ECE1786: Creative Applications of Natural Language Processing

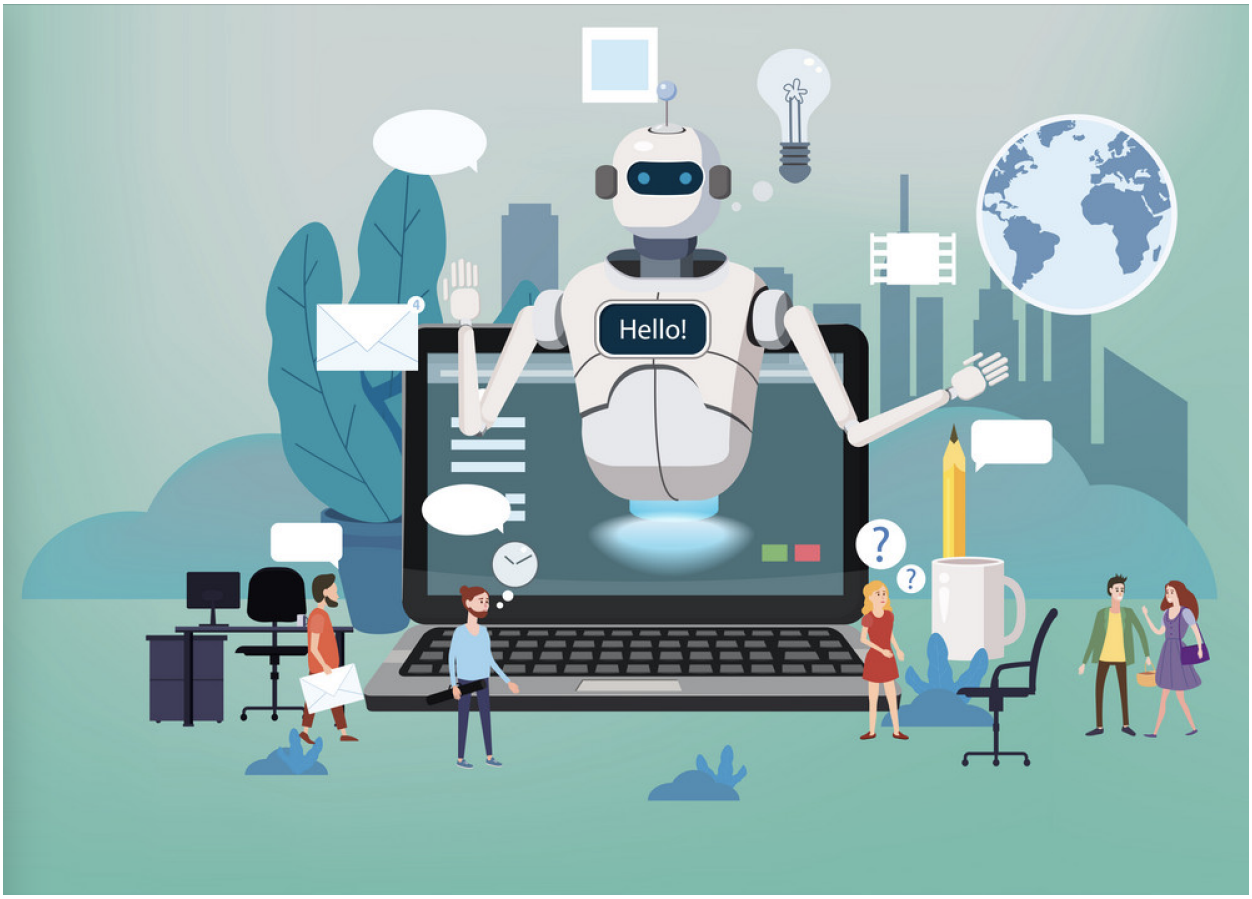
Word Count: 1975

Penalty: 0%

NexaHome

Final Report

December 2023



Chat Bot (Shutterstock)

Zifeng Zhu
Yichen Xiao

Permissions

Yichen Xiao

- permission to post video: yes
- permission to post final report: yes
- permission to post source code: no

Zifeng Zhu

- permission to post video: yes
- permission to post final report: yes
- permission to post source code: no

1 Introduction

Canada's history is deeply intertwined with immigration, drawing individuals globally who see it as an ideal place to settle. As of 2021, over 8.3 million people, roughly a quarter (23.0%) of its population, were either current or former permanent residents or landed immigrants. This figure marks the highest ratio since the country's Confederation, surpassing the 1921 peak of 22.3% and is the most significant among the G7 nations.

