Find Your Perfect Car

David Del Moral
College of Engineering and Computer Science
Texas A&M University-Corpus Christi
Corpus Christi, TX, US
ddelmoral@islander.tamucc.edu

Ean Shi
College of Engineering and Computer Science
Texas A&M University-Corpus Christi
Corpus Christi, TX, USA
eshi@islander.tamucc.edu

Matthew Tomme
College of Engineering and Computer Science
Texas A&M University-Corpus Christi
Corpus Christi, TX, US
mtomme@islander.tamucc.edu

Ricky Godsey
College of Engineering and Computer Science
Texas A&M University-Corpus Christi
Corpus Christi, TX, US
rgodsey@islander.tamucc.edu

Zachary Kao
College of Engineering and Computer Science
Texas A&M University-Corpus Christi
Corpus Christi, TX, USA
zkao@islander.tamucc.edu

***Abstract*—In this proposal we suggest and explore the development of an AI chatbot designed to assist consumers in searching for the vehicle that best suits their wants and needs. We propose the development of an AI chatbot that leverages natural language processing (NLP) and machine learning algorithms coupled with a comprehensive database of vehicles to engage users in a natural, conversational interface that can understand their preferences and provide personalized car recommendations. The goal is to simplify the car selection process and reduce the stress and burden placed on buyers. We explore related solutions and propose potential approaches to creating a user-centric and efficient automotive search process.**

I. Introduction

In the modern automotive market, consumers are overwhelmed by choices. There is a myriad of car models with many features and specifications. In addition to different manufacturers, models, price, and fuel efficiency, modern car buyers need to decide fuel types, autonomous assistance features, connectivity, and infotainment. This complexity can lead to stress that can lead to dissatisfied purchases. With the advancements of artificial intelligence (AI) and natural language processing (NLP), there is an opportunity to enhance the car-buying experience by using artificial intelligence that can answer individual needs.

This paper proposes the development of an AI-assisted chatbot designed to guide users in finding a car that best suits their specific wants and needs. The central problem addressed by this research is the difficulty consumers encounter in navigating the vast array of vehicle options and making informed choices. Traditional methods of car shopping often involve time-consuming research and reliance of generic advice, which may not fully align with an individual's unique lifestyle, budget, and preferences. The chatbot helps the user by simplifying the decision-making process using user inputs to deliver sensible recommendations from a comprehensive database of vehicles.

The support of artificial intelligence in car-buying goes beyond mere convenience; it saves money and reduces stress to search for a car that can improve satisfaction. By employing AI technologies, this proposal seeks to contribute to transforming the automotive sales experience into a confident encounter.

II. Related Works

As AI chatbots and database search engines are not novel ideas, we will discuss related works that we can draw inspiration from.

A. Akinator

Akinator is a web game that employs machine-learning and a vast database of knowledge to deduce fictional and non-fictional characters and objects [1]. It is similar to the game *Twenty Questions*, in which players ask each other simple yes or no questions to deduce what the other player is thinking about. This process of using user input to narrow down possible results based on a database of knowledge serves as the basis for our proposed research. The problem with an *Akinator* style approach is that *Akinator's* database is built by users and adjusted by user input. This data provided by users can potentially be incorrect, so we propose the use of a carefully curated and comprehensive database to ensure accurate data for our chatbot to draw information from.

B. ChatGPT

ChatGPT is a popular AI chatbot that uses NLP to generate human-like conversations and written content [2]. It leverages a generative pre-trained transformer (GPT), a type of large language model (LLM), that is fine-tuned for conversational applications. It produces human-like responses through a combination of supervised learning and reinforcement learning from human feedback. *ChatGPT* provides a user-friendly and conversational interface for users to interact with that we would like to adopt. However, although *ChatGPT* can be used to search for cars, it is not possible to verify the validity of the information it provides. *ChatGPT* is notorious for potentially providing inaccurate or false information, known as hallucinations [6]. For our

purposes, we would like to ensure that all information provided by our chatbot is accurate and citable.

C. CarGPT

CarGPT is a website that allows a user to enter features, styles, or other preferences for their ideal car along with their budget to find cars that match their criteria using AI to provide results [18]. This is a good representation of our ideal implementation; however, it does not provide an unexperienced user with an adequate number and type of categories to choose from right away which could lead to user dissatisfaction. The website also only uses the euro to represent currency, which can cause confusion for people who do not live in an area that utilizes the euro and forces users to convert from their native currency to euros to use the application. We would like to ensure that these are not going to be problems that our users encounter in our program.

