Fine-Tune Large Language Model for Behavioral Activation Chatbot

—

FEINABSTIMMUNG EINES GROSSEN SPRACHMODELLS FÜR EINEN CHATBOT MIT VERHALTENSAKTIVIERUNG

Seminar Thesis

of

Mingze Li

Date of Submission

2384385

05.03.2024

At the Department of Economics and Management

Institute of Information Systems and Marketing (IISM)

Engineering Seminar: Human-Centered Systems

Reviewer: Prof. Dr. Alexander Mädche

Supervisor: Florian Onur Kuhlmeier, Sven Scheu

Table of Contents

List of Figures

1 Introduction

In the realm of treating affective disorders, like anxiety or depression, Behavioral Activation (BA) stands out as a therapeutic method focused on engaging individuals in psychologically beneficial activities, deterring them from harmful ones, and addressing obstacles to positive experiences (Dimidjian et al., 2011, p. 3). The advent of chatbots, software applications designed for simulating conversations (Adamopoulou & Moussiades, 2020), presents a novel avenue for delivering BA. Leveraging Large Language Models (LLMs), which are machine learning model trained on vast datasets to interact in human-like language (Cloudflare, 2023), offers an unprecedented opportunity to tailor mental health interventions.

The use of chatbots introduces several advantages, including personalized interaction, immediate feedback, enhanced user engagement, and high accessibility (Chandel et al., 2019). These aspects, combined with the capability for real-time data analysis, and the ease of integrating incentive mechanisms which is inspired by MUSS (Rohani et al., 2020, p. 9), underscore the potential of chatbots to improve mental health support. However, the challenge lies in effectively adapting LLMs to meet the nuanced requirements of BA within a chatbot framework, ensuring that interactions are not only relevant and sophisticated but also therapeutically beneficial.

This seminar thesis explores the development of a chatbot designed to leverage Large Language Models (LLMs) for BA therapy, addressing the gap between LLMs' broad capabilities and the therapy's specific needs. The aim is to enhance digital mental health interventions by integrating a chatbot that combines LLM's conversational AI strengths with the personalized and targeted approach required for effective BA, thereby contributing to improved mental well-being through digital innovation. To address the research question within the BA context, a BA Chatbot has been designed and developed.

2 Theoretical Foundation

The BA Chatbot leverages the Retrieval-Augmented Generation (RAG) framework and a fine-tuned GPT-3.5-turbo model, blending state-of-the-art natural language processing with the specific needs of BA. This approach underscores the potential of AI to transform therapeutic practices, providing a personalized, effective tool for mental health support.

2.1 Behavioral Activation

Behavioral Activation (BA) is a popular targeted psychotherapy aimed at overcoming depression by changing behavior. BA seeks to enhance engagement in rewarding activities, reduce depressive behaviors, and tackle barriers to positive experiences, while techniques such as activity scheduling, mood tracking, etc. are utilized to shift behavior towards more adaptive patterns, addressing both direct and indirect factors contributing to depression (Dimidjian et al., 2011, p. 3). Extensive research and empirical evidence have demonstrated that BA is as effective as medication and more traditional forms of psychotherapy in treating depression (Dimidjian et al., 2011, p. 4). BA is a cost-effective and straightforward approach that focuses on changing behaviors without the need for complex psychological analysis. This simplicity makes it accessible to a wide range of individuals.

2.2 Retrieval Augmented Generation

A Retrieval-Augmented Generation (RAG) is a mechanism that extracts relevant information from external sources and incorporates it into the core framework (Lewis et al., 2020). The RAG can be simplified as three key steps: retrieval, integration, and generation. In the BA Chatbot, the RAG retrieves relevant information from external sources using cosine similarity, integrates the retrieved results with user input during prompt construction, and then inputs this combined information into a Large Language Model (LLM) to generate answers. By employing RAG, the BA Chatbot not only enhances its performance through an expanded knowledge base derived from external sources, but also achieves greater accuracy and relevance in its responses. Additionally, this approach allows for potential updates to the data, ensuring the information remains current. Therefore, the tasks of this chatbot are implemented based on the RAG framework.