To bridge the everyday assistance gap for Canadian immigrants by connecting them with proficient household service providers that speak the same language, we aim to develop a customer service chatbot that can address user queries in real time. And we will prioritize the customer's language choice with our newly designed tasker matching criterias to help Canadian immigrants surmount the language barrier. NexaHome will be able to take clients' requirements about household service area, time availability, language and gender preferences, interpret these descriptions and intelligently match them with the most suitable taskers based on their expertise and alignment with the user's needs.

Our range of services encompasses cleaning, plumbing, moving help and furniture assembly. The chatbot will act as a hub of skilled individuals ready to assist with daily errands, making the transition to new environments stress-free and more manageable. The usage of large language models (LLMs) is essential for this project as we need to clearly understand the customers' descriptions and instructions about their daily errand challenges, especially important for communicating with Canadian immigrants whose first language is not English.

2 Illustration / Figure

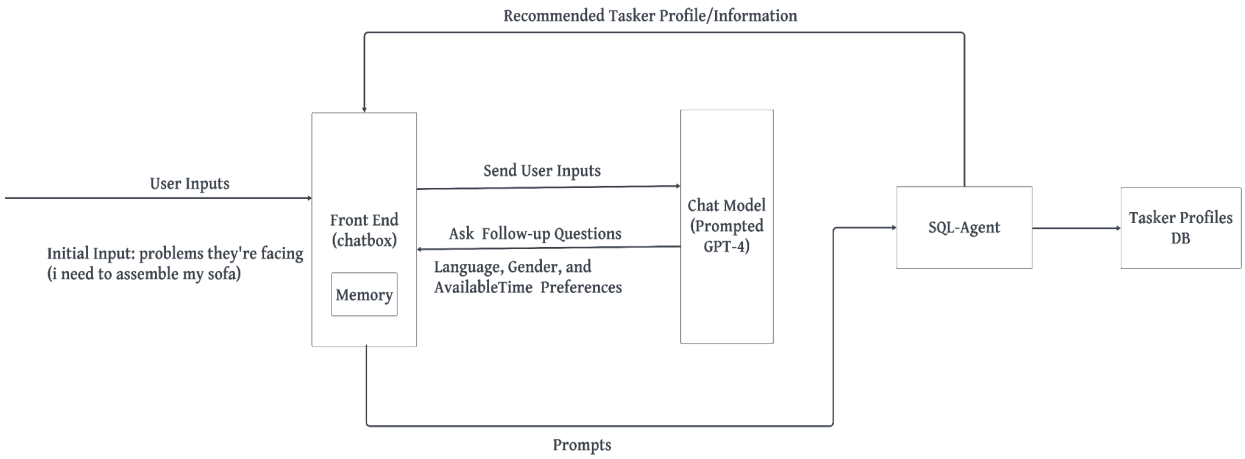


Figure 1. Main Model Architecture

3 Background & Related Work

3.1 An Overview of Chatbot Technology

The research paper summarizes the evolution of chatbots and discusses its applications in multiple fields like Marketing and Supporting System etc. The paper depicts the history of chatbots firstly and introduces technological concepts, classification, and architectures of various chatbots, and discusses the motivations and social impacts of chatbot design. The paper emphasizes the essential impact of NLP (Natural Language Processing) on the future chatbots, and the potential enhancement of chatbots on human-computer interaction (Adamopoulou & Moussiades, 2020).

3.2 Revolutionizing Business Communication: Exploring the Potential of GPT-4 in Corporate Settings

The research paper discusses the potential impact of GPT-4 on business communication. This paper highlights the ability of GPT4 to process vast amounts of data and respond similarly to humans, and to transform existing norms and practices in corporate settings to develop areas like customer service, and internal communication. On the other hand, it also discusses potential challenges and ethical concerns about the role of GPT4 in the business environment (George et al., 2023).