D. Traditional Search Methods

Traditional search methods for car buyers involve manual searches, the use of websites such as cars.com [4] to find available cars nearby, and a reliance on car dealership sales representatives to find information and narrow down choices. These options either place the burden of the search on the consumer or a sales representative who may not have the consumer's best interest as their priority.

E. Scholarly Research

In this section we will discuss previous research related to our proposed research and how it can be built off or applied to our research.

1) Automobile Consumer Search Behavior

The proposed chatbot aims to provide value to consumers by decreasing customer dissatisfaction through improving the search process. To achieve this, research into vehicle consumer behavior and satisfaction must be done. A relevant study done by Punj in 1983 [12] explores the relationship between searching for vehicle information and customer satisfaction. Their research made three major conclusions. First, they found that people with less prior general car knowledge found greater benefit in searching for car information. Second, they found that search activity indirectly increases customer satisfaction by providing a means to achieve a better purchase. Third, they found that less prior car knowledge a person had, the more search activities they performed.

We can draw several conclusions from this research that will provide a basis for our proposed chatbot to improve customer satisfaction. As consumers with a large prior base of car knowledge are less likely to need to search for information and will still have the information necessary to make a satisfying purchase, we will be targeting consumers with a small prior base of car knowledge. These consumers will benefit from searching for information, but as Punj's research showed, consumers do not directly draw satisfaction from the search activity itself. Customer satisfaction is gained from the application of their knowledge to make an informed and good purchase. Therefore, we aim to increase the efficiency of the vehicle information search. By lowering the amount of searching a customer must do and increasing the amount of useful car information they draw from the search activity, we can increase the efficiency of customer satisfaction.

2) Personalized Product Searches

Online product search is not confined to just retrieving products relevant to the consumer's search. It also focuses on finding items that satisfy the consumer's individual needs and preferences. This makes the personalization of product searches valuable to both consumers and producers. We aim to provide personalized searches through our chatbot, so relevant research into this field is applicable to us. A relevant study by Ai in 2017 [13] proposed the use of a hierarchical embedding model using a deep neural network to improve personalized product searches. Their research found that their hierarchical embedding model for personalized product search significantly outperformed the state-of-the-art base-lines on Amazon benchmarks. This research indicated that personalization of searches is fruitful. By using an AI chatbot to assist consumers in searching, we can leverage the personalization inherent in the technology to increase the effectiveness of vehicle searches.

3) Inquisitive Chatbots

Chatbots can respond to natural language input, but in many scenarios, the user's query may not contain adequate information to provide an answer. In these cases, the chatbot needs to be inquisitive so that it can be interactive and collect the data required to answer the query. A relevant study by Sadhana in 2016 [14] researches this topic. This study achieved an inquisitive chatbot by using a chat engine that can identify what types of information is missing, which can then be obtained by issuing inquisitive queries. We propose a similar approach in the development of our chatbot by utilizing a database, from which missing information can be identified.

4) Chatbots for Searching

In our research, we found several instances of chatbots being employed to improve upon traditional search methods. One relevant study was done by Cantador in 2021 [15] on the impact of using a chatbot to search for open government data. This chatbot was used by non-expert users and found that the proposed system outperformed traditional search methods.

Another relevant study by Pitel in 2024 [16] utilized popular AI-assisted search engines such as ChatGPT to investigate the accuracy of queries in relation to cancer variant interpretation. They found that these popular and public search engines tended to falsely overestimate cancer variant interpretation. They concluded that AI-assisted search engines are a useful supplement in a highly specialized field, but that it required expert human oversight. To account for this, we propose the use of a expertly curated database for our chatbot to use as its knowledge base so that it does not draw false conclusions from past information.

III. Proposed Approaches (Tentative)

We propose the development of an AI chatbot that will leverage NLP and a comprehensive vehicle database to provide users with a user-friendly solution to searching for a car that suits their needs and preferences. To achieve this, we propose an incremental approach to development. The first priority is the creation of a chatbot that uses simple yes and no answers to prompts to narrow down the search results from a database. Once this basic model of the chatbot is created, we can proceed to implementing more complex and user-friendly AI features similar to ChatGPT.