2.3 GPT 3.5

GPT, developed by OpenAI, is a transformer model designed for text generation, employing pre-training and fine-tuning methods. It employs a method of pre-training and fine-tuning to understand and generate human language. GPT-3 (Brown et al., 2020) and its successor, GPT-3.5 Turbo (OpenAI, n.d.-c), are trained as autoregressive language models with billions of parameters, boasting in-context learning capabilities. The GPT-

3.5 Turbo model is capable of understanding and generating both natural language and code. It has been optimized for conversational applications through the Chat Completions API, though it also performs well on non-chat tasks. Models in the GPT series, including ChatGPT based on GPT-3.5 and subsequent versions, have garnered significant attention for their outstanding natural language processing capabilities (Ye et al., 2023, p. 22). GPT-3.5-turbo is under consideration for chatbot because it is anticipated to be the most suitable model for most users, offering optimal results and ease of use.

2.4 Fine-tuning

Fine-tuning (OpenAI, n.d.-a) is the process of adjusting a pre-trained model's parameters slightly to adapt it to specific tasks or datasets, thereby improving its performance on those tasks. GPT-3.5 Turbo (OpenAI, n.d.-c) offers enhanced processing speed and efficiency for handling complex data, making it ideal for both large-scale and real-time applications. It stands out for its cost-effectiveness, versatility in chat and non-chat tasks, and superior performance in applications ranging from chatbots to content creation. Furthermore, its ease of integration through the Chat Completions API (OpenAI, n.d.-b) makes it a flexible tool for developers in optimizing natural language processing applications. These advantages make GPT-3.5 Turbo an ideal option for fine-tuning and deploying natural language processing applications, from consumer-facing chatbots to enterprise-level automation tools.

3 Related Work

A mental health LLM-based BA-based AI chatbot has been designed to deliver personalized BA and remote health monitoring (Rathnayaka et al., 2022). While the chatbot showcases advanced cognitive skills and utilizes a robust user interface (UI) for scheduling activities, it lacks the incorporation of incentive mechanisms to prevent users from disengaging with BA therapy. While MUBS offers a novel approach to BA for depressive disorders by providing personalized activity recommendations via a smartphone-based system(Rohani et al., 2020), it lacks the integration of chatbot technology. This omission overlooks the benefits of chatbots, such as personalized interaction, immediate feedback, enhanced engagement, constant accessibility, and the capability for real-time data analysis and incorporation of incentive mechanisms.

Chatbots' interactive nature and support availability anytime and anywhere significantly contribute to improving the user's therapy experience and adherence.

The BA Chatbot's contribution lies in its innovative integration of mood tracking, activity recommendations, schedule management with Google Calendar API, and four incentive mechanisms with LLM capabilities to fulfill the specific needs of BA, enhancing digital mental health interventions.

4 Methods

The BA Chatbot's architecture is built on the Retrieval-Augmented Generation (RAG) framework, encompassing retrieval, integration, and generation stages. It begins by using cosine similarity to retrieve relevant information from two external sources, which is then combined with user input to construct prompts. These prompts are structured to guide the model's responses, incorporating Instructions that set the task, User Input that adds specificity, and the anticipated Response that showcases the model's output diversity and relevance. This composite data is processed by GPT-3.5-turbo, enhancing the chatbot's functionality by leveraging varied external sources to enrich its knowledge base, thus improving the accuracy and relevance of its responses.

The BA Chatbot integrates two main feature sets:

- LLM-based features: BA introductions， activity recommendation， four incentive mechanisms ， mood intervention， mood tracking
- Google Calendar API based feature： schedule management.

4.1 External Sources

RAG allows the BA Chatbot to reference and integrate external data sources when answering questions, significantly expanding its knowledge base beyond the information contained in its training dataset. There are 2 external sources available: Knowledge Hub and Schedules Hub.