---

4 Data and Data Processing

4.1 User Input Data

For our project, we meticulously gathered 35 instances of typical household service requests from online sources. These samples were intentionally retained with their original grammatical inaccuracies. This approach was chosen to reflect the realistic language proficiency of many Canadian immigrants, who may not be native English speakers. Additionally, we created and integrated 15 examples using ChatGPT, ensuring these were grammatically flawless. This was done to represent a spectrum of English proficiency levels within our target demographic. These 50 examples were then categorized and aligned with the four service offerings of our chatbot: cleaning, plumbing, moving assistance, and furniture assembly. A representative sample from this collection is: “My sink is clogged”. We further enriched this dataset by incorporating various parameters such as language preference (e.g., English, Chinese, French), gender preference (e.g., Male, Female), and

availability (e.g., this Wednesday at 4pm). This comprehensive dataset forms the foundation of our user input data for the project.

4.2 Tasker Profile Data

We employed ChatGPT to generate 20 illustrative tasker profiles, adhering to the profile schema delineated in Figure 2. This schema includes several key attributes: name, nationality, gender, languages spoken by the tasker, available time slots (encompassing both weekdays and specific time ranges), tasker rating, and areas of expertise. Notably, the language skills, AvailableTimeSlot, and Expertise fields are formatted as JSON arrays. Subsequent to their generation, these profiles underwent a thorough cleaning process. This step ensured the elimination of grammatical errors and the refinement of the content to be succinct, relevant, and strictly pertinent to our service categories and other specified criteria. Post-cleaning, these profiles were meticulously entered into our database, ensuring their readiness for future application. Figure 3 provides a visual representation of the tasker profile table in our database.

