

Financial Forum Future: Comparing and evaluating ChatGPT’s performance giving financial advice using Reddit questions and answers

Sathvik Samant* (ssamant26@lawrenceville.org), Aditya Dhar* (adidhar46@gmail.com), Shreya Kochar (shreya.kochar@columbia.edu), Aneesha Sreerama (sreerama.a@northeastern.edu), Andrew Wang (andrew.wang.2@stonybrook.edu), Anirudh Sreerama (anirudh.s@berkeley.edu)

Abstract

As artificial intelligence, and particularly Large Language Models (LLMs), rapidly advance, there is a growing interest in how they can be used in the world of financial advising. In order to examine, investigate, and evaluate the performance of such LLMs in the role of a financial advisor, this experiment utilizes financial questions asked on the Reddit forum “r/Financial Planning” in order to test the GPT-4 LLM. The results were compared to actual Reddit comments, assessing the model’s response content, length, and advice. By evaluating the model’s advisory competency, this study explored the role of AI in financial forums, its ethical consequences, and potential threat to employment and existing systems. It found that while AI can present accurate information, it fails in its delivery, clarity, and decisiveness.

Keywords: GPT-4, Financial forums, Finance, Generative AI

1. Introduction

Given the rapid development of artificial intelligence(AI), particularly within the past year, a growing focus has been given to the applications of these technologies in various, interdisciplinary fields. Due to this advancement, there has been a growing call for innovation within the financial sector, including financial advisory. In particular, growing research seeks to explore and develop these AI technologies to help users develop financial literacy, and give financial recommendations.

One such content creation tool, ‘Chat Generative Pre-Trained Transformer’(ChatGPT), was released globally in late 2022. Using various data inputs, ChatGPT¹ generates explanatory sentences in conversational-style settings, which greatly increases efficiency in meeting user needs. While it is an extremely useful and powerful tool in a multitude of fields and settings, ChatGPT is a significant tool in generating financial recommendations, acting in a similar capacity as an advisor. Although this tool does not take in raw financial data, it can respond to specific situational text blocks, and provide options and responses.

This research particularly aims at comparing ChatGPT’s financial advice with advice given on Reddit forum “r/FinancialPlanning,” in order to understand how the model responds to specific financial situations in comparison to real human advice. This comparison was automated using natural language processing (NLP) techniques in order to quantitatively understand similarities between responses. Primarily, this study focused on content, or what feedback the model suggested, as well as structure, length, and potential ethical implications.

Although prior research has been conducted in the field of LLMs in finance, many focus on financial markets or market trends, instead of personal finance and financial advisory². Other similar research that has been conducted often compares GPT-4 responses to real financial advisor advice, while ours particularly focuses on other Reddit users to explore how such LLMs might shape the future of such forums. While this research does not focus on evaluating the accuracy and content of GPT’s responses, it instead focuses on comparing the responses with real user responses to understand its capabilities in performing similar to a real user.

To understand the future role of ChatGPT in financial advisory spaces, this research was guided by multiple core questions:

- How does GPT-4’s financial advice perform in comparison to the advice given on “r/FinancialPlanning” forum?
- How can we quantitatively determine these similarities and differences?
- What are ChatGPT’s strengths? Weaknesses?
- What do Chat GPT’s strengths and weaknesses tell us about the role that ChatGPT plays in financial forums, advisory, and the future?

The research question being addressed in our project is whether LLMs, specifically GPT-4 is competent in giving financial advice compared to a human being and how financial

¹OpenAI (2023). GPT-4 Technical Report. ArXiv, abs/2303.08774.

²Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., & Bendersky, M. (2022). What do LLMs Know about Financial Markets? A Case Study on Reddit Market Sentiment Analysis. Companion Proceedings of the ACM Web Conference 2023.