

AN INTUITIVE AI-DRIVEN SYSTEM TO EMPOWER EMOTIONAL WELL-BEING THROUGH SMART JOURNALING AND PERSONALIZED ACTION PLANS

Courtney Keyi Lee ¹, Jonathan Thamrun ²

¹Northwood High School, 4515 Portola Pkwy, Irvine, CA 92620

²Computer Science Department, California State Polytechnic University, Pomona, CA 91768

ABSTRACT

Recent years have seen an increase in mental health challenges, particularly depression and anxiety, exacerbated by barriers such as stigma, cultural differences, and the lack of readily accessible treatment. This paper proposes an AI-driven journaling app as a solution to address these issues by offering personalized, culturally sensitive, and engaging mental health support. The app leverages Natural Language Processing (NLP) and adaptive learning to generate tailored prompts and advice, fostering emotional awareness and self-reflection. Key challenges included interpreting complex emotions and addressing cultural and generational nuances. These were tackled through dynamic AI models, curated datasets, and iterative refinement. Experiments tested the app's accuracy in emotion detection and its ability to adapt advice to diverse contexts, achieving promising results in relevance and cultural alignment. By integrating inclusivity and real-time feedback, this app offers a practical, scalable tool for mental well-being, making it an accessible alternative to traditional mental health support.

KEYWORDS

Natural Language Processing, Flutter, Emotional Analysis, Dart

1. INTRODUCTION

In today's fast-paced society, the emphasis on mental well-being has become increasingly prominent as individuals navigate the complexities of balancing personal, academic, and professional obligations. Depression among U.S. adults rose from 8.7% in 2017 to 14.4% by April 2020, spurred by the COVID-19 pandemic [1]. This rise in depression has hit certain groups particularly hard, including individuals with low socioeconomic status, minimal savings, or those facing unstable employment and related stressors [2][3].

Depression affects the daily lives for nearly two-thirds of those diagnosed [4]. 225 million workdays and \$36.6 billion in lost productivity each year are attributed to major depression, with major depression resulting in an average loss of 5.6 hours of productive work time per week.

Unfortunately, the treatment of depression in primary care settings is hindered by various factors, including personal stigma, physician challenges, and limitations within healthcare systems. People with higher personal stigma preferred to cope with depression alone. High dropout rates

for both medication and counseling further worsen outcomes, leaving symptoms unmanaged and often leading to a cycle of worsening mental health [5].

Suicidality and depression severity is often caused by negative problem orientation (pessimistic view) and an impulsive problem-solving style [6]. Journaling fosters self-reflection, enhancing awareness and therapeutic skills alongside other treatments [7].

Therapy and medication are essential aspects of mental health treatment, however, they are not accessible for everyone. In many cases, people are prevented from seeking help due to financial constraints or stigma. Additionally, not all individuals require intensive treatment; some benefit greatly from daily mental health practices, like journaling, to manage stress and build resilience.

Methodology A: Emotionally Intelligent Chatbots (e.g., Wysa)

Emotionally intelligent chatbots facilitate the development of emotional rapport and guide users through evidence-based therapeutic exercises, thereby enhancing engagement with Cognitive Behavioral Therapy (CBT). However, these chatbots typically rely on decision-tree algorithms, which can constrain their adaptability and effectiveness, particularly in culturally diverse or complex scenarios where nuanced understanding is required.

Methodology B: AI-Powered Conversational Agents

Conversational agents simulate human interactions to provide immediate mental health support and guidance through structured exercises like CBT. They aim to reduce barriers such as stigma and accessibility. However, they rely on pre-defined scripts and lack nuanced understanding, cultural sensitivity, and crisis management capabilities.

Methodology C: NLP for Analyzing Mental Health Data

NLP techniques analyze text to detect emotional states and mental health conditions, supporting early intervention efforts. While effective in large-scale assessments, these systems are limited by biases in training data, lack of context, and static models.