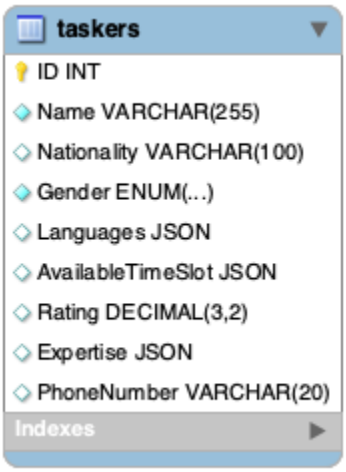


Figure 2. Tasker Profile Schema

ID	Name	Nationality	Gender	Languages	AvailableTimeSlot	Rating	Expertise	PhoneNumber
1	Alice Smith	Canadian	Female	["English", "French"]	[{"time": "08:00-10:00", "weekday": "Monday"}, {"time": "14:00-16:00", "weekday": "Tuesday"}]	4.80	["Cleaning", "Plumbing"]	183-883-2916
2	Bob Johnson	British	Male	["English"]	[{"time": "09:00-11:00", "weekday": "Wednesday"}, {"time": "13:00-15:00", "weekday": "Friday"}]	4.30	["Moving Help", "Furniture Assembly"]	288-285-1301
3	Carlos Rivera	Mexican	Male	["Spanish", "English"]	[{"time": "10:00-12:00", "weekday": "Thursday"}, {"time": "16:00-18:00", "weekday": "Saturday"}]	4.50	["Cleaning"]	784-215-1455
4	Diana Chen	Chinese	Female	["Mandarin", "English"]	[{"time": "11:00-13:00", "weekday": "Tuesday"}, {"time": "15:00-17:00", "weekday": "Thursday"}]	4.70	["Plumbing", "Furniture Assembly"]	195-112-7526
5	Edward Müller	German	Male	["German", "English"]	[{"time": "09:00-11:00", "weekday": "Monday"}, {"time": "14:00-16:00", "weekday": "Wednesday"}]	4.60	["Moving Help"]	421-811-9851
6	Fiona Lee	American	Female	["English"]	[{"time": "10:00-12:00", "weekday": "Monday"}]	4.20	["Cleaning"]	253-367-1606
7	George Kim	Korean	Male	["Korean", "English"]	[{"time": "09:00-11:00", "weekday": "Tuesday"}, {"time": "14:00-16:00", "weekday": "Thursday"}]	4.10	["Plumbing"]	743-327-6774
8	Hannah Martinez	Mexican	Female	["Spanish", "English"]	[{"time": "13:00-15:00", "weekday": "Wednesday"}]	4.40	["Furniture Assembly"]	734-851-7353
9	Ian Nguyen	Vietnamese	Male	["Vietnamese", "English"]	[{"time": "10:00-12:00", "weekday": "Friday"}, {"time": "15:00-17:00", "weekday": "Saturday"}]	4.00	["Moving Help", "Cleaning"]	772-319-9929
10	Julia Ivanova	Russian	Female	["Russian", "English"]	[{"time": "11:00-13:00", "weekday": "Monday"}]	3.90	["Plumbing", "Furniture Assembly"]	899-305-6884
11	Kevin Zhao	Chinese	Male	["Mandarin", "English"]	[{"time": "08:00-10:00", "weekday": "Tuesday"}, {"time": "15:00-17:00", "weekday": "Thursday"}]	4.30	["Cleaning", "Furniture Assembly"]	815-152-5087
12	Linda Ochoa	Colombian	Female	["Spanish", "English"]	[{"time": "09:00-11:00", "weekday": "Monday"}, {"time": "14:00-16:00", "weekday": "Wednesday"}]	4.60	["Plumbing"]	747-454-6902
13	Mohamed Ali	Egyptian	Male	["Arabic", "English"]	[{"time": "10:00-12:00", "weekday": "Sunday"}]	4.20	["Moving Help"]	363-978-8787
14	Nina Patel	Indian	Female	["Hindi", "English"]	[{"time": "09:00-11:00", "weekday": "Friday"}, {"time": "16:00-18:00", "weekday": "Saturday"}]	4.50	["Cleaning", "Furniture Assembly"]	459-903-2937
15	Omar Khaled	Jordanian	Male	["Arabic", "English"]	[{"time": "11:00-13:00", "weekday": "Tuesday"}]	4.10	["Plumbing", "Moving Help"]	555-413-2343
16	Patricia Santos	Brazilian	Female	["Portuguese", "English"]	[{"time": "10:00-12:00", "weekday": "Wednesday"}, {"time": "14:00-16:00", "weekday": "Friday"}]	4.70	["Cleaning"]	348-296-1275
17	Quentin Dubois	French	Male	["French", "English"]	[{"time": "13:00-15:00", "weekday": "Monday"}]	4.00	["Furniture Assembly"]	787-109-4700
18	Rachel Green	Australian	Female	["English"]	[{"time": "09:00-11:00", "weekday": "Tuesday"}, {"time": "15:00-17:00", "weekday": "Thursday"}]	4.40	["Moving Help", "Plumbing"]	598-735-2743
19	Samuel Lee	Korean	Male	["Korean", "English"]	[{"time": "08:00-10:00", "weekday": "Wednesday"}]	3.80	["Cleaning", "Furniture Assembly"]	733-959-5816
20	Tina Mendoza	Argentinian	Female	["Spanish", "English"]	[{"time": "10:00-12:00", "weekday": "Friday"}, {"time": "14:00-16:00", "weekday": "Sunday"}]	4.30	["Plumbing", "Moving Help"]	370-416-8791

Figure 3. Tasker Profile Table

5 Architecture and Software

The diagram demonstrated in section 2 is a flowchart for a chatbot system. Here's a breakdown of the flowchart:

- User Inputs: This is where the user starts interacting with the system by providing input data, like a household problem they're facing as their initial query (e.g., I need to assemble a sofa).
- Front End (chatbox): The chat box represents the user interface where the user interacts with the system. It has a memory component used to retain context or previous interactions. Gradio with chatbot mode is used for the system's UI so it is convenient to establish the chatbot environment. Furthermore, the memory in the front end is from the Langchain library.
- Chat Model (Prompted GPT-4): This box represents the backend chat model, specified as GPT-4 that has been prompted. An important portion of the prompts used in this model is “You are a polite customer service chatbot having a conversation with a human. Once a user mentions something about household

inquiries such as cleaning, plumbing, moving help or furniture assembly, answer with something that involves an empathic interaction with the user's situation to enhance the user's comfort and engagement." Then, the second piece of the prompts in this model will instruct the chatbot to ask follow-up questions based on the specific household service need and the information that the user has already provided.