4.1.1 Knowledge Hub

The knowledge hub features two columns: 'Entities' and 'Description.' The 'Entities' column lists key terms, which cover 94 emotions like fear and happiness, concepts such as BA, the BA Chatbot functions for constructing prompts like activity recommendations, and user commands for schedule management including 'add schedule,' 'get schedule,' and 'delete schedule.' The 'Description' column provides detailed explanations for each entity. With this knowledge hub, the BA Chatbot is equipped to retrieve meaningful information pertinent to the user's context.

4.1.2 Schedules Hub

The Schedules Hub includes 10 columns closely associated with Google Calendar Events: event_id, summary, start_date_time, end_date_time, time_zone, location, status, attendees, creation_date, and last_modified_date. Users are required to fill in these data fields when managing schedule information to Google Calendar.

4.2 BA Chatbot Tasks

The BA Chatbot encompasses four core tasks: mood track, activity extract, schedule management, and a main task, all rooted in the RAG framework, fundamentally designed for question-answering. The user's input serves as the question, prompting the RAG to retrieve relevant information encoded from external sources via cosine similarity and FAISS. This retrieved data, merged with user input during prompt construction, is then fed into a LLM to generate responses. The process involves invoking the OpenAI API's `client.chat.completions.create` method, specifying the model and messages. The system extracts and returns the first response choice as the answer within the given conversation context.

4.2.1 Mood Track

Mood Track (Huguet et al., 2016) is used to detect user’s emotion. Mood Track is a LLM-based feature and is crucial as it allows the chatbot to detect depression or anxiety, enabling it to offer immediate comfort, suggestions, or emotion regulation techniques, thus supporting the user's mental well-being. Although each user input returns the most relevant information from the Knowledge Hub, often matching the expressed emotion stored as external knowledge within the Knowledge Hub, it remains essential to proactively analyze the user's mood every five user inputs.

This process begins by retrieving chat history, which serves as the foundation for emotion analysis. Using the `get_answer_with_single_question` function, the system compares the chat history against predefined emotions in the Knowledge Hub, ensuring precision through a similarity threshold. The analysis involves constructing a prompt that instructs the BA Chatbot to identify the user's emotion based on the chat history, followed by capturing the current timestamp with `get_current_timestamp()` to mark the assessment moment. An emotion entity, comprising the deduced emotion and timestamp, is then formed and added to Knowledge Hub for organized data management. This systematic procedure culminates in a return statement that combines the emotion entity and its description, allowing for efficient mood tracking and intervention.

4.2.2 Activity Extraction

Activity extraction is a critical process that precedes Schedule Management, designed to discern the user's intent by analyzing their latest input and comparing it against predefined commands in the Knowledge Hub, like the approach used in mood tracking. Initially, the system retrieves the user's query to understand their scheduling request, whether it's to add, get, or delete an event. It then accesses the Knowledge Hub to identify possible actions and employs a similarity measure to accurately deduce the user's intended action. Upon recognizing the action, the system captures the current timestamp, marking the precise moment the action was identified. This action, along with the timestamp, is consolidated into a singular entity, "action and timestamp," which is subsequently added to the Schedules Hub. The data is organized and sorted for straightforward retrieval. Finally, the system generates feedback, providing a concise summary of the identified action and its timestamp, facilitating an efficient and user-friendly scheduling process.

4.2.3 Schedule Management

Schedule Management, utilizing the Google Calendar API, facilitates adding, retrieving, and deleting Google Calendar events. Activity Extraction matches user commands (e.g., add, get, delete schedule) with data in the Knowledge Hub to accurately process scheduling requests. This ensures responsive and precise handling of calendar tasks based on user inputs. Three key functions are outlined to support these scheduling management efficiently.

