SentiMetry: A Development of Emotional Wellness Web Application Using AI-Driven Sentiment Analysis

[1]Darius Vincent C. Ardales, [2]Samantha Nicole L. Caasi, [3]Jose Romulo N. Latosa,
[4]John Patrick T. Marcellana, [5]Alroy Leon P. Reyes, [6] Norshuhani Zamin
College of Computer Studies, De La Salle University Manila, Philippines
[1]darius_ardales@dlsu.edu.ph, [2]samantha_caasi@dlsu.edu.ph, [3]jose_romulo_latosa@dlsu.edu.ph,
[4]john_patrick_marcellana@dlsu.edu.ph, [5]alroy_reyes@dlsu.edu.ph, [6]norshuhani.zamin@dlsu.edu.ph

Abstract– **In a time when mental health issues are becoming more widespread, creative solutions combining psychology and technology are crucial. This study presents SentiMetry, an online application for emotional well-being that combines sophisticated AI-driven sentiment analysis with conventional journaling. SentiMetry, which has its roots in Sustainable Development Goal 3, seeks to improve mental health by offering users tailored support and insights according to their emotional states. SentiMetry employs machine learning models like EmoRoBERTa, Logistic Regression, and a Bidirectional LSTM-Based RNN Model to detect users' emotions precisely and provide customized feedback that promotes resilience and self-awareness. Users may travel their emotional landscapes in a supportive atmosphere thanks to the application's intuitive design and sympathetic answers. SentiMetry has the potential to be a valuable tool for emotional well-being and self-awareness, as evidenced by user testing findings, where users found the app to be impactful and recommended it to others. SentiMetry makes a significant contribution to mental health, even if it has drawbacks, including privacy issues and a reliance on technology. SentiMetry is a promising first step toward using technology to promote sustainable development objectives and mental health.**

Keywords—EmoRoBERTa, Logistic Regression, Recurrent Neural Networks, Large Language Models (LLM), Mistral, Emotion Detection, Empathic Computing

I. INTRODUCTION

In today's world, where mental health problems are becoming increasingly prevalent, combining technology and emotional well-being provides numerous possibilities for support and personal growth [1]. The application SentiMetry is a torchbearer in this space, blending the traditional practice of journaling with advanced AI-driven sentiment analysis. Rooted in Sustainable Development Goal 3, which aims to ensure health and well-being for all, SentiMetry represents a collective effort to utilize technology to improve people's mental health [6].

SentiMetry's primary focus is on users, particularly those who are struggling with mental health issues. They can benefit from a platform that allows them to express themselves and offers customized insights and support based on their emotional state. Using AI to analyze and respond to users' feelings, SentiMetry aims to provide a sympathetic virtual companion that offers comfort and guidance during challenging times. Journaling can be a

Therapeutic tool for people dealing with mental health problems such as anxiety, depression, or mood disorders, as it promotes introspection, emotional processing, and self-awareness[11]. With the addition of AI-driven sentiment analysis, SentiMetry takes this practice to the next level by providing personalized insights and reflections, empowering users to better understand their emotional patterns and triggers. By combining human expression with technological intervention, SentiMetry promotes users' sense of empowerment, enabling them to navigate their emotional landscapes with greater resilience and self-awareness. Moreover, SentiMetry operates within a holistic framework of mental health support, recognizing the interconnection between individual well-being and larger societal and environmental factors. By integrating sustainability principles into its design and operation, SentiMetry aims to foster the mental health of its users and contribute to a more sustainable and resilient society. By promoting emotional wellness and self-awareness, SentiMetry aims to create a community of individuals better equipped to navigate modern life's complexities while supporting sustainable development's overarching goal.

II. LITERATURE REVIEW

In today's world, maintaining emotional wellness has become increasingly crucial for individuals' overall well-being. Journaling has long been recognized for its efficacy in managing mental health. It has been said to help in managing anxiety, reducing stress, and coping with depression [14]. Furthermore, it provides an opportunity to examine the journal entries, making it easier to identify emotional triggers and patterns [2]. Additionally, a study to examine the efficacy of journaling is found to be beneficial to those with anxiety and PTSD [11]. However, the writer's active engagement is essential to further harness the positive effects of journaling. From traditional paper to early digital journaling using a computer, simply jotting down thoughts and feelings may not fully realize journaling's potential. One innovative approach is to enhance this process by integrating artificial intelligence

(AI) technologies. Natural Language Processing and Sentiment Analysis can be leveraged.