- SQL-Agent: This component interacts with the database to retrieve or store information. It's the bridge between the chat model and the database where tasker profiles are stored. The main advantage of SQL agent is that it can answer questions based on the databases' schema as well as on the databases' content effectively. The SQL-Agent is a component of Langchain. The agent is also prompted, and an important portion of the prompt used in this instance is "Find me all tasker's profiles that can help me if {inquiry} and who speaks {language} and whose gender is {gender} and whose available time slot is {available_time}." The values inside {} are retrieved from the chat history between users and the chatbot.
- Tasker Profiles DB: A database that contains profiles of workers who can perform tasks such as furniture assembly. MySQL DB is utilized as the SQL database which stores worker's profiles.

The flow shows that once the user inputs their issue into the front end, the chat model will ask follow-up questions to refine the user's request, taking into account preferences such as language, gender of the tasker, and available time. The memory stores all conversations between the chat model and users. Once enough information is gathered, the front end will create a prompt and send it to the SQL agent, and the SQL agent will retrieve suitable tasker profiles from the Tasker Profiles DB to recommend to the user, directly sending profiles to the front end. The "Recommended Tasker Profile/Information" is an output

presented to the user after processing their inputs and preferences, offering them a selection of taskers that fit their criteria.

---

6 Baseline Model

The diagram outlines the baseline model in which user inputs, such as a request to assemble a sofa, are processed by the GPT-4 AI model. This model is prompted with a dataset from the tasker profile database (DB) stored in a CSV file format, which contains the worker's information. After processing the input with the help of the dataset, GPT-4 generates an output that includes a recommended tasker profile, which means the most suitable service provider from the database to complete the user's requested task. The process is a direct input-output flow without any intermediary steps. In other words, users should make their requirements as specific as possible so that gpt-4 can select the correct tasker profile. The baseline model is implemented in the ChatGPT; firstly, the prompt with the dataset is imported to allow gpt-4 to read and learn the dataset. Then, the user can type their requirements (input data) into GPT-4 to receive the recommended worker's profile.

