Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models

赵伟皓

05.14.2024

1 Research questions

Can ChatGPT and other language models predict future stock returns using news headlines?

2 Why are the research questions interesting?

• The generative AI and LLMs' potentials in different fields have received widespread attention.
• However, there is rare research on using LLMs to predict stock returns in the financial field.
    1). People do not value LLMs because these models are not explicitly trained for returns prediction.
    2). LLMs are of value for processing textual information to predict stock returns.
• Due to above controversies, the performance of LLM in predicting market trends is an unresolved issue.
• ChatGPT surpasses traditional sentiment analysis methods in predicting stock returns.

3 What is the paper's contribution?

***(1)* Literature uses ChatGPT in the context of economics**
Previous: Using ChatGPT to help solve portfolio selection problems, but without positive performance.
This paper: Studying the potential of LLMs in financial markets(especially decision-making process).

***(2)* Literature uses text analysis and ML to study financial research questions**
Previous: Generating features, showing chances for text mining. Testing features in predicting returns.
This paper: GPT extracts signals from news headlines to outperform sentiment measures in predictions.
• A novel evaluation tech to understand GPT's reasoning by predicting the correctness of recommendations.

***(3)* Literature uses linguistic analyses to extract sentiment and predict stock returns**
Previous: Using the sentiment of firm news to predict future individual stock returns.
This paper: Focusing on whether LLMs extract additional information that predicts stock market reactions.

***(4)* Literature on employment exposures and vulnerability to AI-related technology**
Previous: Studying the extent of job exposure and vulnerability to AI-related technology, and impacts.
This paper: The potential of LLMs in adding value to participants in processing info to predict returns.

4 What hypotheses are tested in the paper?

H1: ChatGPT's assessment score predicts the subsequent daily stock returns.
H2: Complex LLMs provide more accurate return forecasts compared to basic ones.

a) Do these hypotheses follow from and answer the research questions?
• Yes, these hypothesis provide assistance in answering research questions.
Do these hypotheses follow from theory or are they otherwise adequately developed?
• ChatGPT's predictability outperforms traditional sentiment analysis methods from leading data vendors.
• GPT-4 has the highest Sharpe ratio, indicating predictability is an emerging ability in complex LLMs.
• The predictability exists in both small and large stocks, indicating insufficient market response to news.
• Predictability is stronger among stocks with bad news, consistent with limits-to-arbitrage.

5 Sample: comment on the appropriateness of sample selection procedures.

The three datasets accurately reflect the problem that the paper aims to test. Matching with RavenPack assures that only relevant news are used for testing.

6 Dependent and independent variables: comment on the appropriateness.

The dependent variable (stock returns) and the independent variable (vector containing the ChatGPT score or other scores) is relatively straightforward.

7 Regression model specification: comment on the appropriateness.

The paper used simple linear regression to control for fixed effects, thus intuitively demonstrating how ChatGPT scores affect stock returns.

8 What difficulties arise in drawing inferences from the empirical work?

The paper compares GPT-4 with basic models, but the prompt strategy for the basic language model is somewhat hasty, the exploration of basic models seems to have no value for comparison and reference.

9 Describe at least one publishable and feasible extension of this research.

Future research can explore the role of LLM in shaping market behavior and its potential positive and negative impacts on the financial system.

References

[1] Ko, Hyungjin, and Jaewook Lee. 2023. "Can Chatgpt Improve Investment Decision? From a Portfolio Management Perspective." SSRN Electronic Journal.

[2] Chin, Andrew, and Yuyu Fan. 2023. "Leveraging Text Mining to Extract Insights from Earnings Call Transcripts." Journal Of Investment Management 21 (1)..