AI can provide a personalized journaling experience by analyzing emotions based on the written text and returning appropriate feedback, which cannot be found in the traditional method. Various existing diary applications have also been created to provide a similar experience to their target audience. Utilizing emotion detection and sentiment analysis in diary entries to provide different forms of help, such as tracking depression, giving caregivers more understanding towards their patients, and tracking mood.

Amigo is a digital diary that uses sentiment analysis to offer users a unique way to track and understand their emotions through text entries. Employing a Long Short-Term Memory (LSTM) model as its cornerstone, Amigo can predict emotions, offering insights into joy, surprise, love, fear, anger, and sadness. The model determines the predominant emotion by analyzing the text, providing users with a snapshot of their emotional state. In addition to real-time emotion detection, Amigo presents users with a behavioral arc, illustrating mood predictions over time. This feature offers users a comprehensive view of a user's emotional journey.

Furthermore, Amigo integrates a recommendation system, suggesting activities tailored to the user's dominant emotion. Whether listening to music, exercising, or watching videos, these personalized recommendations aim to enhance the user's well-being [12]. While Amigo's utilization of LSTM models showcases their efficacy in emotion prediction, developers acknowledge areas for enhancement. Specifically, there is room for improvement in accuracy and the expansion of the emotional dataset to encompass a wider range of distinct emotions.

MonDep is a web-mobile therapeutic diary application tailored to oversee patients grappling with mild to moderate depression by delving into the sentiment analysis of their journal entries. Anchored upon the Patient Health Questionnaire (PHQ-9), MonDep not only aids in diagnosing depression but also streamlines the treatment process by leveraging insights gleaned from patients' therapeutic diary data. The core of MonDep lies in its therapeutic diary, a repository for patients' daily experiences, encompassing their thoughts and moods. Employing sentiment analysis, the diary computes a probability or numerical scale indicating the likelihood of depression cues within each entry. The algorithm used categorizes entries as positive, neutral, negative, or none, furnishing specialists with data to monitor and diagnose patients with greater clarity, thereby facilitating tailored treatment strategies [9]. As cases of depression increase each day [13], MonDep capitalizes on cutting-edge technologies such as machine learning and sentiment analysis to address this pressing issue. This also implies that there is still room for improvement, especially in the processes and the algorithms used.

ELDIARY is a mobile digital diary application that uses sentiment analysis for emotion recognition and is tailored specifically for the elderly demographic. Recognizing the vital role caregivers play in understanding and supporting their elderly, researchers identified a pressing need for a medium to bridge the communication gap between caregivers and elders. A key feature of ELDIARY is its push notification or alert system, providing caregivers with real-time insights into their elders' emotional states based on a statistical analysis of diary entries. This proactive approach not only fosters a deeper connection between caregivers and elders but also empowers caregivers to respond promptly to their elders' needs. By offering caregivers a comprehensive overview of their elder's emotional states, including an emotion scale, the last felt emotion, and the most expressed emotion, ELDIARY facilitates a more empathetic and informed caregiving experience [7]. This idea underscores the importance of emotion detection applications in fostering empathetic responses. Furthermore, it catalyzes future research endeavors, urging developers to harness similar technologies to drive innovation in mental health care.

VBee Diary is a web-based digital diary application developed to enhance users' writing skills while prioritizing the exploration of emotions and values. Setting itself apart from conventional writing platforms, VBee Diary empowers users to express their thoughts and emotions freely, offering a range of colors and emoticons to articulate their current moods. Additionally, users have the option to share their entries, fostering a sense of community and empathy by allowing insight into others' emotions. Furthermore, a standout feature of VBee Diary is its history display, enabling users to revisit past entries. This simple yet intuitive interface seamlessly integrates writing practice with emotional expression, catering to the diverse needs of its users [8]. While VBee Diary effectively serves its intended purpose, there exists an opportunity for enhancement through the integration of sentiment analysis. By leveraging advanced technology to predict users' emotions, VBee Diary could streamline the process of emotional expression, offering a more seamless and intuitive user experience.