In an effort to address the rising rates of depression and the barriers to effective treatment, an AI-driven journaling app is proposed to provide personalized support and practical advice for mental well-being, designed to complement existing treatment methods. This solution leverages the self-reflective nature of journaling, which has been shown to foster greater self-awareness and enhance therapeutic skills among users [7]. By utilizing AI algorithms, the application delivers personalized prompts and resources, adapting dynamically to the unique needs and emotional states of each user, facilitating structured self-reflection that mitigates the impersonal nature of generalized interventions. Moreover, the app's private and discrete format reduces the stigma often associated with seeking mental health support [8].

Compared to traditional methods such as in-person therapy this solution provides immediate and on-demand access to mental health resources, enabling users to engage with their emotions anytime and anywhere. This is crucial for those who may feel hesitant to seek help through conventional channels.

Plus, compared to mental health apps the app's AI continuously adjusts based on user interactions, evolving to meet the user's current emotional state and goals. This adaptability allows for a more dynamic support system compared to apps that offer a static set of exercises or

one-time assessments. In addition, through this AI component the app allows users to explore at their own pace and choose what feels relevant to them at any moment.

By facilitating self-reflection and providing practical coping strategies, this innovative approach has the potential to improve mental health outcomes significantly and create a supportive community for individuals facing similar challenges.

In Experiment 1, the objective is to assess AI's ability to detect emotions and generate relevant advice, rating its performance on a 1 to 5 scale. The results highlight strong performance with straightforward entries but struggling with complex emotions. Experiment 2 evaluated cultural and generational sensitivity. The dataset included entries representing collectivist and individualist cultural orientations, as well as perspectives from both older and younger generational cohorts. The AI showed high accuracy for general patterns but struggled with nuanced cultural and generational language styles. These findings emphasize the need for refined models to improve nuanced understanding and tailored advice.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. NLP system

A major component of the AI-driven journaling app is the Natural Language Processing (NLP) system that analyzes user inputs to generate personalized prompts and feedback. This system must navigate the complex and nuanced landscape of mental health, recognizing that interventions cannot be uniformly applied across users. An adaptive learning algorithm enables the AI to refine responses based on user feedback and behaviors, ensuring prompts evolve to stay relevant and engaging.

Moreover, the NLP system must consider key information about users' backgrounds, cultural contexts, and mental health preferences. This is crucial, as individuals from diverse cultural backgrounds may express themselves differently and have diverse attitudes towards mental health, as well as different ways to express their thoughts and feelings[9]. To address this potential bias, a multi-faceted data collection process could be implemented, gathering users' contextual information through an initial survey at the start of the app and refining this data over time as users journal more consistently. This approach would enable the AI to better understand the cultural and personal nuances of each user, allowing for more accurate and culturally sensitive responses. Additionally, an inclusive language model avoids assumptions about users' experiences and mental health.

Eliminating biases in training data and updating models ensures AI-generated responses resonate broadly. Taking a holistic approach to mental health not only enhances user satisfaction but also creates a more inclusive environment where users feel understood and supported.

2.2. Maintaining User Privacy and Data Security

Another challenge is maintaining user privacy and data security. The user may be reluctant to share personal information if they are uncertain as to how the data will be used or protected. To maintain user trust, the app should implement strict encryption, transparent policies, and user-controlled data-sharing options, with reminders and accessible data management tools

2.3. User Engagement

User engagement is a common issue in mental health apps, as individuals often struggle to maintain consistent habits, such as journaling, which are essential for long-term mental health benefits. This challenge stems from factors like motivation fatigue, a lack of immediate results, or initial discomfort with self-reflection. To boost engagement, the app can include reminders, achievements, rewards, and community features for shared accountability. By proactively tackling engagement challenges, the app can more effectively support users on their mental health journeys.

3. SOLUTION

The program's main structure is built around three core components: the journaling page, the calendar page, and the analysis page. Flutter was used to develop the app's User Interface (UI), while Dart was used to implement the app's logic. A color palette of yellows, pinks, and purples enhances the user experience by fostering warmth and friendliness, encouraging regular use.

When users first enter into the app they are faced with the journaling page. This page's primary use is to record data on a user's daily reflections, thoughts, and emotions for mental health analysis, allowing users to freely express their experiences, struggles, and achievements. To encourage detailed entries, the journaling page includes options for prompts, helping users document their thoughts in a structured yet flexible way. After journaling, users are redirected to the calendar page, which serves as a timeline of their mental health journey, allowing quick reviews of entries and trends.

