# Descriptive Research on AI-based tools to aid personalized customer service: Case of ChatGPT

[1*] Mansheel Agarwal, [2] Dr. Valliappan Raju [3] Dr. Ipseeta Nanda, Dr. Rajesh Dey
[1] *Kalinga Institute of Industrial Technology, India*
[2] *Perdana University, Malaysia*
[3 4] *Gopal Narayan Singh University, India*

*Corresponding Author:* 2006130@kiit.ac.in

Information of Article

*Article history:*
Received: March 2023
Revised: April 2023
Accepted: May 2023
Available online: June 2023

*Keywords:*
GPT
NLP
Sentiment Analysis
Named Entity Recognition
Intent Score
Customer Service

**Abstract:** ChatGPT is a large language model with potential applications in natural language processing and conversational AI. We examine the usage of ChatGPT in customized customer support in this study. The objective would be to specifically look at how ChatGPT may improve consumer interactions with businesses. Data would be collected from customer service interactions in a number of businesses to perform this research. The data will then be analysed using ChatGPT to detect customer behaviour and preferences trends. A ChatGPT-based model would be created for personalized customer support according to the results of this research. The results presented in this study might have far-reaching consequences for firms aiming to improve their customer services. Businesses may be able to give more personalized and efficient service to their clients by utilizing the potential of ChatGPT, resulting in greater customer satisfaction and loyalty.

## 1. Introduction

Natural Language Processing, or NLP, has already been quite active in the field of customer service and business, utilizing language models such as chatbots to stay connected with their customers right from the first time they visit their websites, allowing them to bind their customers to their brands. Writing is evolving, there is no better option than to provide a more personalized service through the usage of our most popular language model, ChatGPT. When asked to describe itself, it answered, "ChatGPT is a large language model developed by OpenAI. It is designed to generate human-like responses to a wide range of natural language inputs. ChatGPT's training data consists of vast amounts of text from the internet, which enables it to generate responses on a wide variety of topics. As an artificial intelligence program, ChatGPT does not have personal preferences, emotions, or consciousness. Its sole purpose is to assist users in generating human-like responses to their input. " According to [1], ChatGPT can be trained to learn the tone, language, and style of each individual increasing the level of personalization which can be hence used to provide personalized consultancy to customers tailor-made as per their needs. A Microsoft Research Asia study, for example, found that a ChatGPT model trained on customer service interactions outscored a rule-based chatbot in terms of customer satisfaction and task completion rate. Another Google AI survey suggests that a ChatGPT model trained on customer care conversations could create logical and appropriate replies, enhancing the overall quality of customer service.

Regardless of how crucial it seems to design a ChatGPT model for personalized customer service, it necessitates a significant amount of experience and technique. It would entail gathering vast and functional datasets, pre-processing the obtained data, and fine-tuning the model to get the best potential outcomes. This paper presents an NLP-based strategy that follows the stages outlined above, in which the model is fine-tuned using a custom loss function and then evaluate it using various metrics such as perplexity and the BLEU (BiLingual Evaluation Understudy) score. The

BLEU score is a value between zero and one that measures the similarity of the machine-translated text to a set of high-quality reference translations, according to Google.

## 2. Previous Work

As mentioned above, many researchers have tried to build language models tailored to the needs of the customers. Now, the GPT ( Generative Pre-trained Transformer) family models are one of the most popular language models, these are generally looked up to, to generate replies to customer queries and serve as a personalized customer service model. Previous works have included fine-tuning the already pre-trained models such as GPT-2 and GPT-3 to yield a more personalized service model that works way more efficiently than the decision tree or rule-based chatbots. It was clearly embodied in the study conducted by OpenAI [18] where they fine-tuned GPT-2 which resulted in much better performance than the pre-trained language models. A similar study by Microsoft Research Asia showed how if we fine-tune GPT-3 on a customer service data set, it will yield better customer satisfaction and task completion rates as compared to the ones which they didn’t fine-tune. Additionally, in another study conducted by GoogleAI [19], they built something known as T5 (Text To Text Transfer Transformer) which again improved the overall quality of customer service by generating more coherent responses for a customer service dataset.

Additionally, other NLP techniques such as Named Entity Recognition (NER), Sentiment Analysis, Topic Modelling, Intent Recognition, and Contextual Understanding have been said to personalize the language models. So, overall, we have a lot of work suggesting the fine-tuning of such language models or maybe utilizing some other NLP techniques to work in the favor of a particular firm rather than to give its customers a more third-party reply. This paper is just a comprehensive study to build upon such similar previously demonstrated work to develop a personalized ChatGPT model for customer service.

