

# **Improving a Customer Support Chatbot**

A Technical Report submitted to the department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

**Naomi Nichols**

Fall 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Briana Morrison, Department of Computer Science

Rosanne Vrugtman, Department of Computer Science

# Improving a Customer Support Chatbot

CS4991 Capstone Report, 2023

Naomi Nichols  
Computer Science  
The University of Virginia  
School of Engineering and Applied Science  
Charlottesville, Virginia USA  
nln5awc@virginia.edu

## ABSTRACT

As a Product Management (PM) Intern, I collaborated with a team of developers to analyze customer satisfaction and suggest improvement areas for an AI-based customer support chat bot used by thousands of customers. Leveraging SQL data analysis, customer interviews, and internal and external research, I determined that improving the accuracy of the answers generated by the bot and redesigning the self-service and issue escalation menu would have the greatest positive impact on customer satisfaction. I refined these suggestions with the development team and then outlined a strategy document to describe multiple approaches to improving these areas. As the development team followed my release plan, I analyzed usage data to determine potential future feature upgrades. The improvements made to the chat bot will assist customers in using the company's product by delivering fast and reliable answers, as well as reduce customer support costs while laying the foundation for future innovation. As my internship concluded, the development team continues to work on the improvement areas I outlined and is investigating implementation methods for my feature suggestion.

## 1. INTRODUCTION

To reduce customer support costs associated with employing a full team of customer support staff, my company investigated the potential of leveraging AI to

help solve users' issues with an internal pipeline product. The pipeline product helps enforce security and compliance requirements of builds as set by company policy, government regulations, and technical standards. Software engineers configure the tool for their project and leverage it iteratively as they develop.

Customer support for the pipeline tool previously consisted of three methods: referencing documentation, posting a question to the company's Stack Overflow site, and submitting a ticket to the support staff. Developers have access to tens of pages of documentation regarding the usage and customization of the tool, but as the tool has evolved, much of this documentation can be inaccurate, confusing, or misleading. Stack Overflow questions are monitored by support staff but also open for other developers to weigh in on common issues, leading to diverse collaboration across the organization. Finally, the support tickets are answered by a full-time, contracted support staff member who typically looks at an issue within two days. While these methods eventually resulted in resolving issues, there was a noticeable lack of real-time support.

The chatbot was originally developed by a volunteer team as an entry for a company-sponsored hackathon. Leveraging a large language model trained on the pipeline tool's documentation, the chatbot integrated with the company's chat software for an intuitive customer experience. Shortly after its initial

development, the chatbot was launched as the entry point for customer support. A notable change from the previous support experience was that users could only escalate issues to the support staff through the bot rather than the previously direct method.

## **2. RELATED WORKS**

A number of firms have leveraged AI-based chatbots for real-time customer support. With the introduction of tools such as ChatGPT, these chatbots can be trained on a topic area with relative ease and leverage features like topic and emotion detection to create a more natural customer service experience (George, et al., 2023).

While AI technology in this area is promising, researchers warn about limitations of these large language models because of a phenomenon called hallucinations. AI chatbots like ChatGPT respond to factual questions with convincing replies, but these replies may or may not be based on facts. Responses containing hallucinations cannot be verified by source material but rather are “guesses” based on the provided information due to knowledge gaps within the training data (Bang, et al., 2023).

## **3. PROJECT DESIGN**

In order to assess customer satisfaction with the chat bot and propose improvements and new features, I leveraged customer interviews, SQL data analysis, and internal and external research.

### **3.1 Customer Interviews**

I began by interviewing a variety of users of the pipeline product, developers working for my company, not only to familiarize myself with their customer support experience, but also to learn more about the product itself and what types of issues developers face when using it.

The interview participants were users who had completed the semi-annual product

feedback survey and indicated a willingness to be contacted for further conversations. I met with six customers and discussed their experience with the pipeline product and the customer support bot, as well as asked for their suggestions for areas of improvement.

Synthesizing the common themes and experiences from a variety of users helped contextualize the issues with our customer support approach and provided the foundation of my recommendations for improvement. The most common themes from these interviews were major issues with bot hallucinations, confusion regarding the process for submitting a ticket to the support team, and difficulty navigating relevant documentation.

### **3.2 Data Analysis**

Using my insights from the customer interviews as a guide, I next analyzed conversation transcripts between users and the bot as well as customer support tickets to further refine my recommendations for improvement.

