-is-algorithmic/

[2] D. Vonko, "Neural networks: Forecasting profits," 2021. [Online]. Available: https://www.investopedia.com/articles/trading/06/neuralnetworks.asp

[3] B. Wang, "Mesh-Transformer-JAX: Model-Parallel Implementation of Transformer Language Model with JAX," https://github.com/kingoflolz/mesh-transformer-jax, May 2021.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *CoRR*, vol. abs/1706.03762, 2017. [Online]. Available: http://arxiv.org/abs/1706.03762

[5] L. Gao, S. Biderman, S. Black, L. Golding, T. Hoppe, C. Foster, J. Phang, H. He, A. Thite, N. Nabeshima, S. Presser, and C. Leahy, "The pile: An 800gb dataset of diverse text for language modeling," *CoRR*, vol. abs/2101.00027, 2021. [Online]. Available: https://arxiv.org/abs/2101.00027

[6] G. Chen, Y. Chen, and T. Fushimi, "Application of deep learning to algorithmic trading," 2017. [Online]. Available: http://cs229.stanford.edu/proj2017/final-reports/5241098.pdf

[7] P. Mehta, S. Pandya, and K. Kotecha, "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning," *PeerJ Computer Science*, vol. 7, p. e476, Apr. 2021. [Online]. Available: https://doi.org/10.7717/peerj-cs.476

[8] Y. Gu, T. Shibukawa, Y. Kondo, S. Nagao, and S. Kamijo, "Prediction of stock performance using deep neural networks," *Applied Sciences*, vol. 10, no. 22, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/22/8142

[9] K. Chen, Y. Zhou, and F. Dai, "A lstm-based method for stock returns prediction: A case study of china stock market," in *2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 2823–2824.

[10] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach," *Applied Soft Computing*, vol. 70, pp. 525–538, Sep. 2018. [Online]. Available: https://doi.org/10.1016/j.asoc.2018.04.024

[11] Y. Shi, Y. Zheng, K. Guo, and X. Ren, "Stock movement prediction with sentiment analysis based on deep learning networks," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, Nov. 2020. [Online]. Available: https://doi.org/10.1002/cpe.6076

[12] Y. Wang, D. Wang, S. Zhang, Y. Feng, S. Li, and Q. Zhou, "Deep q-trading," 2017. [Online]. Available: http://cslt.riit.tsinghua.edu.cn/mediawiki/images/5/5f/Dtq.pdf

[13] A. Nan, A. Perumal, and O. R. Zaiane, "Sentiment and knowledge based algorithmic trading with deep reinforcement learning," 2020.

[14] Y. Li, W. Zheng, and Z. Zheng, "Deep robust reinforcement learning for practical algorithmic trading," *IEEE Access*, vol. 7, pp. 108 014–108 022, 2019. [Online]. Available: https://doi.org/10.1109/access.2019.2932789

[15] M.-Y. Day and C.-C. Lee, "Deep learning for financial sentiment analysis on finance news providers," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, Aug. 2016. [Online]. Available: https://doi.org/10.1109/asonam.2016.7752381

[16] J. Schmitz, "Stock predictions with state-of-the-art transformer and time embeddings," Sep 2020. [Online]. Available: https://towardsdatascience.com/stock-predictions-with-state-of-the-art-transformer-and-time-embeddings-3a4485237de6

[17] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," *CoRR*, vol. abs/2005.14165, 2020. [Online]. Available: https://arxiv.org/abs/2005.14165

[18] B. Wang and A. Komatsuzaki, "GPT-J-6B: A 6 Billion Parameter Autoregressive Language Model," https://github.com/kingoflolz/mesh-transformer-jax, May 2021.

[19] A. Neuhierl, A. D. Scherbina, and B. Schlusche, "Market reaction to corporate press releases," *SSRN Electronic Journal*, 2010. [Online]. Available: https://doi.org/10.2139/ssrn.1556532