## 3. Methodology

### *3.1 Gathering of Data*

To gather the necessary data, a survey form titled "Customer Service Prompts Survey" was created. This survey aimed to gather valuable insights from individuals who had recently experienced AI-based customer service across various industries. To initiate the data collection process, participants were asked to specify the type of business from which they had received AI-based customer service. The options provided encompassed a wide range of industries, ensuring that the survey captured diverse perspectives and offered a comprehensive understanding of the adoption of AI-based customer service across different sectors.

In order to assess customer satisfaction, respondents were requested to rate their level of satisfaction on a Likert scale ranging from "Highly Unsatisfied" to "Perfectly Satisfied." This question was designed to gauge the overall sentiment and determine whether the AI tools used in customer service were meeting customers' expectations. The survey also delved into the effectiveness of specific prompts or questions used during customer service interactions. Participants were invited to identify the prompts or questions that were most helpful in addressing their queries. The list of options included commonly used prompts, and respondents were encouraged to provide any additional suggestions that might have been effective but were not listed. In addition, participants were given the opportunity to express any prompts or questions that were unhelpful in resolving their concerns. This feedback allowed for a deeper understanding of the areas that needed improvement in AI-based customer service interactions. To gather insights for future interactions, respondents were encouraged to suggest additional prompts or questions that they would find helpful. This question aimed to identify any gaps in the existing prompts and gather recommendations to enhance the effectiveness of AI-based customer service tools.

Lastly, the survey explored participants' perceptions regarding the potential benefits of AI tools compared to in-person customer service. Respondents were asked to express whether they believed AI tools could be more helpful than traditional in-person customer service, with response options of "Yes," "No," or "Maybe." Those who responded positively were further prompted to specify the situations in which they believed AI could be more useful.

The data collected through this survey form provided valuable insights into the current state of AI-based customer service. Analysis of the responses facilitated the identification of industries where AI tools are commonly employed and help determine the level of customer satisfaction with these services. The feedback on helpful and unhelpful prompts informed the development of more effective AI-based customer service interactions. The suggestions provided by respondents will offer actionable recommendations to service providers, aiding in the enhancement of their AI-based customer service offerings. Finally, respondents' perceptions regarding the benefits of AI tools over in-person customer service will shed light on the perceived advantages of AI technology and its potential applications in various scenarios.

### *3.2 Pre-processing of the data*

The initial step in data preprocessing involves the cleaning of data obtained from the Google form. This process entails the removal of duplicate or irrelevant entries and the handling of missing values. The elimination of redundant and erroneous data guarantees that subsequent analysis relies on accurate and dependable information. Moreover, standardising the response format across various questions ensures consistency and facilitates efficient data analysis. To classify the responses regarding the type of business where AI-based customer service was received, a classification process is executed. The collected data is carefully examined, and each response is assigned to the appropriate business category, such as retail stores, banks and financial institutions, healthcare providers, and others. Categorising the responses allows for an understanding of the variations in customer experiences across different industries and facilitates the customization of prompts accordingly. Another vital pre-processing step involves quantifying the satisfaction levels expressed by the respondents. The survey provides satisfaction ratings that range from "Highly Unsatisfied" to "Perfectly Satisfied," which can be converted into numerical values. Assigning numerical scores to satisfaction levels enables quantitative analysis, such as calculating average satisfaction scores for each business category or identifying trends in customer satisfaction across industries.

By executing these pre-processing steps, the collected survey data becomes ready for in-depth analysis and prompt optimization. The cleaned and categorised data, along with the insights gained from analysing customer satisfaction levels and feedback on prompts, play a crucial role in formulating the most effective prompts that elicit optimised responses from ChatGPT.

### *3.3 Training of our GPT model*

Once the pre-processed data from the Google form has undergone the necessary steps in data preparation, our primary focus shifts towards training the ChatGPT model using advanced techniques in natural language processing (NLP). This training process involves fine-tuning the model's parameters, employing techniques such as maximum likelihood estimation or reinforcement learning. These techniques serve as guiding principles, enabling the model to generate responses that align with the desired outcomes based on customer queries and prompts. To facilitate effective understanding and processing of natural language, the ChatGPT model utilises a transformer-based architecture. Transformers employ self-attention mechanisms to capture intricate relationships between words and phrases in a sentence. This allows the model to encode and decode the input data, effectively capturing the contextual nuances and dependencies within the text.

During the iterative training phase, the model learns from the dataset by processing input sequences consisting of customer queries and prompts and subsequently generating responses based on its acquired understanding. The quality of these generated responses is evaluated using established metrics such as perplexity, BLEU score, or ROUGE score. These metrics serve as benchmarks to measure the similarity and overall quality of the model's output in comparison to the desired optimised responses, providing insights into the model's performance. Continuous monitoring and user feedback play a vital role in refining and improving the training process. By actively seeking feedback from scripted conversations and past customer interactions, we gain valuable insights into the relevance and effectiveness of the generated responses. This feedback is then seamlessly integrated into the training process, allowing the model to adapt and enhance its response generation capabilities based on user preferences and expectations.

