

# 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, ChatGPT<sup>1</sup> 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 advisory<sup>2</sup>. 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

---

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

<sup>2</sup>Deng, 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.

spaces will be impacted with AI in the picture, specifically in terms of ethics and future employment. This project is unique as there has not been too much research done specifically on how AI financial advice compares to advice given by a financial advisor. Our project is exciting as AI in financial spaces brings in opportunities to investors that were not available before such as unbiased and completely personalized financial advice based on the investor's objective and risk-tolerance. Another exciting aspect of AI in financial spaces is the ability for the AI to potentially analyze data and financial trends at a much faster speed than any human being. Lastly, this project is both important and pertinent as it seeks to quantitatively analyze the current progress of AI in the space, as well as identifying drawbacks, dilemmas, and areas where improvement is needed.

## 2. Literature Review

ChatGPT, an Artificial Intelligence model developed by OpenAI, has been gaining attention and interest in various fields, including finance. The combination of AI in the field of finance has garnered increasing interest due to its many capabilities such as improved risk management for investors, fraud detection, customer service, and compliance capabilities that could facilitate a deeper understanding of the compliance requirements of multilevel financial advice. Several studies have explored the use of ChatGPT in the investment domain and have examined its potentials and limitations.

In a study by Neilson<sup>3</sup>, the use of ChatGPT in assisting investment decisions was researched. They found that ChatGPT can be a valuable tool in this context, as it can quickly cut the time taken to process large amounts of data and generate relevant insights for investors who are looking for simple financial recommendations. However, in the study, it was determined that ChatGPT could not effectively operate with complex financial advice and required professional guidance in order to adhere to regulatory requirements. ChatGPT was ultimately determined in the current AI space to be more of a tool that could help save time and cost rather than an independent entity. Another study conducted by Cao<sup>4</sup> went into detail about many of the benefits of an AI empowered economy and financial spaces. Specifically, Cao addressed how AID (Artificial Intelligence and Data Science) may be used to take on complex economic and financial challenges while also being utilized to create an information as well as intelligence-driven economy and financial system. On the other hand, a study by Biswas et al.<sup>5</sup> investigated some of the more specific limitations and challenges of using ChatGPT in investment decision making. They highlighted the potential concerns such as the fact that ChatGPT is not reflective on a dynamic market and rather is dependent upon historical data. The study also highlighted how data security and the transparency in ChatGPT could pose a potential

threat to investors as data could potentially be manipulated or biased leading to a negative outcome for the investor. All of the limitations presented in the study could have implications for the reliability and trustworthiness of the generated investment insights.

While each of these studies address the benefits and limitations of AI models such as ChatGPT in giving financial advice, research analyzing the effectiveness of ChatGPT's financial advice compared to finance advice given on social media platforms, such as Reddit, has yet to be analyzed.

## 3. Methods

### 3.1. Tools

This study used the Jupyter Notebook editor for python Programming. To collect the data used for this research, the Reddit API was used, while OpenAI's GPT-4 API was used to generate LLM responses. The following packages were used for data manipulation, analysis, and presentation: Pandas, Seaborn, SpaCy, SciPy, Natural Language Toolkit(NLTK), and Matplotlib.

### 3.2. Data Collection

In order to collect the raw user data, the Reddit API was due to its accessibility, as well as long standing relevance in financial spaces. Due to the focus of the study on financial advisory, the API was used to scrape from the forum "r/FinancialPlanning," a popular forum to ask and provide personal financial advice, which is subscribed by over 685K users. In this subreddit, users ask for financial advice in posts, while fellow users can respond to them in comments, which receive upvotes based on community perception of their accuracy and quality. Due to time and material constraints, only the 100 most recent posts were scraped, alongside their most upvoted comment in order to gauge a popular response to the question. Using the Pandas library, this data was stored as a dataframe in order for later manipulation and analyses.