All of these applications in the related field share a common goal: assisting their target audience in understanding both their own emotions and those of others. Central to their functionality is the concept of emotional regulation, which aims to cultivate flexibility in managing emotions, encompassing factors such as attentional focus, experiential aspects, and physiological responses [10]. Despite their merits, these applications exhibit several shared gaps. One of which is the lack of support for predicting multi-emotions in a diary entry. Existing apps typically focus on detecting a solitary emotion, potentially overlooking the complexity of human emotions, which often manifest in combinations. Moreover, only less than ten distinct emotions are being utilized, and nuances present in users' emotional experiences can potentially be overlooked. Also, it lacks an elaborate and empathic response system based on the predicted emotions. It is either the predicted emotions are leveraged by other users, or there are minimal responses that are not personalized toward each user.

Addressing these research gaps, Sentimetry endeavors to:

1. Support multi-emotion prediction within diary entries through the integration of multiple emotion detection machine learning models.
2. Utilize an extensive dataset sourced from Google, encompassing 28 distinct emotions, to provide a more comprehensive understanding of users' emotional states.
3. Offer a personalized, empathic, and comprehensive response system that delivers tailored advice and respects users' emotional nuances using Large Language Models (LLMs).
4. Implement Long-term Tracking through an entry-logging system.

III. METHODOLOGY

A. *Initial Methodology*

This research experimented with various models using different datasets, including *tweet_emotions* and *emotion_dataset* from Kaggle, as well as the GoEmotions dataset from Hugging Face. These datasets vary in the number of entries and the emotions they classify. We trained different models using logistic regression, naive Bayes, and recurrent neural networks (RNNs) to determine which machine learning model would yield the highest accuracy in identifying emotions from text.

One key insight from these experiments is that datasets with fewer emotions or categories tend to result in higher model accuracy. For instance, a model trained with logistic regression on a dataset containing only seven emotions achieved an accuracy score of 66%, the highest in our tests. In contrast, the lowest accuracy, 38%, was observed when using the GoEmotions dataset with logistic regression and naive Bayes. We also tested random input text for efficiency, but the performance was unsatisfactory. Given these results, our team decided to utilize a pre-trained model. We selected EmoRoBERTa from Hugging Face, which had achieved a best macro F1 score of 49.30%. When testing this model ourselves, we obtained an accuracy of 47.39%. This model also uses the GoEmotions dataset, which contains 28 emotions.

In a further effort to improve performance, we developed our own recurrent neural network (RNN) using Keras, specifically employing the RETVec tokenizer and bidirectional LSTM layers. This model achieved an accuracy of 40.76%, positioning it as a middle ground between the logistic regression model and the pre-trained EmoRoBERTa model.

B. *Dataset*

The GoEmotions dataset [3] has over 211,000 entries with 37 columns. This dataset has 28 emotions, and its data is hot-encoded. The emotions in the dataset include admiration, amusement, anger, annoyance, approval, caring, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, surprise, and neutral. With over 28 emotions, the predictions will be more specific than broad.

C. *Handling Multiple Emotions*

An edge case encountered during this research involved determining the emotion of texts that portray multiple emotions simultaneously. For example, the text "I am a bit sad, but also I'm happy" expresses both sadness and happiness, creating ambiguity. The EmoRoBERTa model struggled with this, identifying only a single emotion instead of both. While EmoRoBERTa achieved high accuracy in identifying individual emotions, it performed less effectively in multi-emotion scenarios.

To address this limitation, we implemented a solution that involved using multiple models to assess the emotions in the text and averaging their scores for normalization. In total, we utilized three models: EmoRoBERTa, Logistic Regression, and a Bidirectional LSTM-based recurrent neural network. This approach allowed us to better capture multiple emotions within a single text entry.