The Add Schedule function allows users to create and add a new event to their Google Calendar by collecting event details such as summary, location, start/end times, time zones, and attendees through user inputs. It validates these inputs, sets defaults for any missing values, and checks for valid date/time and email formats. If the inputs are valid, it uses the Google Calendar API to insert the event and updates Schedules Hub with the event's information, including IDs, timestamps, attendees, and reminders. Invalid inputs or API errors result in feedback to the user without adding the event. The Get Event Id function retrieves all instances of a specific event from a user's Google Calendar using the event's unique ID. It initializes the Google Calendar API service with credentials from a 'token.json' file. If successful, it returns the event instances; otherwise, it catches and returns an error message, indicating any issues encountered during the API call. The Delete Event function synchronizes event deletions between a user's Google Calendar and Schedules Hub. It uses the event's ID to remove it from the calendar via the Google Calendar API and then updates the Schedules Hub by excluding the deleted event. The process returns an updated Schedules Hub alongside a confirmation or error message, ensuring that event data remains consistent across both the calendar and local storage.

4.2.4 Main Task

In the main tasks, the implementation of prompts within the BA Chatbot focuses on LLM-based features. Additionally, the main tasks involve invoking mood tracking and schedule management functionalities. This integration ensures a comprehensive approach to BA, leveraging the strengths of language model-based features for personalized interaction and Google Calendar API for efficient schedule management. To augment user engagement and motivation, the BA Chatbot incorporates four strategic incentive mechanisms:

1. **Public Declaration**: encourage users to share their goals and activities publicly. The social pressure and potential for public accountability can be a strong incentive.
2. **Partner Supervision**: encourage users to share their goals and activities to their friends and family so that they can supervise users' activities.
3. **Schedule Reminding (Dimidjian et al., 2011)**: remind users to check in their completed schedule.
4. **Compliment and Praises**: encourage users when they complete activities and remind users to remember and share the joy of successful completion of schedules.

These mechanisms are designed to enhance user engagement, ensure accountability, and boost motivation by leveraging social dynamics, offering positive reinforcement, and providing personal support, thereby augmenting the effectiveness of BA in mental health interventions.

Transitioning to the chatbot's technical setup, the process will be detailed from initialization to interactive loops, illustrating the seamless integration of technology with therapeutic principles. Initial steps involve importing necessary libraries, ensuring authentication for Google Calendar API access, and setting up the OpenAI client with a specific fine-tuned model. The BA Chatbot greets users with a default BA instruction, encouraging open communication. Prompt construction is tailored to guide the chatbot in providing concise, context-based responses, handling scheduling commands, recommending activities, and offering support in case of detected depression or anxiety with following instruction:

- **General Instruction**: You are a helpful chatbot that based on BA treatment. Your answer must be less than 3 sentences. Given some context, you must give a suitable response based on the context.
- **BA Introduction**: Explain BA understandably at first and chatbot makes a self-introduction.
- **API Instruction**: You used Google Calendar API so that user can use order to add event, get event and delete event. Add event order is add schedule. Get event order is get schedule. Delete event order is delete schedule.
- **Activity Recommendation**: Find activities that user likes and encourage the user to take part in them.
- **Mood Intervention**: If depression or anxiety is detected, chatbot need offer comforting words, suggestions, or techniques for emotion regulation.
- **Four incentive mechanisms**: Public declaration, partner supervision, schedule reminding, and compliment and praises

The chat operation commences with a loop for user inputs, allowing commands related to schedules or emotional support, and continues with mood tracking and schedule management based on these inputs. Responses generated by the OpenAI client consider the user's mood, activity preferences, and any schedule-related actions, ensuring relevant and supportive interactions. Throughout the session, messages are added to the history for accuracy in responses, with the history saved for future reference under specific conditions. This operational flow underscores the chatbot's capability to merge BA with advanced technology, ensuring accessibility and therapeutic efficacy.

4.3 Fine-tuning

Fine-tuning is the process of adjusting a pre-trained model's parameters slightly to adapt it to specific tasks or datasets, thereby improving its performance on those tasks (OpenAI, n.d.-a).

