CS310 Project Specification: Chameleon - An Adaptive ChatBot

Brandon Hong u2101112

12th October 2023

1 Problem Statement and Motivation

Conversations are essential for personal well-being and to form meaningful connections with others. While there are a myriad of benefits into having a meaningful conversation, many individuals find themselves without the opportunity of having conversational partners for effective communication or having a safe space to convey private knowledge.[1] Lacking communication skills or conversational partners often leads to difficulties in maintaining relationships and having effective communication, contributing to feelings of loneliness and depression.[2][3] Given that negative experiences, social anxiety, and insecurity are often cited as significant contributors, being able to have conversations can be instrumental in supporting individuals with their mental health and confidence to mitigate the impacts of these contributing factors. [4]

With the advancement of smart conversational agents in recent years, resulting in a surge in demand for this cutting-edge technology for a diverse range of users, including those seeking for a conversational partner.[5] While current chatbots excel at responding user queries at high speed and accuracy, they often fall short of acting as conversational partners due to their configuration, which is to serve as helpful assistants rather than attentive listeners.

In an attempt to address the problem, Chameleon intends to utilise the benefits of self-talk and self-compassion to provide a comfortable environment for users to engage in a conversation and disclose their thoughts in.[6] As chatbots become increasingly prevalent in our society, this project aims to leverage this technology by developing an adaptive chatbot capable of emulating the user's conversational demeanour and mirroring their personality traits, allowing a conversational experience akin to talking to oneself.[7]

2 Methodology

2.1 Problem Approach

The main approach of the project is to profile users using the Myers–Briggs Type Indicator (MBTI) and develop machine learning models to analyse the user's conversational style, including tone, vocabulary and response time.[8][9] By profiling the user, we can gain a deeper insight on the general tendency or conversational style that the user is likely to adopt. The Myers-Briggs Type Indicator classifies personality types into 16 distinct categories based on four dichotomies:

- **Extraversion (E)** vs. **Introversion (I)**: This dimension measures how individuals gain energy.
- **Sensing (S)** vs. **Intuition (N)**: This dimension reflects how individuals gather information.
- **Thinking (T)** vs. **Feeling (F)**: This dimension relates to decision-making.
- **Judging (J)** vs. **Perceiving (P)**: This dimension refers to how individuals approach structure and planning.

This information will be utilized to preconfigure the chatbot's initial starting state. However, it is important to keep in mind that the MBTI is one of many simplified personality assessment theory and may not capture the full complexity of an individual's personality. To mitigate this inaccuracy, we will use natural language understanding (NLU) techniques to extract data from user responses using intent classification and entity recognition, and machine learning models to regenerate responses for the chatbot.[10] It is unclear at this stage which machine learning model will be adopted or if a custom model is required until testing is performed to evaluate the accuracy and suitability of a model. However, these are the models that are currently considered for response generation[11][12]:

- **Reinforcement Learning** Define rewards and penalties for different bot responses from user feedback to optimise the bot's behaviour overtime by maximising rewards and minimising penalties.

- **Neural Networks** Recurrent neural networks are a a deep-learning-based solution for the construction of a chatbot. [13][14] Additional research is required to fully understand this approach.
- **Sequence-to-Sequence Models** Enables the chatbot to map input sequences (user messages) to output sequences (bot responses), allowing it to learn from conversational data.
- **Transformer-Based Models** Using Large Language Models (LLM) such as GPT-3 or BERT, By fine-tuning these models on conversational data, the chatbot can adapt its responses to user inputs.
- **Rule-Based** The simplest approach that lacks flexibility. Rule-based models are used to define explicit rules and conditions for generating responses. It will be used to enforce specific behaviours in the chatbot, ensuring adherence to ethical guidelines which is further discussed under section 6. (Ethical Issues and Considerations)

This is a list of the performance metrics that will be used to determine if a model is well suited for the project. Base on the objective of the chatbot, the first three metrics will be prioritised.

- **Personality Reflection** Determined by how accurately the chatbot mimics the personality. This can be assesed with user feedback and appropriateness of responses based on the desired personality.
- **Adaptation Speed** Shows how quickly the chatbot adapts to the user or changes in user behaviour and preferences during a conversation.
- **Quality of output text** Assess the quality of responses. Responses should have good readability and diversity while adhering to desired personality type.
- **Hardware Requirements** Evaluate the resource utilization of the model, such as memory and processing power, to ensure that it can run smoothly. (Can be a limiting factor for Transformer-Based Models)
- **Robustness** Examine how well the chatbot handles unexpected or off-topic conversations. A robust chatbot should gracefully handle situations outside its scope with fallback actions.