### 3.3. Generating GPT-4 Responses

In order to generate LLM responses to community questions posed on Reddit, this study utilized the GPT-4 API's Chat Completion API. This API uses GPT-4, an advanced LLM notable for its dialogue and summarization capabilities, as well as passing the bar exam amongst the top 10% of test takers. The questions posed on the forum were inputted into the API, and their responses were collected within the dataframe.

### 3.4. Data Pre-processing

To ensure that the data was understandable by the SpaCy similarity algorithm, the NLTK library was first implemented in order to pre-process our data in order to improve the accuracy of the natural language processing algorithms. This processing included: removing stop words, which include removing articles, pronouns, and conjunctions that don't add significant information to the text; making the data lowercase and removing any symbols; and lemmatization, or converting words into their meaningful base form. These preprocessing methods helped to improve the performance and accuracy of the NLP conducted.

<sup>3</sup>Neilson, B. (2023). Artificial intelligence authoring financial recommendations: Comparative Australian evidence.

<sup>4</sup>Cao, L. (2020). AI in finance: A review. SSRN Electronic Journal.

<sup>5</sup>Biswas, S., Joshi, N., & Mukhopadhyaya, J. N. (2023). ChatGPT in investment decision making: An introductory discussion. Unpublished.

### 3.5. SpaCy Similarity

One method of comparison was utilizing the spaCy similarity algorithm to quantify the similarity of the two sentences using a score from 0 to 1, semantically. For this analysis, spaCy used a large model containing word vectors. This similarity is by comparing multi-dimensional meanings or word vectors of words within the text. This similarity score was collected within the dataframe, and the Matplotlib library was used to generate graphs and plots to visualize the results.

### 3.6. Sentiment analysis with Twitter-roBERTa

Based on empirical observation of the Reddit responses, many instances of negative, neutral, and positive sentiment were observed. For instance, some AI responses or Reddit comments indicated approval and praise, while others indicated potentially negative sentiment. Due to this observation, sentiment analysis of both the AI response and Reddit user comments was conducted in order to gauge differences and similarities in sentiment between them. While the content of this data surrounded financial matters, its structure and content aligned most closely with models trained on social media data, rather than financial news and markets data. Hence, the “Twitter-roBERTa-base for Sentiment Analysis” model was employed for this analysis due to the similarity between its training content and the collected data used in this study. The labels (negative, neutral, positive) were collected within the dataframe for analysis.

### 3.7. Bi/Tri-gram analysis

Often, meaningful phrases or sequences of words are more informative and descriptive than singular words. N-gram analyses allow for deeper word and sentence structure analysis to complement NLP. Hence, bigram and trigram analyses were implemented on GPT response and Reddit comment data in order to reveal word sequence patterns that appear across collected data. The NLTK tokenizer and ngrams library was used for this analysis, alongside the collections library. These word patterns were sorted by frequency, in order to be analyzed later.

### 3.8. Response Structure

In order to demonstrate the structure and content of GPT-4 responses, an example is provided below. Bold text indicates a header, or large piece of financial advice. This response provides six individual financial strategies, alongside an introduction and conclusion.

”As an AI language model, I cannot provide personalized financial advice, but I can offer some general insights to help you make an informed decision. It’s essential to consider your individual circumstances and consult with a licensed financial advisor who can provide personalized advice based on your specific situation. Purchasing life insurance, especially term life insurance, is a common strategy to provide financial protection for your family in case of your untimely death. Here are some points to consider:

