Conference Paper Title*

*Note: Sub-titles are not captured for https://ieeexplore.ieee.org and should not be used

1st Danyal Mirza
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Linnéa Soto Carlsson
Department of Computer Science
Uppsala University (UU)
Uppsala, Sweden
Linnea.sotocarlsson.7666@student.uu.se

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—Linnea + Danyal**

***Index Terms*—component, formatting, style, styling, insert.**

I. Introduction

Linnea + Danyal

The objective of the project was to design and develop "Bart", a bartender capable of reading the client's emotional state and providing drink recommendations aligning with the detected mood while engaging in a pleasant conversation. In addition, the system takes into account the individual drink preferences of the clients to ensure a satisfying experience. These specific objectives were chosen in the beginning of the project, as the group members believed they were essential for creating a meaningful project connected to the course work. The work for this project was divided into two primary subsystems: the Perception Subsystem and the Interaction Subsystem.

- **Linnéa** and **Danyal** were responsible for developing the perception subsystem. They built a video server capable of detecting faces and recognizing emotions, with the help of a Random Forest (RF) machine learning model to classify mood based on facial expressions.
- **Adrià** and **Walter** focused on the interaction subsystem. Their work involved setting up the Furhat robot, integrating a Large Language Model (LLM) into Furhat, and designing the system to translate detected emotions and client preferences into prompts for generating personalized drink recommendations.

Teamwork within the group was collaborative and efficient, where meaningful discussions helped solving problems encountered throughout the project. The writing of the report was divided equally among the members, and each person primarily contributed to their specific area of work.

II. Methodology

Linnea + Danyal

This section describes the methodology and system architecture used in the project.

A. Overall system design

Linnea + Danyal

The scenario chosen for this project is as said in previous sections a bartender. The primary task of Bart, our bartender, is to read the customer's emotional state and offer drink recommendations based on their mood. Additionally, the system personalizes the recommendations by taking into account the client's individual drink preferences from engaging in a pleasant conversation with the user. This ensures a more enjoyable experience.

HIGH-LEVEL GOALS:

1) **Emotion Recognition:** Detect the client's emotional state accurately based on facial expressions.
2) **Personalized Drink Recommendations:** Suggest a drink that fits both the emotional state and the individual preferences of the client.
3) **Engaging Conversation:** Maintain a pleasant and natural conversation with the client while offering drink recommendations.
4) **Real-Time Interaction:** Provide real-time responses to emotions and preferences without significant delays, ensuring smooth and interactive engagement.

What specific choices did you make to achieve the high-level goals?

B. User perception sub-system

Linnea + Danyal

C. Interaction sub-system

This section discusses the second subsystem of Bart the bartender. It goes over the high-level design and its implementation. Afterwards it evaluates the obtained results.

1) Design: The interaction sub-system in this project consists of the virtual furhat, Gemini and a connection to the video server. Furhat is used as the interface between the system and the user and Gemini is used to generate a response to the input of the user. The response is based on what the user says and what emotion is detected in the face of the user at the beginning of the conversation. This response is then sent to Furhat so that it can reply to the user. This process is repeated until the right drink is found for the user.

2) Implementation: The typical flow of this subsystem is as follows: A connection is set up between this subsystem and the video server that was described in the previous section. Once this connection is set up, Bart greets the user and asks what kind of cocktail the user wants. From this moment, the subsystem gathers the detected emotions of the user, until the user is done with speaking. From these emotions, we determine the dominant emotion of the user. When the user finishes their reply to Bart, we combine the dominant emotion with the reply of the user and send it to Gemini. Gemini is a LLM that is instructed to act like a bartender that recommends cocktails that fit the mood of the customer. We decided to use Gemini because it is free for people with a google account up to a rate limit. Our application does not use up a lot of tokens, so we should not reach this rate limit. Once Gemini has generated a response, it is forwarded to Bart. Then, Bart is instructed to say the response. After Bart is done with talking, it listens again for input of the user. The input of the user is again combined with the dominant emotion and sent to Gemini. This cycle continues until the user decides on a cocktail. After this the whole cycle repeats from the beginning, assuming there is a new customer.

3) Result: In the project proposal, we described several objectives that we expect our bartender to fullfill. We can test our product by confirming it adheres to the set objectives.

One of the objectives is that our product gives recommendations to the user, based on the user's emotional state. We add the mood of the user to the prompt that we send to Gemini. Because of this, Gemini will take both the emotion of the user as well as the user's speech into account when generating a response.

Another objective is that Bart should take the user's preferences into account. We achieve this by letting Gemini generate questions that explores the needs and preferences of the user. Gemini incorporates current and previous answers of this user into the recommendation.

The last objective is that the conversation with the system should be pleasant. This is a more subjective goal, but we try to make the conversation as pleasant as possible by continuously adjusting the prompts given to Gemini. By doing this, we hope to make the conversation as pleasant as possible.

III. General Discussion

Adria

A. Overall Pipeline

Adria

B. Challenges

Adria

C. Use of ChatGPT and other tools

ChatGPT has been used to assist in resolving error messages in the code and to discuss different ways to optimize machine learning techniques. The information provided by ChatGPT has been cross-verified with other sources, including GeeksforGeeks and the official Python documentation for specific packages. Using ChatGPT helped streamline our workflow by providing quick solutions to technical challenges and offering alternative approaches to improve model performance. Additionally, it saved significant time during debugging and allowed us to focus on implementing and testing new ideas, ultimately enhancing our understanding of machine learning concepts.

AI has also been used as a proof reader of our project report and gave us valuable help with formulation of sentences and handling of grammatical errors. This significantly improved the flow and readability of our report.

D. Ethical Issues

Adria

Conclusion

Adria

References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.