4.3.1 Data Preprocessing for Fine-Tuning

The dataset "verhaltensaktivierung.parquet," provided by supervisors Florian Onur Kuhlmeier and Sven Scheu, originates from a rule-based chatbot study. It is intended for use in developing an LLM-based chatbot. The result of data preprocessing for fine-tuning is . The preprocessing includes five key steps:

1. Select rows where ‘flow id’ equals 'verhaltensaktivierung-2' that contain the most important data that we want to use in the fine-tuning.
2. Narrow down to rows where content type is either 'text', 'question', or 'payload', as they are key components of prompt construction and sort the dataset by ‘created at’.
3. Transfer the selected dataset by ‘Payload’ with ‘direction’ mapping to extract conversations into a format that can be used by the LLM.
4. Process format error checks and perform data analysis according to the OpenAI cookbook guidelines (Wu & Fishman, 2023).
5. Remove examples that violate OpenAI content policies.

4.3.2 Fine-tuning Procedures

Fine-tuning GPT-3.5-turbo for the BA Chatbot involves a detailed process aimed at customizing the model's responses to better align with BA context, enhancing its ability to communicate effectively and engage users. The process begins by initializing the OpenAI client, setting the stage for subsequent data handling and interaction with the model. The next step involves uploading the training data with a specific purpose tagged as "fine-tune," ensuring the model is tailored to meet the unique requirements of BA applications. Upon uploading, the file object's ID is stored and printed, serving as a crucial reference throughout the fine-tuning process.

A fine-tuning job is then created using the file object ID, with the model type specified for customization. The initiation of this job generates a unique ID, which is printed for tracking purposes, allowing for continuous monitoring of the job's progress. Periodic checks are conducted to retrieve the status of the fine-tuning job, ensuring close oversight of its development. Finally, the outcome of the fine-tuning job is assessed; a successful job results in printing the name of the fine-tuned model, while a failure prompts a notification. This systematic approach not only tailors the BA Chatbot's responses to be more context-relevant and engaging but also enhances the overall user experience by making interactions more personalized to the user's current emotional state.

5 Results

The evaluation of the BA Chatbot, conducted through a SurveyMonkey user survey with 25 diverse participants, aimed to assess its mood tracking accuracy, impact on BA, and technical performance. This section delves into the survey results, revealing how the BA Chatbot supports mental health management and where improvements are needed. The findings underscore the BA Chatbot's positive influence on users' well-being and identify technical issues as areas for enhancement to further improve user experience.

5.1 User Survey

In Figure 1, the BA Chatbot has been evaluated by user survey supported by SurveyMonkey collected from 25 test participants. These participants ranged in age from 22 to 38 years and comprised students and professionals in fields such as mathematics, computer science, information systems, and economics. Their educational backgrounds varied from bachelor’s degrees to doctorates, with a gender distribution of 12 males and 13 females.

