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

Working paper, 2023

Zhenting Hu

May 14, 2024

1. What are the research questions?

- Whether LLMs have the ability to predict stock returns?
- What are the models' reasoning mechanisms?

2. Why are the research questions interesting?

- Few studies are using LLMs to predict stock returns.
  - Prior studies use textual analysis and ML methods.(see Chin and Fan (2023))
- The application of LLMs remains controversial
  - They are not explicitly trained for stock return prediction.
  - They're more capable of understanding the context of natural language because of the data scale.

3. What is the paper's contribution?

- Contribution to literature on potential of LLMs in financial markets.
  - Prior studies focus on using historical numerical data to predict.(see Korinek (2023))
  - This paper focuses on LLMs in textual tasks.
- Contribution to literature employs textual analysis and ML methods to study financial questions.
  - This paper uses LLMs, especially ChatGPTs.
- Contribution to study on the evaluation of ChatGPT.
  - This paper propose a novel evaluation technique to understand ChatGPT's reasoning by predicting the correctness of the recommendations.
- Contribution to study analysing news articles.
  - Prior studies use media or firm's news to predict stock returns.
  - This paper focus on understanding whether LLMs add value.
- Contribution to studies on employment exposures and vulnerability to AI-related technology.
  - Prior studies have examined the effect of AI-related technology on employment and productivity.
  - This paper highlight the value added by LLMs in financial markets.

4. What hypotheses are tested in the paper?

- H1: ChatGPT scores are in positive relation with stock returns.
- H2: ChatGPT scores perform better than traditional ways.

a) Do these hypotheses follow and answer the research questions?

• Yes.

b) Do these hypotheses follow from theory? Explain logic of the hypotheses

• Yes, H1 focuses on the main research questions, H2 digs deeper on the topic of a horse race test

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

• Sample periods(Sept 2021 to Oct 2022) is well-suited for explaining the out-of-sample predictability of LLMs.

6. Dependent and Independent Variables.

Prompts used for independent variables are comprehensive.

7. Regression/prediction model specification.

• Maybe other control variables are necessary except for prediction scores.

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

• Though there are comprehensive prompts, there is a significant challenge in validating these interpretations.

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

• Other more complicated materials for ChatGPT analysing (news abstractsshort articles)