A. Basic Model

To achieve a basic model of our chatbot, we propose the creation of a chatbot similar to *Akinator* [1]. To achieve this, we have found a tutorial on how to create such an application [3]. Using this tutorial as our basis, we can expand upon it and customize it to work with cars instead of characters. For our chatbot to have an accurate and comprehensive database of vehicles to search through, we propose the use of the CarAPI database [5], an automotive API created for developers that provides a free vehicle database.

This basic model of the chatbot will prompt the user with a series of simple questions, such as "What is your budget?" and "Would you prefer to drive gas, hybrid, or electric?". The chatbot will process the answers to these questions, search the database for cars that fit the given criteria, and provide the user with suggestions that meet the requirements they provided.

This model will implement the core features that we believe are integral to this project. It will provide a user-friendly text-based interface that will search a reliable and comprehensive database and provide users with customized suggestions. Using this model as our base, we will improve its user-friendliness by leveraging modern NLP to provide a human-like and conversational interface that can process more complex input.

B. GPT Model

To achieve a human-like and conversational model of our chatbot that can process complex input, we propose the use of GPTs to create an interface similar to *ChatGPT*. This model will build off the basic model and improve upon how the user interacts with the chatbot. Through the use of GPTs, we hope to create a version of our chatbot that can process more complex user input. This will allow us to create a more conversational approach to searching for a car. For example, the user may be able to start a conversation with the chatbot by listing multiple criteria they want fulfilled, and the chatbot will process all the information and provide follow-up questions.

To accomplish this, we have investigated the use of several generative AI APIs. While OpenAI's custom GPTs look to be the most promising and well-known solution available, the custom GPT service is a premium service [10] and as such is not a feasible option for us. Instead, we propose the use of a free alternative, Google's *Gemini* [7] and *Google AI Studio* [8]. Google offers free models of their generative AI APIs [9] which makes it an appealing and powerful option for us to leverage in the development of this project.

C. Natural Language Processing

The Chatbot uses NLP (Natural Language Processing) to communicate with customers. Conversations are converted from texts into meaningful information that can power automated search of Data Sets. Chatbot is objective and efficient because it uses NLP to filter out subjective dialogues.

Customer's queries are tokenized and dissembled into units of texts. They are lemmatized to the root form and changed to lowercase. Punctuations, special characters, and numbers are removed. The texts are transformed into structured data. Word embeddings capture the semantic relationship of texts. The meaning of texts is analyzed by dependent and semantic parsing, they identify the grammar and tone of the conversation. The Chatbot uses Natural Language Toolkit in Python programming language as a conversation medium.

IV. System Design (Tentative)

We propose a system that integrates an AI chatbot with a vehicle database. To accomplish this, we have investigated several tools and libraries that will assist in the development of this project. We will outline our proposed system design choices below.

A. Architecture

Our proposed system architecture consists of a chatbot portion and a database searching portion. The chatbot will be programmed as described in sections 3.A and 3.B. It will be able to process either simple input in response to questions, or complex input in the form of conversational phrases. After processing the user's input, the chatbot will use this information to search the database.

The database consists of a number of cars with their features and specifications. The chatbot will be able to search this database and narrow down the results using the answers provided by the user. These results will then be returned for the chatbot to display to the user.

B. Languages

The go-to language for AI development in recent years has been Python. Its ease of use combined with extensive libraries tailored for AI and machine learning make it an ideal language for AI development. The tutorials we have found also make use of Python, and *Gemini*'s API includes a Python library. Python also includes tools for interoperability with C/C++, which makes it useful to us in the case that we need to implement a high-performance section of code in a language we are familiar with. As such, Python will be our language of choice for the development of this project.

C. Software Tools

We have outlined the tools and libraries in the proposal of this project. Here we will lay them out and explain our reasoning for utilizing them.

1) Google Gemini

Google *Gemini* is a generative AI tool that offers a free tier of their API service with lower rate limits. It boasts the ability to build agents for contexts such as data exploration and content searching [11], which gives us the confidence that *Gemini* will provide us with the means to create the chatbot we envision. In addition to potentially providing a conversational interface for our chatbot, it has the potential to help us search through our database.

2) ChatGPT

ChatGPT, like *Gemini*, is a generative AI tool offering free use of their API services and its ability to answer complex answers and write code as needed. This tool will be primarily used for code editing, but fixing, and conversion of straight code into something that works within the website itself or an executable app.