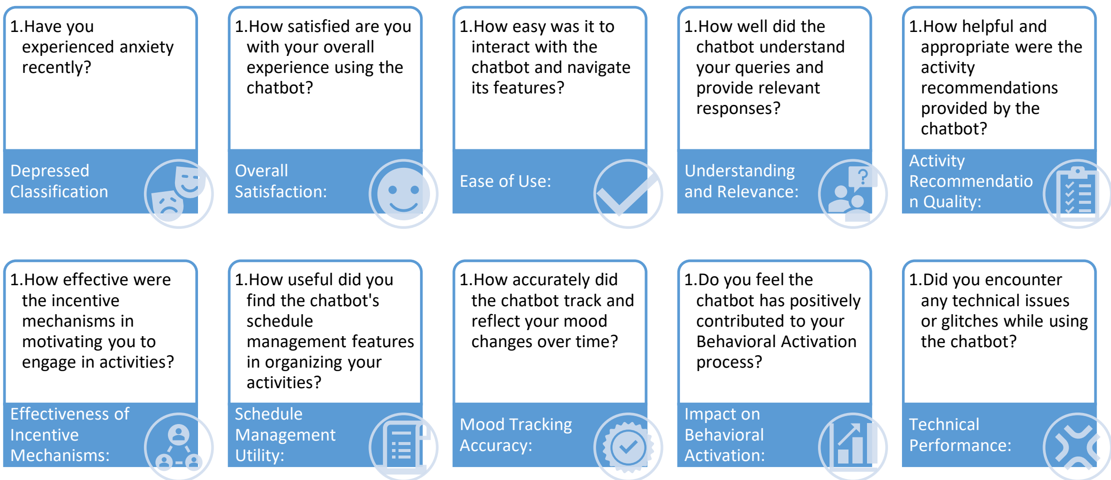

Figure 1: Ten Questions from the User Survey

5.2 Evaluation

Figure 2 has displayed the evaluation result. Mood tracking accuracy is considered high or very high by 60% of respondents, although there is room for improvement to enhance precision and reflect mood changes more accurately. The impact on BA is overwhelmingly positive, with 92% feeling that the BA Chatbot has contributed positively to their process, indicating a significant impact on users' mental health and well-being. However, 20% of users reported encountering technical issues or glitches, pointing to areas where the BA Chatbot's technical performance could be improved.

Overall, the survey results suggest that the BA Chatbot effectively supports users in managing their mental health, providing relevant information and suggestions, and engaging users through its features. However, there are opportunities for enhancements in mood tracking accuracy and technical stability to further elevate the user experience.

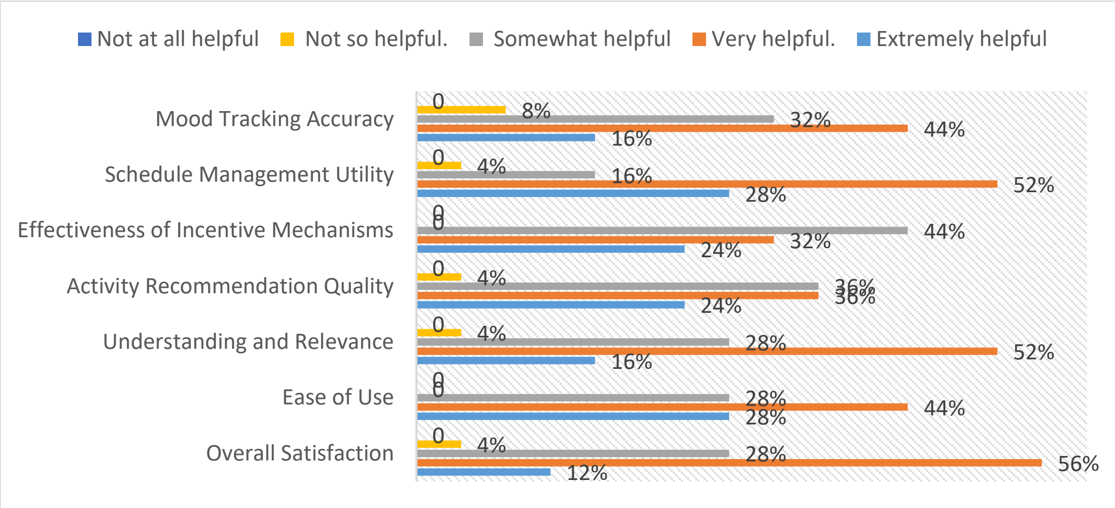

Figure 2: User Survey Results

6 Discussion

This discussion aims to explore the distinct components that constitute the BA Chatbot, specifically focusing on its dual feature sets: the LLM-based features that include BA introductions, activity recommendations, incentive mechanisms, mood intervention, and mood tracking; alongside the Google Calendar API based feature for schedule management. Through an analysis of these elements and a comparison with findings from other studies, the goal is to elucidate the efficacy and unique contributions of the BA Chatbot in the context of BA.

6.1 LLM-based Features

In Figure 3, this example from the "20240301182954.jsonl" file vividly illustrates the efficacy of the LLM-based features. Through a seamless integration of BA introductions, personalized activity recommendations, strategic incentive mechanisms, mood intervention, and continuous mood tracking, the BA Chatbot adeptly supports a user expressing feelings of loneliness and disconnection. By encouraging the user to engage in enjoyable activities such as walking and offering empathy alongside practical advice for overcoming depressive thoughts, the BA Chatbot not only aids in mood elevation but also fosters a sense of accomplishment and belonging. This interaction underscores the BA Chatbot's potential to significantly enhance user engagement in positive behavioral practices, demonstrating its value as a tool for mental health improvement.