2.2 Project Management

At this stage of the project, any initial research performed was to gauge the feasibility of the project and conforming the existence of frameworks to engine the project's objectives. Moving forward, continuous research on technical aspects of the project will be apparent, and the study into personality theories to facilitate an accurate classification of the behaviour associated with each personality type. With this in mind, the project is expected to proceed in a repeating cycle consisting of three components - **Research**, **Development**, **Testing**.

An agile approach to the project will be adopted and is most suitable as development should be performed incrementally and tested at each checkpoint. Due to the nature of the project, identifying any errors in later developments will only prove to be increasingly harder. **Monday.com** will be used to track and monitor the progress of the project. Furthermore, biweekly meetings have been set with the project supervisor to review progress, receive feedback and ensuring the project's progress stays on track.

The **RASA** python framework will be used to develop the chatbot. Rasa is open-source, allowing space for customisation to meet the project's needs and supports machine learning integration with **TensorFlow** for tasks like intent classification and response generation. The details and utilisation of the framework's components will be elaborated upon in the objectives section. Project code will be uploaded to a **GitHub** repository for backup and managed with **Git** as the version control system with commits performed incrementally every few days to avoid losing progress.

2.3 Existing Works and Literature

Replika is a chatbot developed by Luka, Inc primarily capable of recognising and responding to user emotions with sentiment analysis and emotional understanding. It's main **use case** is to engage users in natural language conversations and provide them with emotional support, companionship, and mental health assistance.[15]

Character.AI is based on neural language models. A supercomputer reads huge amounts of text and learns to hallucinate what words might come next in any given situation. In Character.AI, users can customise a chatbot by providing conversation messages as data then rate the responses for continuous learning. It's primary **use case** provides users a fun platform to customise any chatbot they desire.[16]

While both Replika and Character.AI are trained on extensive conversational datasets, each focuses on distinct objectives. Chameleon aims to fuse the strengths of both approaches while introducing a novel concept, by using MBTI

to customise the chatbot. This provides users with self-companionship and a personalised conversational experience which goes beyond conventional chatbots.

3 Objectives

The below objectives define the different technical aspects that make up the chatbot as well as potential improvements which is labelled with *Could* or *Should*.

1. **Chatbot Development (Must)**
   - Develop a chatbot capable of understanding and conversing in English.
   - Implement robust fallback responses for cases where the chatbot cannot generate a meaningful response
2. **User Profiling (Must)**
   - Apply personality type theories to classify users into distinct personality types.
   - Customize the chatbot's tone, conversational style, and choice of language based on the user's personality type.
3. **Natural Language Processing (Must)**
   - Utilize Rasa NLU for intent classification and entity extraction to enhance language understanding. Intent classification enables the system to accurately classify user messages into specific categories, such as determining whether a user is expressing a mood ('happy') or seeking assistance ('asking for help'). On the other hand, entity extraction allows the system to identify and classify important entities within the text, such as names of individuals and locations. This capability enhances the context of conversations and enables our chatbot to provide more personalized and relevant responses.
4. **Machine Learning (Should)**
   This is only labelled as should as we can naively hard-code the bot, however doing so would significantly undermine the project's overall quality.
   - Integrate personality traits into the chatbot's behaviour using machine learning models and appropriate training data.
   - Fine-tune machine learning models to strike the right balance between performance and computational efficiency.
   - Enable the chatbot to adapt to evolving conversations and make context-aware decisions, while exploring different models to optimise accuracy.
5. **User Interface to Converse with the Bot (Should)**
   This is a should objective as we can still converse with the chatbot in terminal using the command **rasa shell**, but this will only allow user testing with a personal laptop.
   - Design a user-friendly interface for users to interact with the chatbot, allowing for seamless communication, input processing, and response generation. (the simplest approach can be done with a index.html file and chatroom.js package)
   - Having a user interface that multiple users can easily access at a time with allow more rigorous and thorough testing
6. **Contextual Management (Should)**
   - While RASA's tracker store and slots provide a valuable means to store contextual data for improved communication, it's important to note that this storage is temporary and limited to the duration of the chatbot's runtime. Consequently, any contextual data will be lost upon termination, and users will need to start anew in the next session.
   - Explore the possibility of integrating a database to store chat history and maintain context even after program termination, with a preference for PostgreSQL.

3.1 Testing and Feedback

Upon starting the development server using the command **rasa shell**, it creates an accessible API endpoint that can be conveniently tested using Postman, an HTTP request testing software. By including a message in the request body, this endpoint returns a list of intents along with their rankings. These intents represent the next response that the chatbot is likely to generate.