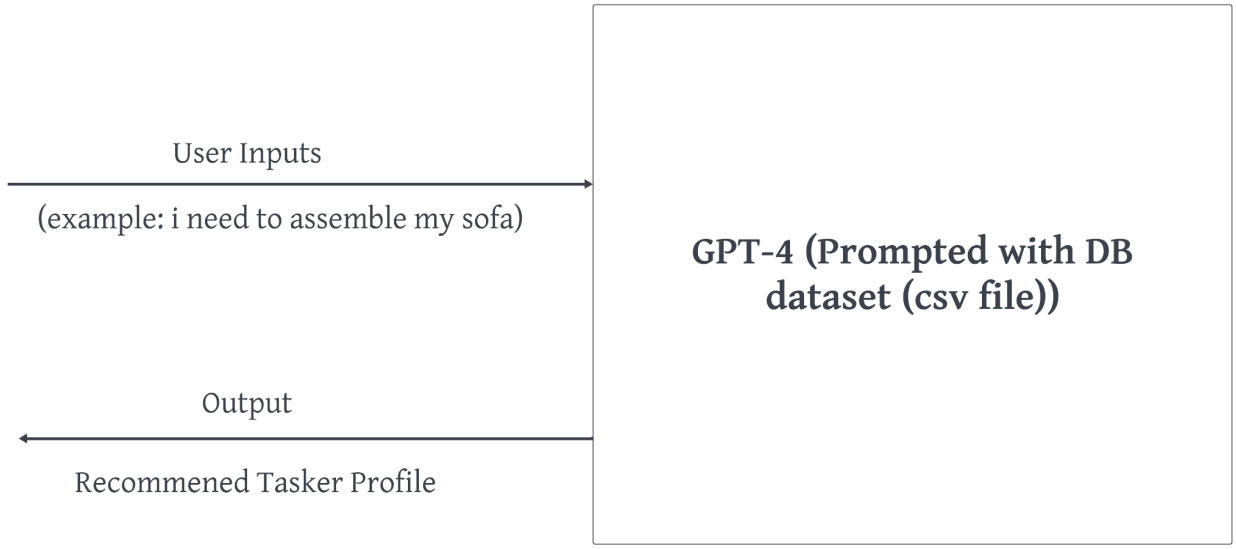


Figure 4. Baseline Model Architecture

7 Quantitative Results

7.1 Comparison Method

- Matching Accuracy

  Evaluates the chatbot's precision in aligning service providers with user-specified needs, crucial for service effectiveness. It is measured by whether the output matches the following four categories of information: household service area, language preference, gender preference and available time slots.

- User Satisfaction

  Measures user contentment with the chatbot's interaction quality, match quality, and request resolution, indicating the chatbot's success and user trust. We invited a

group of 20 people to use our chatbot, and then asked them to fill out a survey to score the service that the chatbot provided. Detailed instructions and evaluation metrics of good service are defined in the background section of the survey.