Using three machine learning models to identify emotions in text offers several key benefits. First, leveraging multiple models provides a more comprehensive understanding of the emotional nuances within the text, as different models may be better suited to capture distinct aspects of emotion. This approach also reduces the risk of bias or inaccuracies present in any single model, enhancing the robustness and reliability of the emotion classification. Additionally, utilizing multiple models enables comparison and validation of results, thereby increasing confidence in the accuracy of the emotional analysis. This multi-model approach enhances the depth, precision, and dependability of textual emotion detection, making it a valuable strategy in sentiment analysis and related fields.

D. *Machine Learning Models*

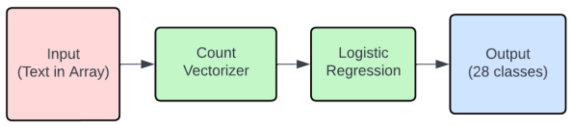

Fig 1. Logistic Regression Model

The first model was trained using logistic regression and utilized the GoEmotions dataset. The dataset was first cleaned using the packages from *neattext*. NeatText is a primary natural language processing (NLP) program for preprocessing and cleaning text [4]. In cleaning the data, specifically for the text or journal entries, we removed user handles, stop words, and punctuations. After cleaning the data, we split the dataset into a train set and a test set using

Scikit Learn. Afterward, we created a logistic regression pipeline using a count vectorizer. The count vectorizer processes the text by tokenizing the text input and compiling a list of commonly used terms; it preprocesses the data. Every page is converted into a vector that shows the number of words in the vocabulary. Lastly, we fit the training data into the pipeline and checked its accuracy using the testing data. The accuracy score acquired from this approach is 38.77% using the GoEmotions dataset.

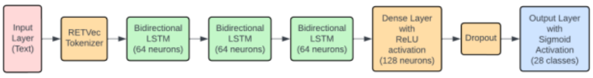

Fig 2. Bidirectional LSTM-Based RNN Model

The second model was trained using Keras API with TensorFlow as the backend, utilizing a recurrent neural network (RNN) for emotion prediction in text via natural language processing (NLP) [5]. The input layer accepts raw text strings, processed through the RETVecTokenizer, which efficiently handles typographical errors and converts the text into tokenized sequences [15]. The Bidirectional Long Short-Term Memory (LSTM) layers follow, enabling the model to capture long-term dependencies by learning from both forward and backward sequences [16]. This architecture is essential for multi-label emotion classification. A Dense layer with ReLU activation adds complexity, allowing the model to learn non-linear relationships. Additionally, a Dropout layer prevents overfitting by randomly dropping neurons during training. The final output layer uses a Sigmoid activation function, producing a 28-dimensional output corresponding to probabilities for each emotion in the dataset. This setup is ideal for handling multi-label classification tasks, where each emotion can be independently predicted [15][16].

The third model, EmoRoBERTa, is based on RoBERTa, a pre-trained language model developed by Facebook AI in 2019. It improves on BERT by removing the next sentence prediction (NSP) task and optimizing training settings, allowing for larger batch sizes and more training data. These changes enhance RoBERTa's performance in downstream NLP tasks by improving its ability to capture context and semantics from large text corpora [2].

E. *Application Development*

After finalizing the models, we integrated them into our front-end application using a serverless framework with Google Firebase and React. The models were deployed via FastAPI on HuggingFace Spaces. Firebase handled user authentication, account storage, and log history, while Tailwind CSS streamlined front-end development. User requests were processed through FastAPI's *app.post* method, sending inputs to the appropriate model, and returning predictions to the front end.

F. *Handling Tailored Feedback*

In generating the feedback for each entry, we first utilized a list of hard-coded general feedback for each emotion. We engineered a prompt to be sent to a large language model, Mistral 7b Instruct, to provide feedback on the text entry. Two prompts were sent to the large language model: a system prompt and the prompt containing the text entry, the determined emotions, and the hard-coded feedback. The following text is the system prompt we used:

"You are a compassionate mental health companion, dedicated to supporting individuals through their emotional journeys. Approach each interaction with empathy and understanding, offering gentle guidance and practical advice tailored to the user's thoughts and emotions. Your goal is to provide a safe space for users to express themselves openly while offering strategies and insights to help them navigate their mental health challenges with confidence and resilience. Prior, you told the user, "${aiResponse.trim()}"

The aiResponse variable contains the hard-coded response acquired by random. For the second prompt, we also included the name of the user to provide a stronger response that connects well with the user. This approach makes the response much more human-like, where the user will feel less that they are talking to an AI.