1. **Family’s Financial Security:** Given your current situation with a stay-at-home wife and young children, having life insurance can provide a financial safety net for your family if something were to happen to you during the term. The death benefit could help cover living expenses, child-care costs, and other financial needs.
2. **Affordability:** The premium you mentioned for a \$1 million, 25-year term policy seems reasonable, considering the coverage amount and the duration. Life insurance premiums tend to be more affordable when you are younger and healthier, making it an advantageous time to secure coverage.
3. **Long-Term Planning:** While you may consider your kids being adults in 25 years, life’s uncertainties can create a need for financial protection even beyond that timeframe. Moreover, having life insurance in place early can lock in lower premiums for the entire term.
4. **Existing Investments:** It’s great to hear that you have some investments outside of retirement accounts, which can provide additional financial security for your family. However, life insurance can offer a lump sum payout that can be used immediately and for specific purposes.
5. **Risk Tolerance and Personal Comfort:** The decision to purchase life insurance is highly personal and dependent on your risk tolerance and comfort level. It’s normal for people to have different opinions about insurance needs. Ultimately, you need to feel confident in your decision.
6. **Consider Future Needs:** As your family grows and circumstances change, your insurance needs may evolve. Review your coverage periodically to ensure it aligns with your family’s changing needs.

In conclusion, life insurance can be an essential component of a comprehensive financial plan, especially when you have dependents and young children. The decision to purchase life insurance should be based on a careful assessment of your financial situation, family’s needs, and long-term goals. Consulting with a financial advisor can provide you with personalized guidance to make an informed decision.”

## 4. Results

The GPT-4 responses had a higher average word count of 184, while the Reddit comments had an average of 135. Despite having different mean word count, with +/- 2SE there is no significant difference between word count length. Figure 1 provides a general baseline comparison of the word counts between the GPT-4 output and the Reddit comments.

### 4.1. Statistical Analysis and Graphing

In order to further present and interpret the analyzed data, various statistical and visualization techniques were used, particularly utilizing the Pandas, Seaborn, and Matplotlib library.

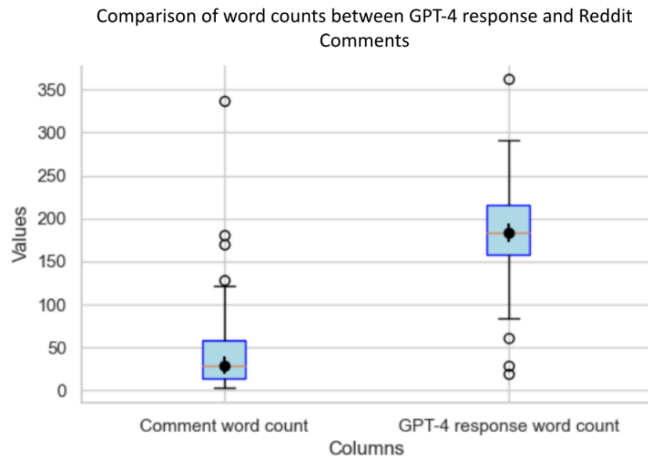


Figure 1: These box plots with error bars indicate average word counts observed between GPT-4 and Reddit comment responses. Outliers are plotted with hollow circles, and the error bars equal 2 standard errors. While there is a difference in averages between the word counts, it is not significant, as indicated by the overlap.

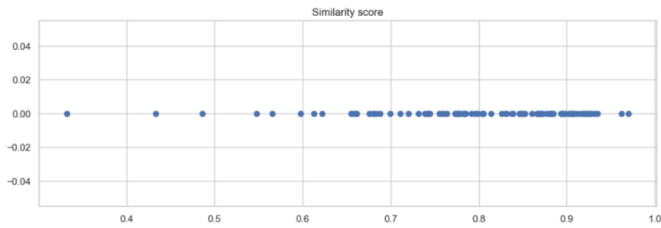


Figure 2: similarity scores between the two texts are predominantly in the highest quartiles.

One of the methods of comparison, the spaCy similarity algorithm, was an NLP approach to quantifying word vector similarity. The mean of the output (similarity score) was 0.803, while the median was 0.834. A similarity score of 1.0 indicates perfect word vector similarity, while 0.0 indicates no similarity at all. Additionally, 97.00% of similarity scores are above 0.5, and 74.00% are above 0.75. This can be seen in Figure 2.