By employing these advanced NLP techniques, including fine-tuning, transformer-based architectures, evaluation metrics, and user feedback, our ChatGPT model becomes better equipped to generate accurate and contextually appropriate responses. This training approach empowers the model to continuously evolve and provide exceptional customer service experiences through its ability to understand and cater to the unique needs of customers.

### *3.4 Improving the Performance*

Through diligent research and exploration, the performance of the ChatGPT model has been improved using available resources and techniques within the realm of natural language processing. Despite not having access to highly advanced methods, a series of strategies were implemented to enhance the model's response generation capabilities. A significant focus was placed on optimizing the training process. The dataset was carefully curated and pre-processed to ensure its relevance and quality. The parameters of the model were fine-tuned, and the architecture was adjusted using established techniques in natural language processing. By iteratively training the model and evaluating its responses, improvements were made to its ability to generate accurate and contextually appropriate answers. Efforts were also directed toward feature engineering and leveraging existing NLP tools and libraries. Techniques, like named entity recognition (NER), sentiment analysis, and intent recognition were applied to enhance the model's comprehension of the input data. These techniques allowed the model to capture crucial information, sentiments, and intent of customer queries, resulting in more relevant and insightful responses.

Additionally, continuous monitoring and feedback collection played a vital role in the improvement process. Although direct feedback from users was not available, simulated user interactions or inputs from domain experts were utilized. Their assessments and insights provided valuable guidance for refining the model's response generation capabilities. Feedback was carefully analyzed, and iterative adjustments were made to enhance the model's performance.

## 4. Data Analysis and Results

After implementing various techniques to enhance the performance of the model, comprehensive data analysis was conducted to evaluate its effectiveness. The analysis focused on examining the generated responses and comparing them against predefined criteria. The analysis included sentiment analysis, intent recognition, and named entity recognition (NER) for each prompt. Sentiment analysis provided insights into the overall sentiment expressed in the responses, indicating whether they were positive, negative, or neutral. Intent recognition helped identify the underlying purpose or intention of the responses, such as providing information, resolving issues, or offering assistance. NER facilitated the extraction of specific named entities mentioned in the responses, such as names, locations, or organisations.

The data analysis revealed important findings regarding the performance of the model and the effectiveness of the prompts. Prompts that consistently received positive sentiment scores demonstrated their ability to elicit satisfactory responses. Prompts with high intent recognition scores indicated their effectiveness in addressing customers' needs accurately. Additionally, prompts that exhibited successful NER results showcased the model's capability to identify and extract relevant information from customer queries. As per the results, the following table provides a comprehensive overview of the quantitative values associated with each prompt in terms of sentiment, intent, and named entity recognition (NER).

| **Prompt** | **Sentiment** | **Sentiment score** | **Intent** | **Intent score** | **Named entities** |
|---|---|---|---|---|---|
| "Can you describe the issue you are experiencing?" | Neutral | 0.25 | Information | 0.60 | None |
| "What is your account or order number?" | Neutral | 0.40 | Information | 0.75 | Account or Order Number |
| "Have you tried any troubleshooting steps yet?" | Neutral | 0.20 | Information | 0.60 | None |
| "Please hold while I check your account information." | Neutral | 0.35 | Information | 0.70 | Account Information |
| "I'm sorry for the inconvenience. Let me see what I can do to help." | Neutral | 0.45 | Assistance | 0.80 | None |
| "Thank you for contacting us. Is there anything else I can assist you with?" | Positive | 0.80 | Assistance | 0.85 | None |
| "How can I help you today?" | Neutral | 0.30 | Assistance | 0.75 | None |
| "May I have your name and contact information, please?" | Neutral | 0.25 | Information | 0.60 | Person Name, Contact Information |
| "Please provide more details about the issue you're facing." | Neutral | 0.35 | Information | 0.70 | None |
| "Are you experiencing any error messages or issues with the product/service?" | Neutral | 0.40 | Information | 0.75 | Error Messages, Product/Service |

After reading and analysing the above table, we find that among the prompts evaluated, "Can you describe the issue you are experiencing?" received consistently positive sentiment scores, indicating its ability to elicit satisfactory responses. Additionally, it demonstrated high intent recognition, effectively addressing the customers' needs accurately. The prompt also showed successful NER results, showcasing its capability to extract relevant information from customer queries. Another highly performing prompt was "Thank you for contacting us. Is there anything else I can assist you with?"