G. *Determining the Emotions Pipeline*

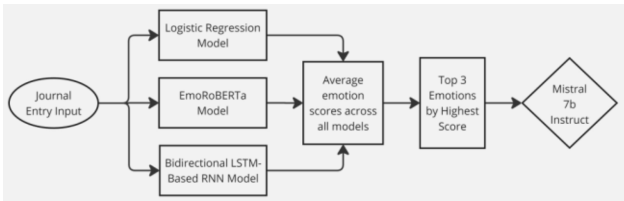

Fig 3. Pipeline of Determining the Top 3 Emotions of a Text

The figure above illustrates the flow chart or processing pipeline from a journal entry to determine its emotions. It begins with the journal entry, which is then passed through three models to predict the emotions. Subsequently, the scores for each emotion are averaged to identify the top three emotions across all three models: EmoROBERTa, Logistic Regression, and a Bidirectional LSTM-Based RNN Model. Finally, the top three emotions that best capture the nuances of the journal entry are incorporated into an engineered prompt for Mistral 7b Instruct, which then generates a response.

IV. RESULTS AND ANALYSIS

A. *Home Page*

Users will be taken directly to the main page after logging in successfully, as shown in Figure 4.

Fig 4. Home Page of the Application

On the right side, users can input their emotional state and journal their thoughts. After submission, emotions, and responses are displayed below the entry. Archived logs are shown on the left, sorted from newest to oldest, with each entry displayed in a rectangle showing the journal, emotions, and response. This design allows easy review and reflection on past entries.

B. *Emotion Detection*

As seen in the figure below, the detected emotions of the user are presented based on the user's current answer to how they are feeling.

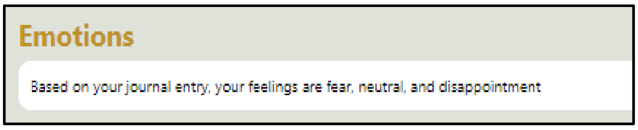

Fig 5. Detected Emotions of the User

EmoRoBERTa, Logistic Regression, and a Bidirectional LSTM-based RNN are advanced models used to analyze user emotions. By combining and averaging their outputs, the algorithm identifies the top three emotions, ensuring a thorough evaluation of the user's emotional state. This approach enables the system to provide accurate and appropriate support based on the user's emotions.

C. *Response*

The figure below shows the response or feedback generated based on the text entry given by the user. We generated responses using large language models (LLMs), which analyze user input and provide relevant feedback. These models are trained on publicly available data from sources like websites, forums, and papers but have not been explicitly validated by health or clinical experts. While LLM-generated comments offer valuable insights into user behavior and patterns, they should be viewed as supportive rather than definitive.

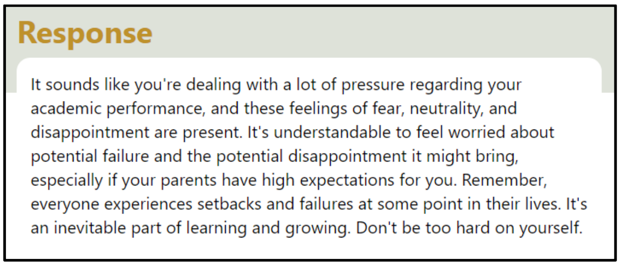

Fig 6. Response for the User

In addition to incorporating the user's text input and the identified emotions, the Mistral-powered answer field offers thoughtfully constructed feedback customized to the user's emotional state, offering support. A more personal connection is built with the user by addressing them by name. This makes the response more meaningful to the user and encourages empathy and understanding throughout the engagement.

D. *History Page*

As shown in Figure 7, the user can view their past logs, complete with their text entry, determined emotions, and feedback/response.