Figures 3 and 4 display the distribution of sentiment labels between GPT-4 responses and Reddit comment responses. Both have a similar distribution of neutral text sentiment, while Reddit comments have a higher proportion of negative sentiment (17) and GPT responses have a higher proportion of positive sentiment (28).

In order to determine the presence of a correlation between word count of Reddit input vs word count of GPT-4 response, the word counts were plotted on a scatter plot using the Matplotlib library. The Pearson correlation coefficient (PCC) was used to determine linear correlation between the two, and returned a coefficient of 0.02, as well as a p-value of 0.8138. At the 0.05 significance level, the correlation between the two is not statistically significant. No linear correlation was determined between the two, and the average GPT-4 response word



Figure 3: The stacked bar chart shows the proportions of each text sentiment based on whether the response is generated by GPT-4 or an actual reddit comment. This helps visualize the proportional sentiment between the two responses, indicating contrast between sentiment levels.

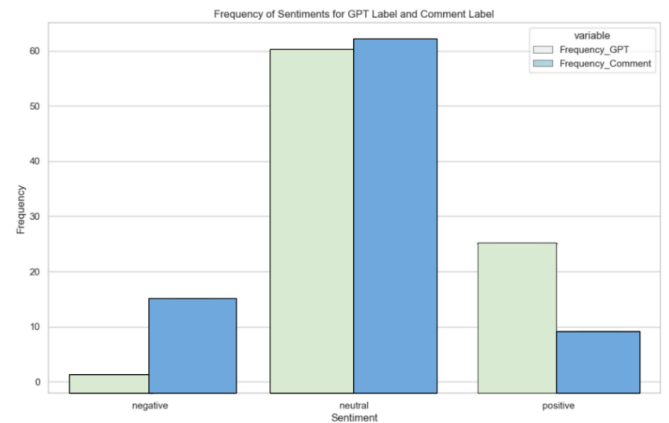


Figure 4: This figure offers a direct comparison of sentiment responses GPT and Reddit comments, light green indicating GPT and dark green indicating comment responses.

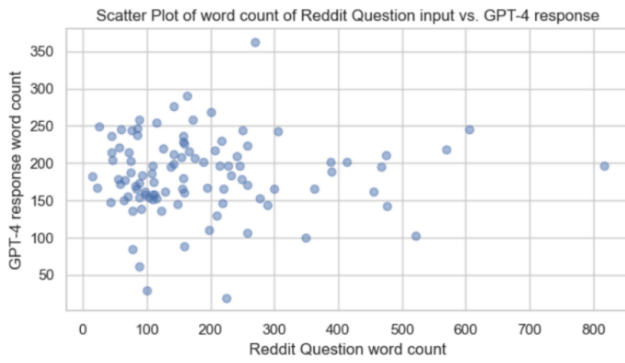


Figure 5: Figure 5 is a scatter plot of the GPT-4 response word count. No linear correlation was determined either through the graph or correlation analysis.

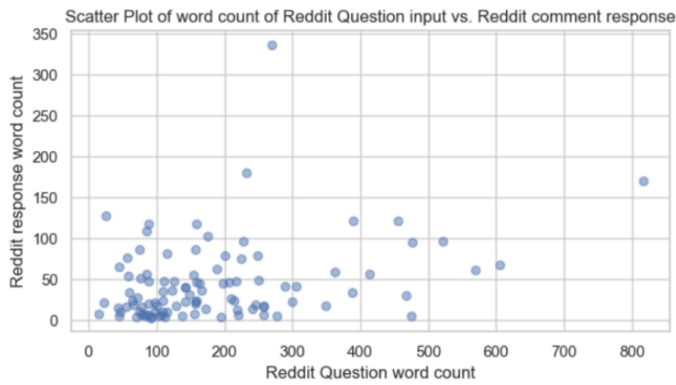


Figure 6: Figure 6 is a scatter plot of the Reddit response lengths, mapped by the x and y axes accordingly. No linear correlation was determined either through the graph or correlation analysis.

count of 185.58 with a standard deviation of 141.45.