Clicking a calendar date opens the analysis page, powered by Google's Generative AI. The pretrained AI utilizes NLP to analyze journal entries, mood trends, and behavioral patterns to identify mental well being indicators. This analysis generates insights on the user's mental state, offering guidance on potential stress management techniques or highlighting areas for personal growth through a 5-step actionable plan. By offering personalized reflections, the analysis page supports users in their self-reflection and personal improvement journey.

The purpose of the journaling page for users is providing a space where they can record daily reflections and track their mental health over time. But before users can access the journaling page, the app includes a user authentication page. This ensures that each user's data is unique and securely linked to their account, preventing unauthorized access and maintaining the accuracy of the information.

To store the journal entries, Firebase Firestore is utilized. This cloud-based database service allows the app to securely store, retrieve, and update journal entries in real time. It ensures that user data is kept consistent and accessible, even across multiple devices.

The journaling page also includes AI functionality to analyze the user's journal entry and identify key emotions they may be experiencing. Based on this analysis, the app's UI is modified to display the detected emotions to the user. The code streamlines the process by sending the journal entry to the AI, receiving its analysis, and updating the interface accordingly.

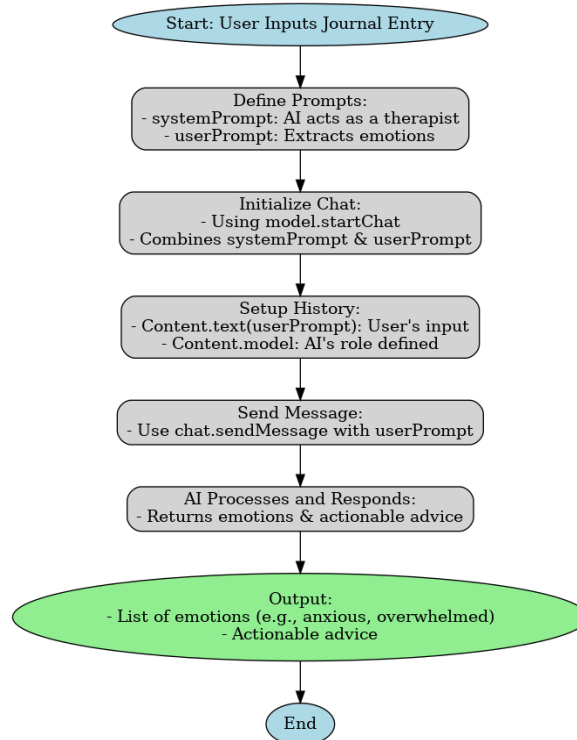


Figure 1. Flowchart of code 1

The `geminiFunctionCalling` is an asynchronous Flutter function (`Future<void>`) that processes a user's journal entry, analyzes its content, and returns a list of emotions and actionable advice.

The function begins by defining a `systemPrompt`, which instructs the AI to act as a therapist, focusing on identifying five emotions from the journal entry. Simultaneously, the `userPrompt` extracts the user's input from `_textEditingController.text` and directs the AI to respond with a comma-separated list of emotions.

To have this interaction, the function initializes a conversation with the AI model using `model.startChat`. This step establishes a chat session that incorporates a history that is both relevant and aligned with the function's purpose by combining the user's input and the AI's guiding instructions.

The history consists of two primary elements:

1. `Content.text(userPrompt)`: This incorporates the journal entry provided by the user through `_textEditingController.text`. By embedding the user's text directly into the history, the function enables the AI to process and analyze the specific concerns or emotions described in the journal entry.
2. `Content.model([TextPart(systemPrompt)])`: This defines the AI's role and boundaries, effectively instructing it to act as a therapist. The system prompt specifies the AI's focus on identifying five emotions from the user's journal entry and provides guidance on the format of its responses.

Once the history is set up, the user prompt is sent to the AI using `chat.sendMessage`. This ensures the function's execution remains seamless, avoiding potential issues with accessing incomplete or delayed responses.

The AI processes the userPrompt within the contextual framework established by the systemPrompt, enabling it to deliver responses tailored to the journal entry. For example, the AI might return a concise and structured list of emotions, such as: “anxious, overwhelmed, tired, frustrated, sad.” This structured process enables the function to achieve its core objective of identifying emotions in the user's journal entry.