Figure 3: Example of User-Chatbot Interaction

6.2 Google Calendar API Based Feature

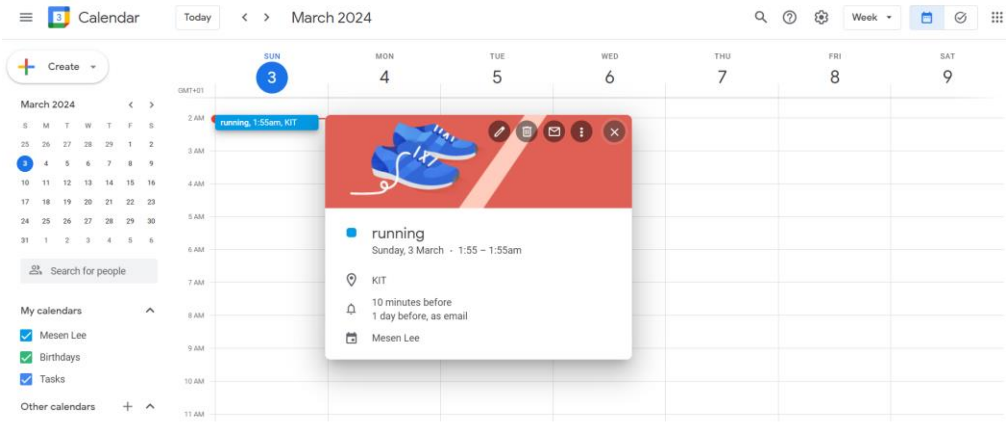

Figure 4: Scream Shot of Add Schedule

An example extracted from the JSONL file "20240303020427.jsonl" illustrates the user's interaction with the BA Chatbot through commands such as "add schedule," "get schedule," and "delete schedule." In this scenario, depicted in Figure 4, a schedule is created on the user's Google Calendar using default values for the current location and date. Subsequently, the user retrieves the schedule and then proceeds to delete it, as demonstrated in Figure 5, where the schedule is successfully removed from the user's Google Calendar. This sequence of actions showcases the BA Chatbot's capability to efficiently manage schedules by leveraging default parameters for creation, retrieval, and deletion of calendar entries.