Finally, we analyzed our n-gram results. The top 10 most frequently found bigrams and trigrams within the GPT-4 response, alongside their frequency, were collected in the below tables. The trigram “consult, financial, adviser” appeared most frequently, while the bigram containing “financial, adviser” appeared 107 times. The outcomes of this work can be seen in Figures 7 and 8.

## 5. Discussion

Based on the findings of this research, many comparisons can be drawn between GPT-4’s response to financial questions and actual Reddit user comments. Its ability and efficiency to provide financial recommendations is comparable to human comments, but unparalleled in speed, efficiency, and language processing skills. Despite this, the results of this study present potential drawbacks of GPT-4’s financial advice including its length, lack of definitive answer, and constant repetition in word structure patterns.

GPT_Response_Trigram			GPT_Response_Bigram		
	3-gram	Frequency		2-gram	Frequency
0	(consult, financial, advisor)	34	0	(financial, advisor)	107
1	(provide, personalized, advice)	20	1	(emergency, fund)	88
2	(financial, advisor, provide)	19	2	(financial, goals)	72
3	(based, information, provided)	18	3	(roth, ira)	54
4	(goals, risk, tolerance)	16	4	(risk, tolerance)	50
5	(consulting, financial, advisor)	16	5	(financial, situation)	49
6	(make, informed, decision)	15	6	(interest, rate)	43
7	(help, make, informed)	14	7	(real, estate)	41
8	(personalized, advice, based)	14	8	(credit, card)	40
9	(credit, card, debt)	14	9	(retirement, savings)	37

Figure 7: These tables demonstrate the results of n-gram analyses, showing word sequence patterns throughout the GPT and Comment responses

Comment trigram			Comment bigram		
	3-gram	Frequency		2-gram	Frequency
0	(credit, card, debt)	4	0	(roth, ira)	20
1	(max, roth, ira)	3	1	(emergency, fund)	15
2	(taxable, brokerage, account)	3	8	(discretionary, income)	8
3	(selling, car, buying)	2	9	(brokerage, account)	7
4	(carrying, credit, card)	2			

Figure 8: GPT responses have higher and more frequent tri/bi-gram patterns than Reddit Comments.

### 5.1. Content Comparison

The first method of analysis employed, the spaCy similarity algorithm, found similarity levels between the word vectors of the GPT-4 responses and Reddit Comments. As seen in Figure 2, the results of this algorithm indicate a clear similarity in word structure and vocabulary between the two texts, demonstrating a clear connection in the advice and financial content that they are recommending. However, Figures 4a and 4b display a large difference in word count between the 2 texts: the GPT-4 response is notably more prolific in comparison to Reddit comments. Reddit comments are substantially shorter and more focused on the question, while GPT-4 responses are longer, and as empirically observed, less focused. Instead of directly providing an opinion or response, which AI cannot do, they suggest many avenues that are potentially viable options. In fact, when asked why it cannot provide financial advice, GPT-4 states that it cannot provide personal finance advice without a “understanding of an individual’s unique financial situation, goals, [and] risk tolerance,” but can provide general information and answer non-specific questions. While GPT-4 is able to provide a multitude of strong suggestions, it rarely ever recommends a specific financial strategy, showing a substantial difference between its advice and Reddit responses. As seen in the “Response structure” section of the results, the response provides multiple, individual steps rather than one compact answer. Moreover, the response ends with a recommendation to consult with a financial advisor— something that few comments do, as seen through the n-gram analysis. Analysis of the differences in structure reveal GPT-4’s current inability to provide opinion or specific financial advice, potentially because of safeguards implemented by OpenAI. GPT-4 is currently not a substitute for peer-to-peer financial advice, but is capable of providing general advice, information, and feedback.