D. Data Sets

For our reliable and comprehensive vehicle database, we will be employing the use of CarAPI's free vehicle database.

The free database is limited to the years 2015-2020, but this is more than enough for us to demonstrate the use and utility that our chatbot will provide. CarAPI provides a free comma-separated value (CSV) file as well as the ability to query and retrieve vehicles using the API in a JavaScript object notation (JSON) format. This acts as the library the AI will pull from when deciding what car is the right one for you.

E. Chat Bot Classifications & Techniques

There are many types of chat bots. A very common type is **Task-orientated.** The goal of these bots is to assist the user with completing a task. This could be anything from purchasing a plane ticket or navigating a website. Task oriented chat bots operate in restricted domains as well. **Non-Task oriented** chat bots operate in a more complex matter. Eventhough responses are still pre-defined, they also generate responses based on pattern recognition. It does this by Generative-Based (generates proper responses during interaction) and Retrieval-Based approaches, (Uses a repository to learn and make more informed responses). **Domain Specific** chat bots can either operate in open or closed domains. Open domains are used for non-specific conversations, and domain specific are used for specific conversations with a goal. By specific conversations it's based off the domain associated (healthcare, or car finder). There are also different interaction types, **Text-Based** and **Voice-Based.** Competitive models tend to include speech-to-text and require more complex programming models. As they need to be able to filter noise and accurately translate speech to text.

As briefly mentioned, chat bots use different models to generate responses, **Rule-Based** and **Self-learning** are the main distinctions between the two. Self-learning models can either be Generative-Based or Retrieval-Based and then these models implement a variety of techniques. **Parsing** and **Pattern Matching** are essential to a good chat bot. They also utilize **Artificial Intelligence Markup Language**, **Markov chain models**, and **Artificial Neural Networks**. Our chat bot needs to utilize complex techniques to Generate responses or retrieve from repositories.

V. IMPLEMENTATION

The first version of our chatbot as described in section III.A is implemented as a basic Akinator-style chatbot and is based off of the previously mentioned Akinator tutorial [3] and uses the CarAPI database [5]. It prompts the user with a series of inquiries, which it uses to search the database and narrow down the vehicle options that it presents to the user as a result. As shown in Figure 1, this version of the chatbot can successfully deliver personalized search results to the user while being fault-tolerant enough to prevent a lack of prior knowledge by the user from negatively affecting the search results.

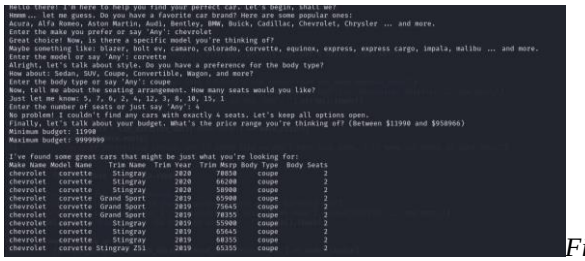

Fig. 1. Basic Version of Chatbot

VI. REFERENCES

[1] Elokence, "Everything you've always wanted to know about Akinator," Akinator.com, 2024. https://en.akinator.com/content/6/Everything-you-039-ve-always-wanted-to-know-about-Akinator (accessed Sep. 29, 2024).

[2] "Introducing ChatGPT," OpenAI, Nov. 30, 2022. https://openai.com/index/chatgpt/ (accessed Sep. 28, 2024).

[3] addykoder, "Youtube_Tutorial/Akinator.py at master · addykoder/Youtube_Tutorial," GitHub, 2021. https://github.com/addykoder/Youtube_Tutorial/blob/master/Akinator.py (accessed Sep. 29, 2024).

[4] Cars.com, "New Cars, Used Cars, Car Dealers, Prices & Reviews | Cars.com," Cars.com, 2019. https://www.cars.com/

[5] CarAPI. https://carapi.app/ (accessed Sep. 28, 2024).

[6] H. Alkaissi and S. McFarlane, "Artificial hallucinations in ChatGPT: Implications in scientific writing," Cureus, vol. 15, no. 2, Feb. 2023, doi: https://doi.org/10.7759/cureus.35179.