Automated testing can be executed using the **rasa test** command, it will execute the test codes written to check specific user paths or stories. The output provides insights into whether the chatbot successfully follows the intended path or reports errors if deviations occur. Additionally, the rasa interactive command offers a step-by-step demonstration of the conversation flow, facilitating the debugging of the chatbot. However, it's important to note that this process can be time-consuming and is typically reserved for cases where conventional debugging methods have proven ineffective. [17]

The chatbot will be tested on users, ideally representing an equal spread of each 16 different personalities. A survey will be conducted to collect user feedback and assess their satisfaction, providing an overall evaluation of the chatbot's performance. This survey will serve as a valuable measure to gauge how well users perceive the chatbot's conversational style and its ability to adapt to their preferences. Additionally, the chatbot can be compared to existing works; however, finding a suitable metric for fair comparisons may be challenging.

3.2 Further Extensions

1. Provide clear and user-friendly documentation that explains how users can interact with the chatbot and outlines its capabilities.
2. Emulate realistic conversation flows by allowing users to send separate messages rather than a single large chunk of text. The bot should respond at a reasonable time, mimicking the pace of a natural conversation.
3. Instead of immediately responding to prompts, the bot can ask user questions to prompt users for more information or encourage them to continue the conversation.

4 Resources and Risks

The below table represents a risk analysis of the project and mitigating actions:

Risk	**Impact**	**Contingency**
Personal Device Breakdown	Productivity will be affected leading to a delayed time schedule	The use of a repository (GitHub) to avoid losing progress. The program can be ran in DCS machines after installing the dependencies.
Machine learning models fail to provide good responses	A core objective that can decrement the quality of the project	Allocate majority of time into researching and fine-tuning the models. If a model does not provide satisfactory results then switch to a different model.
User Feedback Quality	User feedback may be inconsistent or insufficient for training the chatbot effectively	Implement feedback mechanisms such as star ratings and follow-up prompts to encourage users to provide detailed and constructive feedback.[18]
User Adaptation Challenges	Chatbot may struggle to accurately adapt to users' personalities and conversational styles	Conduct regular user testing and iterate based on feedback

Figure 1: Risk analysis table

Currently, no additional resources are required for Chameleon other than the frameworks aforementioned. Only potential resource would be the discovery of a suitable open-source data set that could be utilised as training data for the model. A quick overview of the resources and it's purpose:

- RASA - chatbot development
- TensorFlow/Keras Core - machine learning
- Visual Studio Code - integrated development environment (IDE)
- Git - version control system

- GitHub - backup repository
- Training data set in the form of collected or open-sourced data

5 Legal, Social, Ethical Issues & Considerations

To improve the bot's response, user testing and feedback will be heavily required in this project to verify the performance and accuracy of the chatbot. Furthermore, user data will be collected while remaining anonymous to serve as training data for building machine learning models. It is also important to note that the chatbot should not disrupt user with unpleasant emotions and is within the vision of the project to provide an overall positive experience.

Due to scope and time constraint of the project, it is assumed that data storage will only be stored in a local/personal computer. Data will not be uploaded or used elsewhere, hence no extra software security is currently required to prevent data breaches or leaks. These are some of the following considerations that will be taken throughout the project:

- Collecting and storing user chat history and user data must comply with data protection laws such as the General Data Protection Regulation (GDPR).
- Informed and signed consent that the user is talking to a bot and clear information on how the collected data will be used.[19]
- Ensure that chatbot respect cultural sensitives and differences to avoid unintentional offences to users.[20]
- Establish ethical guidelines for the chatbot's responses, including how it handles sensitive topics, mental health discussions, and potential triggers for users who have experienced loss or trauma.[21]

6 Timetable

Figure 2 depicts a Gantt Chart for the project's planned timeline for term 1. Any remaining tasks will be addressed during the winter break, as the second term is dedicated to refining the chatbot and integrating machine learning to the project.