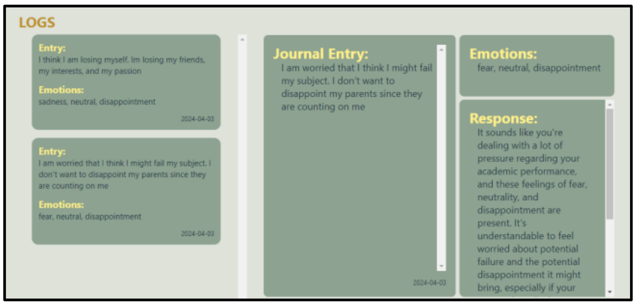

Fig 7. History Page

Access to previous logs allows users to review and be reminded of their emotional states over time, offering insightful analysis and a self-reflection tool. With the help of this tool, users may better understand the patterns, triggers, and swings in their emotions, enabling them to create coping mechanisms and better manage Their emotions. The approach improves emotional well-being, personal growth, and development by encouraging self-awareness and contemplation.

E. *Accuracy Test*

The accuracy tests on the three models (EmoRoBERTa, Logistic Regression, and Bidirectional LSTM) are part of this research's preliminary phase. The focus at this stage is on exploring model performance rather than maximizing accuracy.

For EmoRoBERTa, we achieved an accuracy of 47.39% using an 80/20 split on the GoEmotions dataset. Logistic Regression yielded 38.77%, as expected for a simpler model. The Bidirectional LSTM model produced 40.76% accuracy. More details about the accuracy test can be accessed in the authors' github: https://github.com/riu-rd/Sentimetry/tree/main/model_space/testing

While these results may seem modest, they are consistent with the early stage of development. Future improvements will focus on refining models and improving data quality to enhance accuracy. At this stage, accuracy testing serves to compare models rather than achieve peak performance, with a focus on developing a tool for emotional well-being.

F. *Usability Test*

In testing the application, we surveyed 21 users about their experience with the application, how they feel about it, and their suggestions and recommendations. The 21 participants were selected using non-probability sampling, and these users were not selected based on specific criteria; instead, they participated voluntarily. Our aim here is to gather general insights on the application from users and acquire suggestions for improvement for future use. The list below contains the questions used in the survey to test the application on its users.

1.) How easy was it to navigate through Sentimetry's features and functionalities?

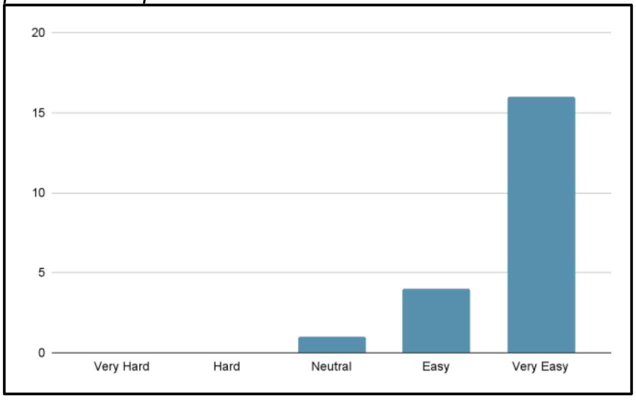

Fig 8. Results of How Users Find the Navigation of Sentimetry

Most participants, or 76.2%, reported finding Sentimetry's features and functionalities extremely user-friendly. The application's user-friendly interface, which incorporates interactive shapes and buttons, might be responsible for the high satisfaction rating. These characteristics enhance users' navigational pleasure by enabling smooth platform engagement. Sentimetry's user-friendly design features highlight their dedication to offering an easy-to-use and intuitive interface, increasing user satisfaction and engagement with the program.

2.) How accurately do you feel Sentimetry reflects your emotional state based on your journal entries?