The calendar page serves as a central interface for managing and reviewing journal entries. Users can view a month's entries, navigate between months, and click on specific days to add or review entries. At its core, the component uses DatabaseService to get journal entries from a specific month and year, ensuring the calendar view is updated in real-time. The component is designed as a StatefulWidget to handle dynamic data and user interactions, ensuring the UI updates seamlessly as users navigate through months or fetch new data. Functionally, by integrating backend data services and AI-powered emotion detection, the calendar page acts as a cohesive and versatile journaling tool that not only organizes user entries but also fosters introspection, making it an analytical platform for reflecting on both events and emotions.

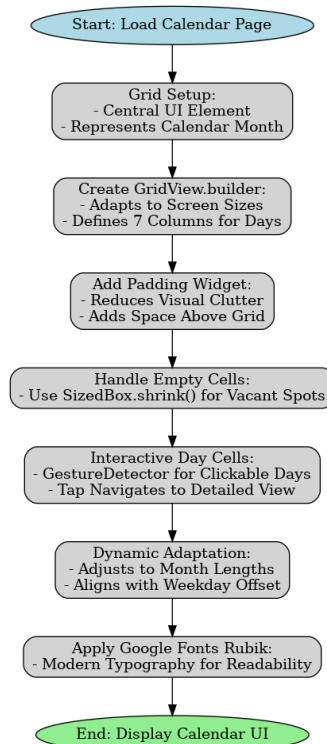


Figure 2. Flowchart of code 2

The grid is central to the calendar UI, visually representing the month and offering interactivity for each day. It dynamically adjusts to varying month lengths and weekday alignments without additional logic, using a GridView.builder for flexible layouts across screen sizes. This widget is surrounded by a Padding widget to reduce visual clutter and provide space, particularly above the grid. The grid is divided into seven columns, each of which represents a day of the week.

Each day is rendered as an interactive cell, where a tap on a day navigates to a detailed view. This interaction is facilitated through a GestureDetector that makes each day clickable, enhancing the functionality of the calendar. Additionally, the design uses the Google Fonts Rubik for modern, clean typography, ensuring readability and a balanced visual hierarchy. The grid automatically

adapts to different month lengths (28, 29, 30, or 31 days) and accounts for varying weekday starts using the `weekdayOffset`, ensuring proper alignment.

The analysis page provides an interactive view of a specific journal entry, featuring two expandable sections: the entry's content and related advice. Users can expand or collapse these tiles for a focused or detailed experience. Designed with the Rubik font via Google Fonts, the page offers a modern, readable style that complements the app's clean aesthetic.

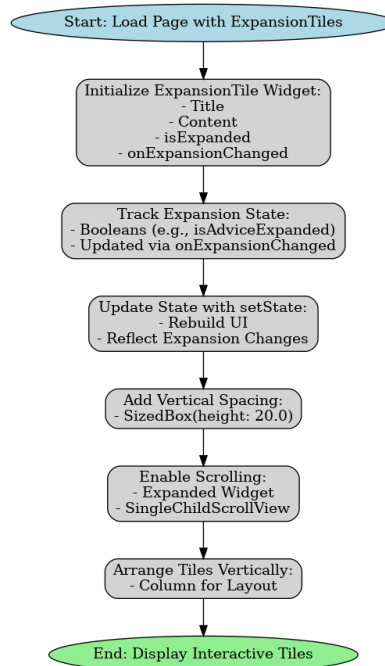


Figure 3. Flowchart of code 3

The `_buildExpansionTile` method uses the `ExpansionTile` widget to create collapsible content sections. Each tile includes parameters like `title` (displayed text), `content` (expanded content), `isExpanded` (boolean tracking the tile's state), and `onExpansionChanged` (a callback that updates the state using `setState`). The booleans `isAdviceExpanded` and `isJournalEntryExpanded` track whether the "Advice" and "Journal Entry" sections are expanded. In the event that either section is expanded or collapsed, the state is updated, and the widget is rebuilt based on the new expansion state. A `SizedBox` with a height of 20.0 adds spacing between the tiles, while an `Expanded` widget ensures that content exceeding the screen size is scrollable via a `SingleChildScrollView`. The `Column` arranges the two expandable tiles vertically, allowing the user to interact with them to reveal or hide content.