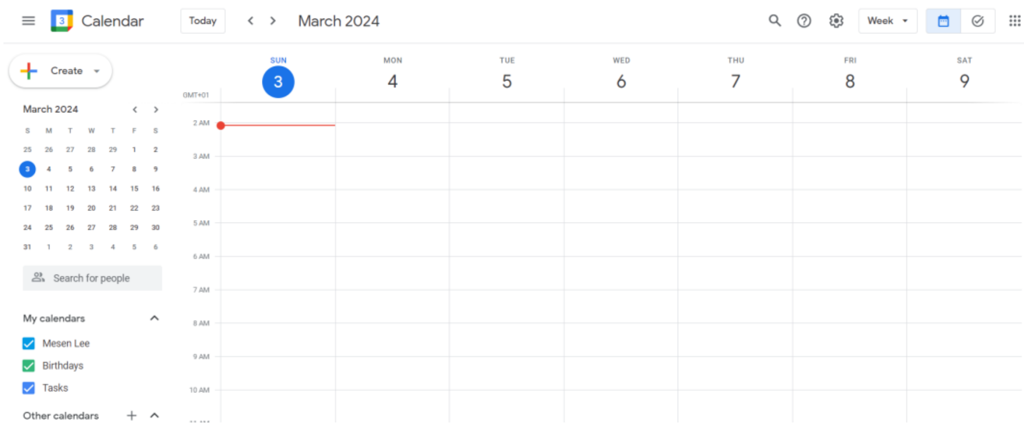

Figure 5: Scream Shot of Delete Schedule

6.3 Comparative Analysis with Existing Solutions

Unlike the BA-based AI chatbot (Rathnayaka et al., 2022), the BA Chatbot incorporates incentive mechanisms, addressing the critical need to sustain user engagement and prevent dropout from BA therapy. Furthermore, by leveraging the familiar and widely-used Google Calendar for scheduling and reminders, the BA Chatbot overcomes the limitations of proprietary UIs and enhances user experience by facilitating participation in activities with others and ensuring adherence through familiar notification systems. In contrast to MUBS (Rohani et al., 2020), the inclusion of chatbot technology in our system not only offers real-time, personalized interaction but also enables immediate feedback and continuous support, harnessing the full potential of chatbots to significantly improve therapy outcomes and user compliance. This comparative analysis underscores our BA Chatbot's unique blend of personalized therapy delivery and practical, user-friendly scheduling, illustrating its potential to advance the field of digital mental health support.

7 Conclusion

The development of a BA Chatbot utilizing a fine-tuned GPT-3.5 model combined with the RAG framework represents a significant advancement in digital mental health solutions, specifically tailored to Behavioral Activation (BA). The BA Chatbot not only introduces users to BA principles but also offers personalized activity recommendations, incorporates four strategic incentive mechanisms, and provides comprehensive schedule management and mood tracking functionalities. By catering to the unique needs of BA, this chatbot stands as a testament to the potential of leveraging advanced AI technologies to enhance the effectiveness of BA interventions, offering a promising tool for individuals seeking support in managing their mental well-being.

Bibliography

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. IFIP international conference on artificial intelligence applications and innovations,

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877-1901.

Chandel, S., Yuying, Y., Yujie, G., Razaque, A., & Yang, G. (2019). Chatbot: efficient and utility-based platform. Intelligent Computing: Proceedings of the 2018 Computing Conference, Volume 1,

Cloudflare. (2023). *What is a large language model (LLM)?* Cloudflare.com. https://www.cloudflare.com/zh-cn/learning/ai/what-is-large-language-model/

Dimidjian, S., Barrera Jr, M., Martell, C., Muñoz, R. F., & Lewinsohn, P. M. (2011). The origins and current status of behavioral activation treatments for depression. *Annual review of clinical psychology*, *7*, 1-38.

Huguet, A., Rao, S., McGrath, P. J., Wozney, L., Wheaton, M., Conrod, J., & Rozario, S. (2016). A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PloS one*, *11*(5), e0154248.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., & Rocktäschel, T. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

OpenAI. (n.d.-a). *Fine-tuning*. OpenAI. https://platform.openai.com/docs/guides/fine-tuning

OpenAI. (n.d.-b). *Introduction*. OpenAI. https://platform.openai.com/docs/api-reference/introduction

OpenAI. (n.d.-c). *Models*. OpenAI. https://platform.openai.com/docs/models/overview

Rathnayaka, P., Mills, N., Burnett, D., De Silva, D., Alahakoon, D., & Gray, R. (2022). A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors*, *22*(10), 3653.

Rohani, D. A., Quemada Lopategui, A., Tuxen, N., Faurholt-Jepsen, M., Kessing, L. V., & Bardram, J. E. (2020). MUBS: A personalized recommender system for behavioral activation in mental health. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,

Wu, M., & Fishman, S. (2023). *Data preparation and analysis for chat model fine-tuning*. OpenAI. https://cookbook.openai.com/examples/chat_finetuning_data_prep

Ye, J., Chen, X., Xu, N., Zu, C., Shao, Z., Liu, S., Cui, Y., Zhou, Z., Gong, C., & Shen, Y. (2023). A comprehensive capability analysis of gpt-3 and gpt-3.5 series models. *arXiv preprint arXiv:2303.10420*.

Affidavit

Ich versichere hiermit wahrheitsgemäß, die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie (KIT) zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, March 12, 2024

Mingze Li

Mingze Li