### 5.2. Sentiment

Sentiment analysis was conducted in order to determine the emotional tone of GPT-4 responses and Reddit comments. Doing so allows for further comparison between GPT-4’s financial advice and peer-to-peer advice. As seen in Figures 3a and 3b, while both are predominantly neutral, GPT-4’s responses are more positive than Reddit comments, while Reddit has a higher proportion of negative sentiment comments. GPT-4 primarily responds with neutral sentiment, but may offer words of encouragement or praise. It never explicitly attacks or criticizes within its response. In contrast, some Reddit responses have negative sentiment, due to occasional critical nature of user comments. Responses with negative sentiment can convey disapproval, which can be important in financial advice settings. With only neutral or positive sentiment responses, GPT-4 can fail to emphasize financial risks or critique current financial decisions, showing a significant difference in its capabilities. GPT-4 responses, however, are able to present a multitude of different suggestions in a positive, unbiased manner.

### 5.3. General Analysis

Although this study does not explicitly evaluate GPT-4 responses for accuracy, it is able to evaluate the performance of

GPT-4 in peer-to-peer forums in financial spaces. GPT-4, while competent in presenting substantial financial advice, is not able to condense and deliver responses in the same structure and format as other users on the forum. At its current stage, GPT-4 does not directly present a risk towards such peer-to-peer advice forums due to its inability to provide direct advice, instead of referring to financial advisors, products, etc. Moreover, the responses collected from the GPT-4 demonstrate a clear pattern: GPT-4 consistently refers to financial advisors to discuss and make financial decisions. At its current capability, GPT-4 cannot access raw financial data, nor directly suggest recommended paths, and does not necessarily pose a direct threat to employment of financial professionals. However, for those with general questions, GPT can provide a starting point with strong ideas, and pathways to explore.

## 6. Study Limitations

While this study compared GPT responses to actual reddit user comments, it did not explicitly check for accuracy of the statements. Although some Reddit users on the forum may be financial experts, many are not. Hence, this study is focused on evaluating GPT-4’s performance in a forum-like financial setting, rather than in a direct financial advisory space similar to other previously conducted research. Moreover, while the similarity score provides an accurate assessment of similarity based on word vectors, it may miss identifying nuances within how words are employed. Similarity is subjective, and this score aims to match similarity in words, proving more useful for our task. Due to time and storage constraints, the dataset of Reddit posts collected for analysis remained limited. This size allowed for quick analysis using similarity and sentiment analysis, and allowed for clear empirical patterns to be identified in the dataset. As well as other fields, employing small sample set analysis serves both speed and efficiency, as well as closer manual analysis of data— however, the size of the dataset often restricts some nuanced or underdeveloped patterns to be clearly identified.

## 7. Summary and conclusions

This research explores the future of AI in financial spaces in terms of its capability to replace the role of a financial advisor as well as the implications that it will have on ethics and future employment in the industry. The GPT-4 LLM and the Reddit API, specifically financial questions from Reddit forum “r/Financial Planning” were used in order to test the GPT-4 LLM. The results were compared to actual Reddit comments, assessing the model’s response content, accuracy, and delivery. Through evaluation of the model’s advisory competency, it was determined that AI-driven algorithms can efficiently analyze vast volumes of data, make data-driven predictions, and optimize decision-making processes in the financial sector. However, while AI holds promise, it also poses significant challenges. The rapid adoption of AI in finance could threaten traditional employment, as certain tasks once performed by human

professionals, such as the role of financial advisors could potentially be automated. Additionally, there is a need to address concerns related to the trustworthiness and ethical implications of AI systems. For example, potential biases in AI are seen in the data from Figure 3a and 3b, they show how AI tends to give a strictly positive or neutral sentiment and almost a negligible amount of negative sentiment while giving financial advice, financial advice from humans on the other hand highlight show that both the negative and positive are take into account while giving advice. Biases like those spotlight the shortage of transparency in AI algorithms that could undermine people's self assurance in AI, raising questions about the reliability and fairness of financial recommendation provided via AI systems in comparison to human financial advisors. To cope with those challenges and make knowledgeable choices, further studies are crucial. Future studies should explore the level of acceptance that people place in AI structures as compared to human financial advisors, because it will assist researchers in recognizing the elements influencing customers' perceptions of AI financial advice. Furthermore, delving into how financial advisors feel about the combination of AI in their career is vital. Knowing their perspectives, worries, and experiences can be a useful resource in developing techniques to successfully integrate AI into their workflow and ensure a harmonious coexistence between AI and humans in financial spaces.