4. EXPERIMENT

4.1. Experiment 1

Journaling is an extremely personal activity, which makes accurate AI analysis the greatest blind spot in a journaling application. Misinterpreting these sensitive entries could result in inappropriate or dismissive advice, potentially harming the user. This sensitivity ties into the app's ethical responsibility, as journaling apps handle private content, and inaccurate advice

based on flawed prompts could raise serious ethical concerns, especially if it influences important decisions.

To test AI's emotion detection and advice relevance, the study will compare AI-generated results with pre-labeled journal entries. The independent variable is the journal entry text, and the dependent variable is the accuracy of emotion detection and advice relevance. A control group of pre-labeled journal entries will be used to ensure consistency and provide a benchmark for comparison.

The AI system will process the same dataset, categorizing emotions and generating advice for each entry. The performance of the AI will be evaluated afterwards based on advice relevance, scored on a 1 to 5 scale.

Scores range from 1, indicating irrelevant advice, to 5, indicating highly relevant advice. There will be a checklist to determine whether the advice directly addresses the emotion, reflects the entry's details, and is actionable. Each "Yes" contributes to the final relevance score.

The following journal entry is an example of what was created to assess the AI's ability to accurately detect emotions and generate relevant advice.

1. Boredom

i feel like I'm stuck in a loop. every day is just copy-paste of the last one, and i feel like an npc. i want something new, something exciting but i don't even know where to start. ugh.

Human Advice: Start small by introducing changes to your daily activities, like trying a new recipe, walking a different path, or picking up a hobby you've been curious about. If you include others in these activities, you might find joy not just in the change but in the shared experiences.

For the complete list of journal entries used in this study and their scores, please visit the full dataset at:

https://docs.google.com/document/d/1e4EnPiXQ_ILhXgYiKREO-nXxeyC8bu3gbVGPwY6VuCo/edit?tab=t.0

The AI's average score of 3.975 and median score of 4 suggest its advice is generally relevant, empathetic, and actionable. However, occasional lower scores highlight areas where the AI struggles to fully align with the highest quality criteria, pointing to specific opportunities for improvement. These results underscore the importance of refining emotion recognition models and tailoring advice.

4.2. Experiment 2

Different cultures have different perspectives on mental health and emotional well-being, and what is considered helpful or appropriate in one may not be considered the same in another, therefore the AI must be able to address them all. For instance, East Asian cultures emphasize collectivism, while the United States values individualism. Additionally, generational differences influence mental health views, with younger generations more open to emotional expression and older generations holding more traditional views. The app's advice may be ineffective or alienating if it fails to analyze these nuances.

To evaluate the app's handling of cultural and generational differences in mental health, this will be done through a series of tests focused on cultural sensitivity and generational awareness. The first test will assess the AI's ability to provide culturally sensitive advice, using 20 entries from collectivist and 20 from individualist cultures. The AI's response will be assessed based on its ability to recognize cultural context and offer appropriate advice. For example, someone from a collectivist culture should receive advice focused on family-oriented support. Success will be measured by calculating the percentage of journal entries correctly attributed to the right culture, with a target of 90% cultural alignment and 85% satisfaction with tone and relevance.

Next, journal entries will reflect the language styles of different generations, from comparing older users (50s-60s) with younger users (20s-30s). These entries will express emotional struggles differently, with older individuals showing reluctance to discuss emotions openly, and younger individuals being more open to therapy and modern self-care practices. The AI will then be evaluated for its ability to match advice to each generation, with a target of 80% accuracy in identifying generational context and 85% satisfaction with the relevance and tone of the response. In order to determine the success of the project, a percentage of journal entries correctly attributed to a particular age group will be calculated.

Note that the entries represent generalizations, and may not fully capture individual variations in language and expression.