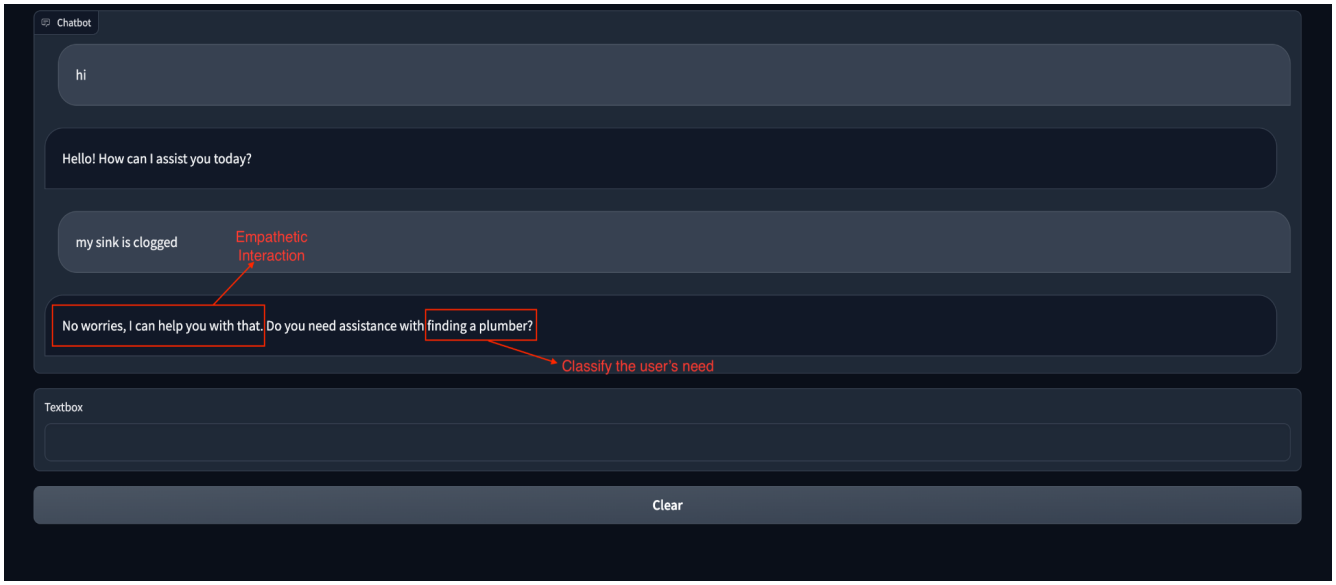
- Tone and Empathy

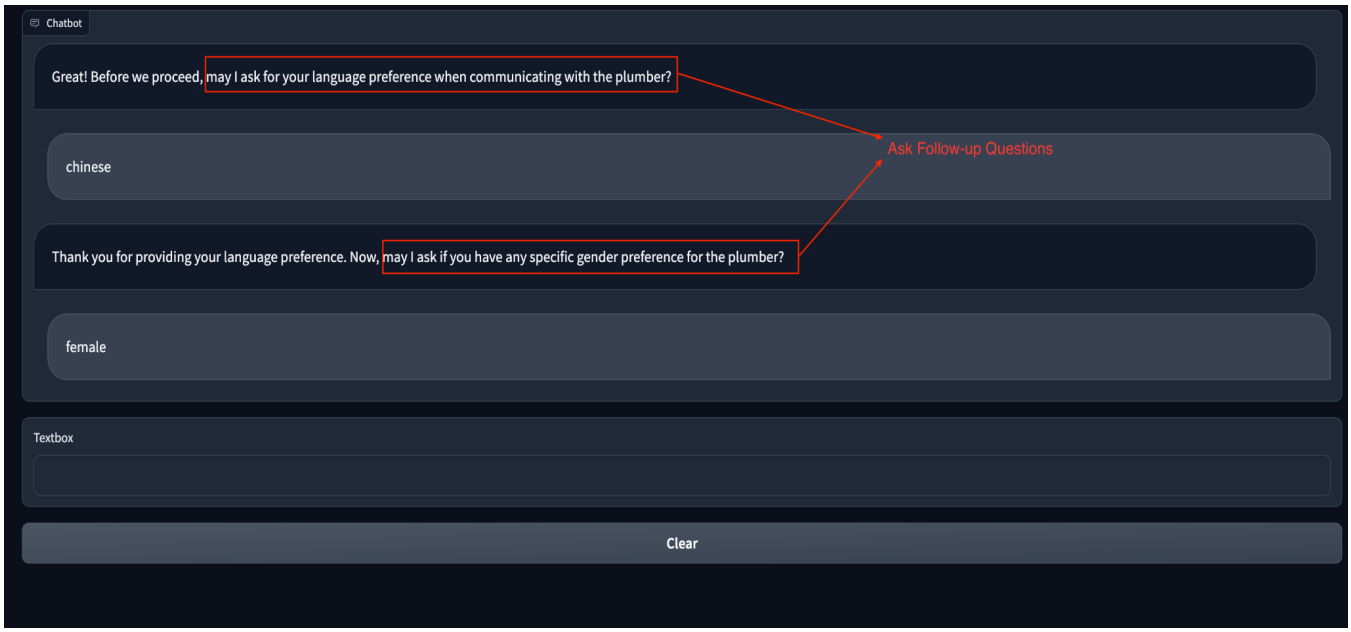
  Assesses the chatbot's contextual, empathetic interaction, enhancing user comfort and engagement, especially important for assisting immigrants in new environments.

7.2 Numeric Results

- Baseline Model
  1. Matching Accuracy: 82%
  2. User Satisfaction: 52%
  3. Tone and Empathy: 20%
- Main Model
  1. Matching Accuracy: 94%
  2. User Satisfaction: 86%
  3. Tone and Empathy: 70%