## References

- Adler, J., & Parmryd, I. (2010). *Quantifying colocalization by correlation: The Pearson correlation coefficient is superior to the Mander's overlap coefficient*. Cytometry. Part A: The Journal of the International Society for Analytical Cytology, 77A(8), 733–742. <https://doi.org/10.1002/cyto.a.20896>
- Biswas, S., Joshi, N., & Mukhopadhyaya, J. N. (2023). *ChatGPT in investment decision making: An introductory discussion*. Unpublished. <https://doi.org/10.13140/RG.2.2.36417.43369>
- Cao, L. (2020). *AI in finance: A review*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3647625>
- collections — Container datatypes. (n.d.). Python Documentation. <https://docs.python.org/3/library/collections.html>
- Deng, X., Bashlovkina, V., Han, F., Baumgartner, S., & Bendersky, M. (2023). *What do LLMs know about financial markets? A case study on reddit market sentiment analysis*. Companion Proceedings of the ACM Web Conference 2023.
- GPT-4 API general availability and deprecation of older models in the Completions API. (n.d.). Openai.com. <https://openai.com/blog/gpt-4-api-general-availability>
- Khader, M., Awajan, A., & Al-Naymat, G. (2019). *The impact of natural language preprocessing on big data sentiment analysis*. International Arab Journal of Information Technology.
- Linguistic features· spaCy usage documentation. (n.d.). Linguistic Features. <https://spacy.io/usage/linguistic-features>
- Loureiro, D., Barbieri, F., Neves, L., Anke, L. E., & Camacho-Collados, J. (2022). *TimeLMs: Diachronic language models from Twitter*. In arXiv [cs.CL]. <http://arxiv.org/abs/2202.03829>
- Neilson, B. (2023). *Artificial intelligence authoring financial recommendations: Comparative Australian evidence*. Journal of Financial Regulation, fjad004. <https://doi.org/10.1093/jfr/fjad004>
- OpenAI. (2023). *GPT-4 Technical Report*. ArXiv.Org. <https://doi.org/10.48550/ARXIV.2303.08774>
- Plisson, J., Lavrac, N., & Mladenic, D. (n.d.). *A Rule based Approach to Word Lemmatization*. Ijs.Si. <https://ailab.ijs.si/dunja/SiKDD2004/Papers/Pillson-Lematization.pdf>
- Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). *Studying Reddit: A systematic overview of disciplines, approaches, methods, and ethics*. Social Media + Society, 7(2), 205630512110190. <https://doi.org/10.1177/20563051211019004>
- Qin, H., Tian, Y., & Song, Y. (n.d.). *Relation extraction with word graphs from N-grams*. Aclanthology.org. <https://aclanthology.org/2021.emnlp-main.228.pdf>
- Rudin, P. (2023, April 4). *GPT4 & the impact of capturing our emotions*. Journal by GetAbstract. <https://journal.getabstract.com/en/2023/04/04/gpt4-the-impact-of-capturing-our-emotions/>
- Sarica, S., & Luo, J. (2021). *Stopwords in technical language processing*. PLoS One, 16(8), e0254937. <https://doi.org/10.1371/journal.pone.0254937>
- van de Schoot, R., & Miočević, M. (2020). *Small sample size solutions: A guide for applied researchers and practitioners*. Routledge.
- OpenAI. (2023). *GPT-4 Technical Report*. Retrieved August 10, 2023, from <https://cdn.openai.com/papers/gpt-4.pdf>