[7] "Build with the Gemini API," Google AI for Developers. https://ai.google.dev/

[8] S. Pichai and D. Hassabis, "Introducing Gemini: our largest and most capable AI model," Google, Dec. 06, 2023. https://blog.google/technology/ai/google-gemini-ai/

[9] "Gemini API Pricing," Google AI for Developers. https://ai.google.dev/pricing

[10] OpenAI, "ChatGPT ," ChatGPT, 2024. https://chatgpt.com/

[11] "Gemini API Docs and Reference | Google AI for Developers," Google for Developers. https://ai.google.dev/gemini-api/docs

[12] G. N. Punj and R. Staelin, "A Model of Consumer Information Search Behavior for New Automobiles," Journal of Consumer Research, vol. 9, no. 4, p. 366, Mar. 1983, doi: https://doi.org/10.1086/208931.

[13] Q. Ai, Y. Zhang, K. Bi, X. Chen, and W. Bruce Croft, "Learning a Hierarchical Embedding Model for Personalized Product Search," International ACM SIGIR Conference on Research and Development in Information Retrieval, Aug. 2017, doi: https://doi.org/10.1145/3077136.3080813.

[14] S. Reshmi and K. Balakrishnan, "Implementation of an inquisitive chatbot for database supported knowledge bases," Sādhanā, vol. 41, no. 10, pp. 1173–1178, Oct. 2016, doi: https://doi.org/10.1007/s12046-016-0544-1.

[15] Iván Cantador, Jesús Viejo-Tardío, M. E. Cortés-Cediel, and M. Pedro, "A Chatbot for Searching and Exploring Open Data: Implementation and Evaluation in E-Government," Biblos-e Archivo (Universidad Autónoma de Madrid), Jun. 2021, doi: https://doi.org/10.1145/3463677.3463681.

[16] B. Pitel et al., "63. An introduction to publicly available AI-assisted chatbot-style search engines for cancer variant curation," Cancer Genetics, vol. 286–287, pp. S20–S20, Aug. 2024, doi: https://doi.org/10.1016/j.cancergen.2024.08.065.

[17] IBM. (2024, October 25). *What is NLP (Natural Language Processing)?*. IBM. https://www.ibm.com/topics/natural-language-processing

[18] CarGPT, "CarGPT," CarGPT. [Online]. Available: <https://www.cargpt.ai/>. [Accessed: Oct. 27, 2024].

VII. Appendices

A. Appendix A - Task Assignments

1) David Del Moral

- Paper writing
  - Related works
- Similar implementation search

2) Ricky Godsey

- Paper writing
  - Proposed approaches

3) Ean Shi

- Paper writing
  - Propose approaches

4) Zachary Kao

- Paper writing
  - Related works
  - Proposed approaches
  - System design
- Scholarly research search

5) Matthew Tomme

- Webpage setup
- Implement prototype of basic version of chatbot

B. Appendix B - Schedule

- Deadline 1: EOD September 29, 2024
  - Project Proposal
  - 3 pages
  - Project Webpage
- Deadline 2: EOD October 27, 2024
  - Project Checkpoint 1
  - 5 pages
  - Progress reports from each member
- Deadline 3: EOD November 24, 2024
  - Project Checkpoint 2
  - 7 pages
  - Progress reports from each member
- Deadline 4: EOD December 8, 2024
  - Final Project Report
  - 10 pages
  - YouTube video
  - Project presentation slides
  - One-page project summary
  - Project source code/data/readme
  - Peer evaluations
- Deadline 5: Finals Week?
  - Project Presentation
  - 15-20 minute presentation including Q&A
  - 5 minute YouTube video demonstrating the application

C. Appendix C – Progress Report 1

1) David Del Moral

For checkpoint 1, I conducted online searches for related works, more specifically, products or concepts that currently exist that share most of the criteria for our project. I described what they do and how it compares to our ideal implementation.

2) Ricky Godsey

For checkpoint 1, I researched chat bots. This includes the different types of models and techniques they use to generate responses. Familiar terms like Neural Networks and self-learning are integral parts of successful chat bots. I expanded upon the proposed approaches section.

3) Ean Shi

For checkpoint 1, I have read about Natural Language Processing and described it in the paper.

4) Zachary Kao

For checkpoint 1, I performed searches for related works on Google Scholar and found relevant studies that pertained to consumer vehicle searches, personalized product searches, and the use of chatbots in searching for information. Using these sources, I expanded upon the related works section.

5) Matthew Tomme

For checkpoint 1, I worked on and created a proof of concept for the Akinator styled car finder. I also updated the website to include the updated documentation on said car finder and got the image included in the document.