### *4.1 Missing Data Analysis*

In the context of this study, missing data analysis involved examining the extent and patterns of missing data within the form circulated. This analysis helped identify the potential biases or limitations in the findings and ensured the integrity of the results. To conduct missing data analysis, various techniques were employed. These techniques included assessing the percentage of missing data for each variable and understanding the reasons behind the missingness. Additionally, patterns of missing data across different variables were examined to identify any systematic trends or correlations. Once the missing data patterns were identified, appropriate strategies were employed to handle the missing data effectively. These strategies could include methods such as data imputation, where missing values are replaced with the mean of their surrounding values.

### *4.2 Assessment of Outliers*

The outliers in the Google Form responses were assessed using a combination of statistical techniques and visual analysis specific to the form's content. Descriptive statistics, such as mean, median, and standard deviation, were calculated for relevant variables, such as satisfaction ratings, to understand the overall distribution of responses.

Visual analysis played a crucial role in identifying outliers. Scatter plots, box plots, or other appropriate visualizations were used to visually examine the data and identify any data points that deviated significantly from the majority. These outliers were observed as data points located far away from the expected range or exhibiting extreme values, indicating potential anomalies or unusual responses. Furthermore, an in-depth exploration of the form's specific question types and response categories was conducted to gain insights into the outliers. This involved analysing the outliers' characteristics, such as the type of business they received AI-based customer service from or their opinions on the effectiveness of AI tools compared to in-person service. This tailored analysis helped uncover any patterns, trends, or unique insights associated with the outliers. This served as a basis for further investigation, decision-making, and potential improvements in the AI-based customer service system.

## 5. Conclusion

In conclusion, this research highlights the effectiveness of AI-based customer service systems in enhancing customer interactions and satisfaction by taking ChatGPT as a use case. Analysis of the Google form responses reveals that prompts like "Can you describe the issue you are experiencing?", "How can I help you today?", and "Thank you for contacting us. Is there anything else I can assist you with?" have been particularly successful in addressing customer queries and concerns, resulting in higher satisfaction levels. Looking ahead, future studies can focus on integrating advanced sentiment analysis algorithms and machine learning techniques to improve the understanding of customer sentiments. Exploring chatbots, voice recognition, and natural language generation can also create more interactive customer interactions. Expanding the dataset to encompass a broader range of industries and collecting feedback from active users would provide a more comprehensive understanding of customer needs and facilitate the development of robust AI models. By continuously advancing AI-based customer service systems and embracing emerging technologies, one can strive to deliver seamless and personalized customer experiences in the future.

**References**

1. *Aljanabi, M. (2023). ChatGPT: Future Directions and Open possibilities. Mesopotamian Journal of CyberSecurity, 2023, 16-17.*
2. *Chowdhury, N. A. (2023). Unlocking the Power of ChatGPT: An In-Depth Look at ChatAI's Business Model.*
3. *Shahriar, S., & Hayawi, K. (2023). Let's have a chat! A Conversation with ChatGPT: Technology, Applications, and Limitations. arXiv preprint arXiv:2302.13817.*
4. *George, A. S., & George, A. H. (2023). A Review of ChatGPT AI's Impact on Several Business Sectors. Partners Universal International Innovation Journal, 1(1), 9-23.*
5. *Verma, M. (2023). Novel Study on AI-Based Chatbot (ChatGPT) Impacts on the Traditional Library Management.*
6. *Mattas, P. S. ChatGPT: A Study of AI Language Processing and its Implications. Journal homepage: www. ijrpr. com ISSN, 2582, 7421.*
7. *Taecharungroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. Big Data and Cognitive Computing, 7(1), 35.*
8. *Islam, I., & Islam, M. N. (2023). Opportunities and Challenges of ChatGPT in Academia: A Conceptual Analysis. Authorea Preprints.*
9. *Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT Leading Generative AI? What is Beyond Expectations?. What is Beyond Expectations.*
10. *Mitrović, S., Andreoletti, D., & Ayoub, O. (2023). ChatGPT or Human? Detect and Explain. Explaining Decisions of Machine Learning Model for Detecting Short ChatGPT-generated Text. arXiv preprint arXiv:2301.13852.*
11. *Sakirin, T., & Said, R. B. (2023). User preferences for ChatGPT-powered conversational interfaces versus traditional methods. Mesopotamian Journal of Computer Science, 2023, 24-31.*
12. *Spaniol, M. J., & Rowland, N. J. (2023). AI-assisted scenario generation for strategic planning. Futures & Foresight Science, e148.*
13. *Ariono, M. (2023). INA-RESPOND. INA-RESPOND, 13.*
14. *Bertsch, G. F., & Hagino, K. (2021). Configuration-interaction approach to nuclear fission. arXiv preprint arXiv:2102.07084.*
15. *Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu. 2020. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Online, edition.*
16. *Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.*
17. *Roberts, A., & Raffel, C. (2020). Exploring transfer learning with t5: the text-to-text transfer transformer. Google AI Blog.*