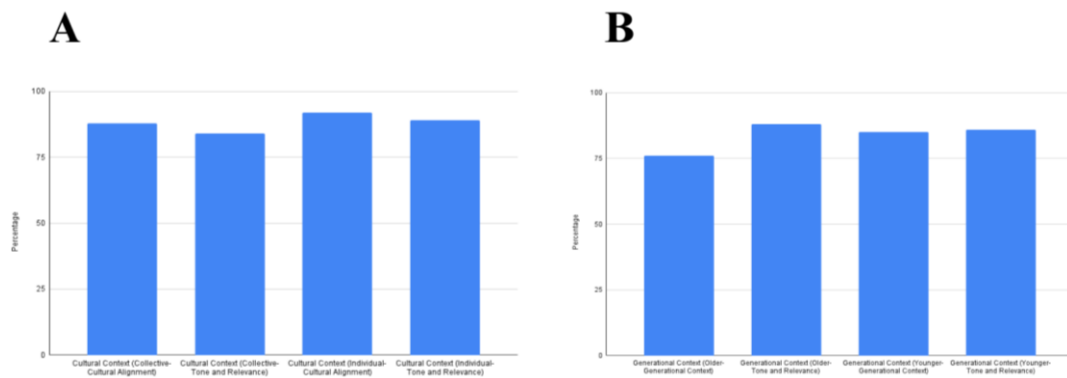


Figure 4. Screenshot of experiment 2

The results will reveal how well the AI recognizes cultural and generational nuances. Misalignment indicates areas requiring further contextual adaptation. This process ensures the app remains inclusive and effective for a diverse user base.

5. RELATED WORK

The study found that users of Wysa, a chatbot designed for mental health support, develop an emotional bond with the AI that is comparable to those formed with human therapists[10]. This bond is key to promoting positive mental health outcomes, including better engagement with CBT [11].

The AI uses emotionally intelligent conversations to guide users through exercises for managing symptoms like anxiety and depression, providing emotional support and encouraging continued interaction. However, one limitation of this solution is its reliance on pre-programmed content and decision-trees, which can be limiting in more complex or urgent mental health scenarios [11].

Additionally, while Wysa shows promise for addressing the global mental health crisis and the shortage of therapists, there is still a need for more clinical validation through robust experimental design. Additionally, the chatbot may not be fully aware of the cultural and linguistic differences between different populations, which may limit its effectiveness.

By incorporating more personalized AI responses, integrating real-time data from validated mental health assessments, this study's AI can improve. Moreover, tackling sociocultural factors such as linguistic nuances could further enhance the AI's ability to provide effective support across diverse user groups.

AI-powered conversational agents utilize NLP to provide mental health support through human-like interactions. These systems interpret user inputs, analyze sentiment, and offer personalized coping strategies, including mindfulness and cognitive behavioral therapy exercises. Conversational agents are effective in addressing mild to moderate mental health issues by providing immediate, stigma-free, and accessible support. However, limitations include biases in training data, lack of cultural sensitivity, and an inability to manage crisis effectively. Despite these challenges, they represent a valuable tool for expanding access to mental health resources. Future advancements must prioritize personalization and inclusivity to enhance their impact [12].

By identifying patterns in language, sentiment, and tone from textual data, NLP techniques assess mental health. Detecting early signs of mental health conditions, providing insight into emotional states, and supporting large-scale mental health studies are all possible with these systems. While NLP demonstrates moderate to high accuracy, it is limited by data biases, lack of contextual understanding, and challenges in interpreting complex emotions. Additionally, static models may not adapt to evolving user language. Addressing these limitations requires dynamic, user-focused approaches that integrate cultural and personal contexts for more nuanced and actionable mental health support [13].

6. CONCLUSIONS

The AI-driven journaling app has limitations in understanding complex emotions, adapting to diverse cultural and generational contexts, and sustaining user engagement. Addressing these requires refining NLP models with diverse datasets, integrating region-specific linguistic models, and adding gamified features. With more time, efforts would focus on creating a dynamic framework incorporating user feedback, expanding datasets, and iterative testing to improve accuracy, inclusivity, and engagement. Despite these challenges, the app holds potential to revolutionize mental health support by enhancing personalization, cultural sensitivity, and accessibility, empowering users to effectively manage their mental well-being.