8 Qualitative Results





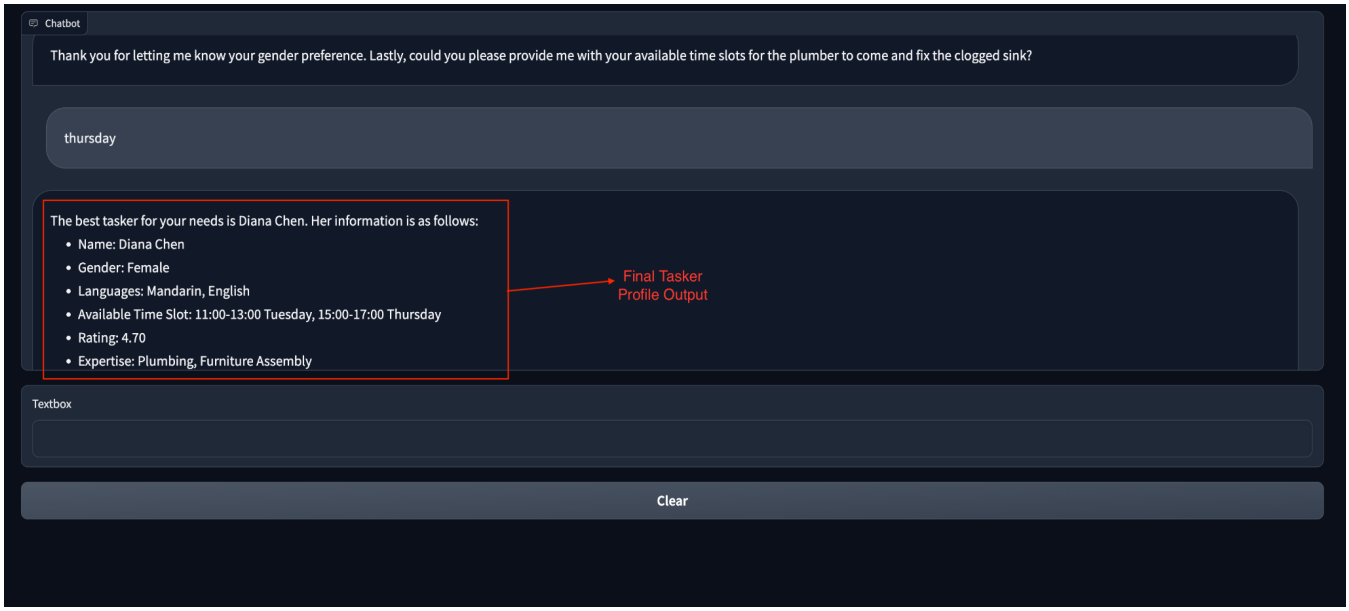


Figure 5. Sample inputs and outputs

The exhibited results clearly demonstrate the chatbot's proficiency in accurately identifying the highest-rated tasker who fulfills all the user-specified preferences, achieving a matching precision of 100% in the illustrated instance. Additionally, the chatbot's capability to respond with phrases like “No worries” in reaction to the user's concerns is indicative of its ability to engage in empathetic interactions. This feature enhances the user experience by providing a more human-like and understanding response. Furthermore, the chatbot is equipped to pose contextually relevant follow-up questions, which are tailored based on the specific details of the user's query and the ongoing conversational dynamics. This functionality underscores the chatbot's advanced interaction skills, allowing it to address user needs more effectively and in a nuanced manner.

9 Discussion and Learnings

In our project, we observed that the main model significantly outperformed the baseline across all key evaluation metrics, underscoring its efficiency and effectiveness. This superior performance is a testament to the model's advanced capabilities in accurately processing user inputs and tasker profiles, ensuring precise matches and fostering empathetic user interactions. The experience gained in prompt engineering was particularly enlightening; our methodical approach to this aspect enhanced the chatbot's responses, illustrating how subtle nuances in prompt design can profoundly influence outcomes.

Looking forward, we recognize the value of integrating real-world data in future projects. Such data offers in-depth industry-specific insights, enhancing the applicability and relevance of our model in real-world scenarios. Additionally, the development of a web scraper for regular database updates emerges as a crucial consideration. This tool would ensure the chatbot's alignment with the latest industry trends and requirements, maintaining its accuracy and relevance over time.

10 Individual Contributions

Zifeng Zhu worked on designing and refining the main model prompts, and he wrote the logic to connect all the prompts to corporate together with the chat history and the database. He also developed the front end user interface and the memory for chat history.

Yichen Xiao designed the database schemas and synthetised data with ChatGPT to create the tasker profile database used in this project. He also worked on developing the SQL agent that takes several filter values to extract the best tasker profile.

The group worked together to collect the user input data used for testing and design the user satisfaction survey.

References

[1] Statistique Canada, 'The Daily — Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as Canadians,' Statistique Canada, 26-Oct-2022. [Online]. Available: https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm. [Accessed: 20-Nov-2023].

[2] E. Adamopoulou and L. Moussiades, 'An Overview of Chatbot Technology,' in Artificial Intelligence Applications and Innovations, I. Maglogiannis, L. Iliadis, and E. Pimenidis, Eds., vol. 584, AIAI 2020. 2020. doi: 10.1007/978-3-030-49186-4_31.

[3] A. S. George, A. S. H. George, T. Baskar, and A. S. G. Martin, 'Revolutionizing Business Communication: Exploring the Potential of GPT-4 in Corporate Settings,' Partners Universal International Research Journal, vol. 2, no. 1, pp. 149-157, 2023. doi: 10.5281/zenodo.7775900.