To identify key areas of dissatisfaction in the bot’s responses, likely due to hallucinations, I leveraged SQL to analyze conversation transcripts from users’ interactions with the chat bot. The chat bot interface allows users to review the bot’s answers with either a thumbs up or a thumbs down. Using this data, sorted by topic, I was able to infer which areas users required the most support in and how satisfied they were with how the bot addressed their questions. A summary of these findings is shown below in Figure 1.

Topics	Percent of Total Sessions Related to Topic	Percent of Related Sessions Earning Positive Feedback	Percent of Related Sessions Earning Negative Feedback
Pipeline Builds	35%	17%	22%
Framework	11%	28%	39%
Verification	8%	29%	36%

Figure 1: Transcript Analysis Results Summary

Additionally, I reviewed the customer support tickets that were filed since the support chatbot was released. These tickets were filed through the bot, meaning that the users were unable to get their issue resolved by the bot and thus, the issue was escalated to the support team to address. Investigating what types of questions the bot is unable to answer could help inform the improvements or new functionality can be added to better support the customer.

Looking at the categories of support tickets filed and reading through customer issues, it was quickly apparent that users were facing widespread and specific issues that needed individualized support. However, the highest-impact category for future functionality to address is permissions.

### 3.3 Additional Research

In addition to interviewing customers, I also interviewed teams within the company pursuing similar endeavors. Developers on both teams then met to discuss our implementations in more detail. This meeting helped both teams notice some differences in our respective AI chatbots implementations and results.

Members of the chatbot development team also received multiple emails from customers regarding the chatbot. Such feedback was also considered in forming the improvement recommendations.

### 3.4 Recommendations

My recommendations for improving the customer experience with the chatbot focused on three areas: answer accuracy, self-service and issue escalation menu, and permissions functionality.

The most pressing complaint customers voiced with the chatbot was the inaccuracy of the answers it provides due to the AI hallucinations. To address these, I developed three improvement paths, shown below in Figure 2. Because the chatbot is trained on our pipeline product's documentation, ensuring the accuracy and relevance of the documentation is imperative to improving quality of chatbot's responses and ultimately, its usefulness to the customer.

Improvement Path	Details
Documentation	Ensure that information is current, concise, and clear, including reflecting recent version deprecation of foundational technology.
	Prioritize documentation review and revision for accuracy and bot-friendliness: <ul style="list-style-type: none"> <li>• Pipeline Builds</li> <li>• Framework</li> <li>• Verification</li> </ul>
Bot Configuration	Leverage internal resources to refine the strategy for training the bot.
	Include documentation sources within bot responses.
Customer Education	Inform customers of best practices for formatting questions.
	Increase clarity of greeting message to better inform users of the chatbots use and capabilities.

Figure 2: Answer Accuracy Improvement Paths

The next recommendation I made for improving the chatbot was redesigning the self-service and issue escalation menu. By presenting clearer options and making the flow for escalating tickets more intuitive to the user, the concerns brought up within the customer interviews and emails can be addressed.

The final recommendation I made to improve the chatbot was to add functionality to address permissions issues. Many users had issues setting up appropriate editing permissions for their team pipeline, but the functionality can be added to the bot to automate the checking and granting of appropriate permissions upon user request.

Ultimately, each of these recommendations was informed by the customer and data analysis I conducted in order to have the highest impact on customer satisfaction regarding our pipeline product's customer support.

#### **4. ANTICIPATED RESULTS**

The recommendations I suggested were well received by both developers and PM leadership. As my internship concluded, the chatbot development team continued to work on the improvement areas I outlined and investigate implementation methods for my feature suggestion. With these improvement areas addressed, the customer experience with the pipeline product will be improved and the costs associated with customer support will be reduced.

#### **5. CONCLUSION**

As users are working with the pipeline product, their efficiency relies on timely and accurate customer support. Making the improvements previously discussed will help achieve both of those goals for our customers, ultimately improving the way work is done across the company.

#### **6. FUTURE WORK**

Once development is complete on the suggested improvements and features, the development team will continue to monitor the evolving customer needs in order to provide effective customer support for the pipeline product.

Establishing the validity of this customer support chat bot will further lay a

foundation that could be used across the company for internal customer service for various products.

#### **REFERENCES**

A. Shaji George, & A. S. Hovan George. (2023). A Review of ChatGPT AI's Impact on Several Business Sectors . *Partners Universal International Innovation Journal*, 1(1), 9–23. Retrieved October 8, 2023 from

<https://doi.org/10.5281/zenodo.7644359>

Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V., Xu, Y., & Fung, P. (2023). *A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity*.