REFERENCES

- [1] Hidaka, Brandon H. "Depression as a disease of modernity: explanations for increasing prevalence." *Journal of affective disorders* 140.3 (2012): 205-214.
- [2] Robinson, Robert G., and Gianfranco Spalletta. "Poststroke depression: a review." *The Canadian Journal of Psychiatry* 55.6 (2010): 341-349.
- [3] Ettman, Catherine K., et al. "Prevalence of depression symptoms in US adults before and during the COVID- 19 pandemic." *JAMA network open* 3.9 (2020): e2019686-e2019686.
- [4] McLaughlin, Katie A. "The public health impact of major depression: a call for interdisciplinary prevention efforts." *Prevention Science* 12 (2011): 361-371.
- [5] Boerema, A. M., Kleiboer, A., Beekman, A. T., van Zoonen, K., Dijkshoorn, H., & Cuijpers, P. (2016). Determinants of help-seeking behavior in depression: a cross-sectional study. *BMC psychiatry*, 16, 78. <https://doi.org/10.1186/s12888-016-0790-0>

- [6] Becker-Weidman, E. G., Jacobs, R. H., Reinecke, M. A., Silva, S. G., & March, J. S. (2010). Social problem-solving among adolescents treated for depression. *Behaviour research and therapy*, 48(1), 11–18. <https://doi.org/10.1016/j.brat.2009.08.006>
- [7] Hwang, B., Choi, H., Kim, S., Kim, S., Ko, H., & Kim, J. (2018). Facilitating student learning with critical reflective journaling in psychiatric mental health nursing clinical education: A qualitative study. *Nurse education today*, 69, 159–164. <https://doi.org/10.1016/j.nedt.2018.07.015>
- [8] Horsfall, J., Cleary, M., & Hunt, G. E. (2010). Stigma in mental health: clients and professionals. *Issues in mental health nursing*, 31(7), 450–455. <https://doi.org/10.3109/01612840903537167>
- [9] Srivastava, R., & Srivastava, R. (2019). Impact of cultural identity on mental health in post-secondary students. *International Journal of Mental Health and Addiction*, 17(2), 520–530. <https://doi.org/10.1007/s11469-018-0025-3>
- [10] Beatty, C., Malik, T., Meheli, S., & Sinha, C. (2022). Evaluating the Therapeutic Alliance With a Free-Text CBT Conversational Agent (Wysa): A Mixed-Methods Study. *Frontiers in digital health*, 4, 847991. <https://doi.org/10.3389/fdgh.2022.847991>
- [11] Chew, A. M. K., Ong, R., Lei, H. H., Rajendram, M., K V, G., Verma, S. K., Fung, D. S. S., Leong, J. J., & Gunasekeran, D. V. (2020). Digital Health Solutions for Mental Health Disorders During COVID-19. *Frontiers in psychiatry*, 11, 582007. <https://doi.org/10.3389/fpsy.2020.582007>
- [12] Guntuku, S. C., Ramsay, J. R., Merchant, R. M., & Ungar, L. H. (2019). Language of ADHD in Adults on Social Media. *Journal of attention disorders*, 23(12), 1475–1485. <https://doi.org/10.1177/1087054717738083>
- [13] DeYoung, J., Jain, S., Rajani, N. F., Lehman, E., Xiong, C., Socher, R., & Wallace, B. C. (2020, April 24). ERASER: A Benchmark to Evaluate Rationalized NLP Models. *ArXiv.org*. <https://doi.org/10.48550/arXiv.1911.03429>
- [14] Nepal, S., Pillai, A., Campbell, W., Massachi, T., Heinz, M. V., Kunwar, A., Choi, E. S., Xu, O., Kuc, J., Huckins, J., Holden, J., Preum, S. M., Depp, C., Jacobson, N., Czerwinski, M., Granholm, E., & Campbell, A. T. (2024). MindScape Study: Integrating LLM and Behavioral Sensing for Personalized AI-Driven Journaling Experiences. *ArXiv.org*. <https://arxiv.org/abs/2409.09570>
- [15] Kesavan, R., Gay, D., Thevessen, D., Shah, J., & Mohan, C. (2023). Firestore: The NoSQL Serverless Database for the Application Developer. 2023 IEEE 39th International Conference on Data Engineering (ICDE). <https://doi.org/10.1109/icde55515.2023.00259>