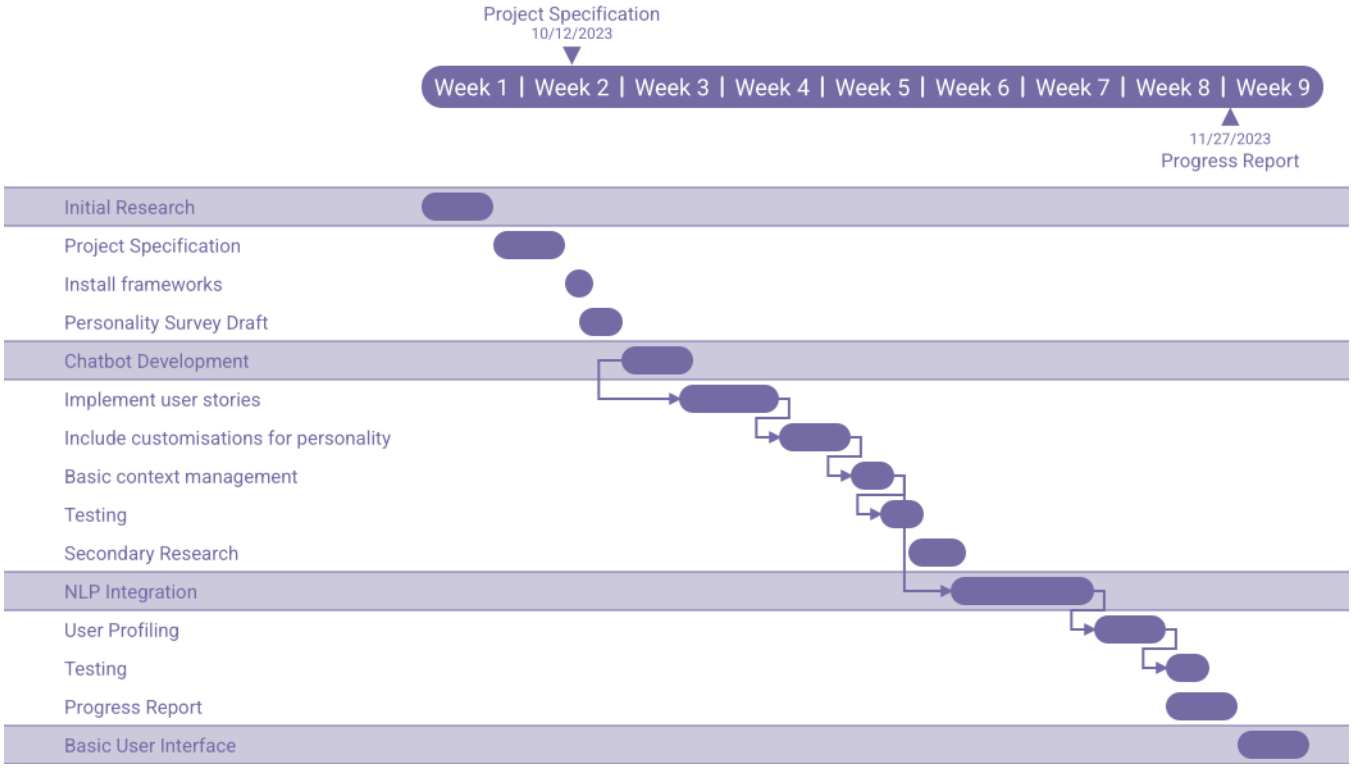

Figure 2: Gantt Chart

7 References

1. Gilburt, H., Rose, D. & Slade, M. The importance of relationships in mental health care: A qualitative study of service users' experiences of psychiatric hospital admission in the UK. BMC Health Serv Res 8, 92 (2008). https://doi.org/10.1186/1472-6963-8-92
2. https://www.clearfocustraining.co.uk/blog/how-poor-communication-can-affect-mental-health/
3. https://www.childrensmentalhealthweek.org.uk/news-and-blogs/2021/february/understanding-the-links-between-communication-difficulties-and-mental-health-challenges/
4. https://www.calmclinic.com/anxiety/impairs-communication
5. Simon D'Alfonso, AI in Mental Health https://www.sciencedirect.com/science/article/pii/S2352250X2030049X
6. https://www.betterhealth.vic.gov.au/health/healthyliving/talking-through-problems Accessed : 07/10/2023
7. Neff, K., 2011. Self-compassion: The proven power of being kind to yourself. Hachette UK.
8. https://www.psychologyjunkie.com/how-to-spot-each-myers-briggs-personality-type-in-conversation/
9. https://www.16personalities.com/articles/lets-talk-to-other-personality-types-the-art-of-conversation
10. https://marutitech.com/chatbots-work-guide-chatbot-architecture/ Accessed : 08/10/2023
11. https://www.ibm.com/watson-advertising/thought-leadership/machine-learning-chatbot
12. https://blog.on-page.ai/how-chatbots-generate-responses/
13. https://towardsdatascience.com/personality-for-your-chatbot-with-recurrent-neural-networks-2038f7f34636
14. Nuruzzaman, M. and Hussain, O.K., 2018, October. A survey on chatbot implementation in customer service industry through deep neural networks. In 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE) (pp. 54-61). IEEE.
15. https://help.replika.com/hc/en-us/articles/4410750221965-How-does-Replika-work- Accessed : 09/10/2023
16. https://beta.character.ai/ Accessed : 09/10/2023
17. https://rasa.com/docs/
18. https://botpenguin.com/glossary/feedback-chatbot Accessed : 09/10/2023
19. https://www.codecademy.com/article/ethics-of-chatbots
20. https://www.chatdevelopers.com/blog/ethics-and-responsible-ai-in-chatbot-development
21. Coghlan S, Leins K, Sheldrick S, Cheong M, Gooding P, D'Alfonso S. To chat or bot to chat: Ethical issues with using chatbots in mental health.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10291862/