Automation of Customer Queries
for TBH Labs via a Rule-based Chatbot

Author: Aung Cham Myae – 6411325 , Tanat Arora - 6410381

Date: 26th September 2023

Abstract

This report presents the development and evaluation of a rule-based chatbot designed to handle large volumes of customer queries for TBH Labs, a tech review Facebook page. Leveraging the capabilities of wit.ai for intent recognition and entity extraction, the system has been further enhanced by integrating it with the gsmarena API for fetching phone specifications. However, the system faces limitations due to the capped number of API calls permissible. The solution proposed here is the creation of an internal database to address these constraints and improve scalability.

Introduction

TBH Labs is a tech review page on Facebook, boasting significant engagement from its user base. Due to its vast outreach, TBH Labs receives thousands of messages daily. While it's a testament to the page's popularity, managing this volume of messages manually poses challenges. Hence, an automated system was conceived to streamline the process and offer timely responses.

Methodology

Wit.ai Integration

The chatbot leverages wit.ai, a renowned platform known for Natural Language Processing (NLP). This aids in discerning user intent and extracting key entities from the queries. Three primary intents identified were:

Phone Specification: Users seeking detailed information about a specific phone model.

Comparison: Users comparing two or more devices to make an informed decision.

Price: Queries related to the current market price of devices.

Glitch Platform

The "glitch" platform was chosen to manage and handle the user intents. It serves as the middleware, effectively processing the parsed data from wit.ai and further executing the appropriate function, such as fetching data from the gsmarena API.

GSM Arena API Integration

For authentic and updated phone specifications, the gsmarena API was incorporated. It's a comprehensive database that provides detailed information about various mobile phones, ensuring users receive accurate responses.

Role of Wit.ai in Chatbot Development

Wit.ai has been instrumental in the development and refinement of our chatbot, particularly in understanding and processing user queries. One of the most noteworthy features of wit.ai is its pre-trained model that supports multiple languages, including Burmese. This ability to recognize and process queries in Burmese significantly broadens the chatbot's utility, especially for a diverse audience.

The advantage of relying on wit.ai's pre-trained model lies not only in its linguistic capabilities but also in the reduced time and effort required for training. With it, we bypass the otherwise intensive process of training a model from scratch for a specific language.

Behind the Scenes: The NLP Techniques Powering Wit.ai

Understanding the background workings of wit.ai requires a deep dive into the realms of Natural Language Processing (NLP). Wit.ai employs an array of sophisticated NLP techniques to achieve its commendable proficiency in intent recognition and entity extraction.

1. **Named Entity Recognition (NER**): NER is a core technique used by wit.ai. It involves identifying and classifying named entities (like names, places, dates) from the input text into predefined categories. For instance, when a user queries about the specifications of a particular phone model, NER enables the identification of the phone's name and subsequently fetches its details.
2. **Machine Translation:** To cater to multiple languages, wit.ai utilizes machine translation. This allows the platform to automatically translate queries from non-English languages (such as Burmese) into English, process them, and then translate the responses back. The result is a chatbot that can effectively communicate across a myriad of languages.
3. **Intent Classification:** This involves determining the user's purpose behind a query. Using deep learning models, wit.ai discerns the underlying intent, be it seeking phone specifications, price inquiries, or comparisons.

4. **Dependency Parsing**: Wit.ai analyzes the grammatical structure of a sentence to understand relationships between words. This aids in more accurately capturing the essence of user queries, especially complex ones.
5. **Contextual Understanding:** Leveraging recurrent neural networks (RNNs) and transformer architectures, wit.ai processes queries in context rather than in isolation. This means the system understands the broader narrative or conversation, ensuring more accurate and contextually relevant responses.
6. **Continuous Learning and Feedback Loop:** As users interact with the chatbot, wit.ai captures these interactions. Mistakes and corrections made are fed back into the system. Over time, this iterative feedback refines the model's accuracy and understanding, making it more robust and reliable.

In essence, wit.ai offers a robust NLP foundation that powers our chatbot, allowing it to interpret and respond to a diverse range of user queries with remarkable efficiency and accuracy.

Sovereignty through Database Ownership

Controlling our database offers unparalleled flexibility. When a system's data source is proprietary, it opens up a multitude of possibilities. It allows for custom tweaks, personalized categorizations, and rapid incorporation of user feedback. Simply put, owning our database would enable us to tailor the chatbot's responses more finely, ensuring a high level of user satisfaction and improved engagement.

Rationale Behind Incorporating GSMarena API

While the idea of creating our proprietary database for phone specifications was considered, the choice to integrate the GSMarena API in the initial phase was driven by practicality. We aimed to test the capabilities and limitations of our chatbot without getting mired in the intricacies of database development from the outset.

Developing a comprehensive database from scratch is a time-intensive venture. Given the rapid rate at which new phones are launched and specifications are updated, maintaining such a database poses substantial challenges. Moreover, GSMarena offers a vast array of phone specifications, something that would have been resource-intensive for us to replicate in the early stages.

Our choice to incorporate the GSMarena API, therefore, was a strategic decision to expedite the chatbot's rollout and to gauge its performance. It was a way to balance time efficiency with providing accurate and comprehensive information.

Challenges and Limitations

Despite the sophisticated design, the system faces challenges:

Volume of Queries: TBH Labs' vast user engagement means that the bot often has to deal with more queries than it can handle.

API Call Restrictions: The gsmarena API has a limitation on the number of calls that can be made within a specific timeframe. This poses a bottleneck, especially during peak engagement times.

Proposed Solution: Building an Internal Database

To circumvent the limitations posed by external API reliance, it's proposed that TBH Labs develop its own internal database. This would entail:

Data Collection: Collate phone specifications from various trusted sources, ensuring data accuracy and comprehensiveness.

Database Maintenance: Regularly update the database with new phone models, ensuring specifications are current.

Integration with the Chatbot: Modify the existing middleware on the "glitch" platform to fetch data from this internal database instead of the external API.

Benefits of the Proposed Solution

Scalability: By owning and managing the database, TBH Labs can ensure that no artificial limits are imposed on the number of queries handled.

Relevance: With direct control over the data, TBH Labs can ensure that the most sought-after information by its user base is readily available.

Conclusion

TBH Labs' initiative to automate customer responses using a rule-based chatbot signifies a step forward in enhancing user engagement and experience. The proposed internal database solution aims to further this objective by ensuring scalability and relevance in the face of increasing user engagement. Once implemented, TBH Labs can truly offer a more seamless and informative experience to its already existing vast user base.

Material: https://github.com/nordmarin/gsmarena-api