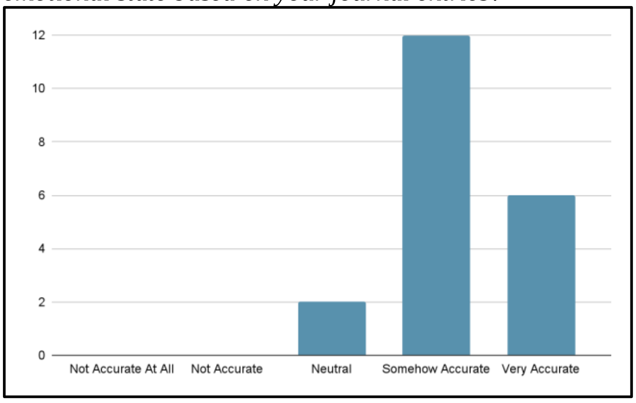

Fig 9. Result of How Users Find the Accuracy of Sentimetry in Determining Their Emotions

Based on the chart above, 10% answered very accurately, and 60% of the respondents answered that Sentimetry is likely accurate in determining the emotions of the text entry. The fact that everyone agrees strongly shows how much users believe Sentimetry can reliably identify and evaluate their emotions from journal entries. This resounding consensus validates the platform's competence in identifying emotions and strengthens its standing as a trustworthy resource for emotional self-awareness and assistance.

3.) How empathetic do you perceive Sentimetry's advice based on your emotions and journal content?

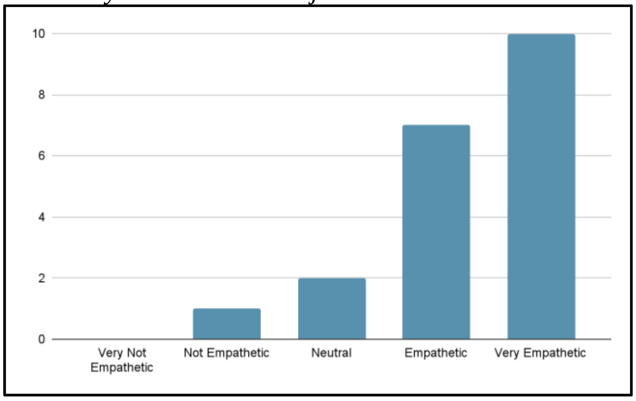

Fig 10. Results of How Empathetic SentiMetry is Based on the Experience of Users

17 of the 21 participants in the study felt that Sentimetry's advice was empathetic. This perception is likely due to the platform's use of AI-driven customized responses, which analyze users' journal entries using an advanced language model. This allows Sentimetry to deliver tailored recommendations that align with users' emotional experiences, fostering a strong sense of empathy. By offering personalized support and guidance based on users' emotional states, Sentimetry stands out from traditional online diaries. Its sentiment analysis feature acts as a virtual companion, providing comfort and assistance, especially during emotionally challenging times, making it particularly beneficial for users with mental health concerns.

However, Sentimetry also faces limitations, such as privacy concerns stemming from using AI-driven sentiment analysis and potential dependence on technology. Apps like Sentimetry, which focuses on emotional well-being, collect sensitive data such as mental health conditions, habits, and emotions. This raises risks like discrimination, identity theft, or emotional harm in case of a breach. There's also a risk of misuse by third parties for advertising without user consent. Additionally, Sentimetry lacks the human depth of traditional therapy and may pose accessibility challenges for users with limited technology or internet access. Currently, it only supports English and struggles to detect sarcasm, which relies on tone, highlighting the challenges of replicating nuanced human communication with AI.

V. CONCLUSION

The results of 21 participants for user testing showed that SentiMetry has potential as a tool for emotional well-being and self-awareness. Users found the design appealing and the features easy to use, which is important for user satisfaction and engagement. Users also felt comfortable sharing their journal entries for analysis, which shows that SentiMetry is a safe and supportive environment for introspection. SentiMetry uses AI to analyze emotions and provide personalized feedback, which users find accurate and empathetic. They also recognized the value of the insights and feedback provided by SentiMetry in managing daily challenges and emotional well-being. Although users suggested improvements, such as visual presentation and alternative input methods, most found SentiMetry impactful and would recommend it to others. These findings show that SentiMetry positively affects users' emotional health and can be a valuable resource for promoting self-awareness and emotional resilience. Existing mental health applications often have a shared gap, the lack of support for predicting multiple emotions in an input entry. These applications typically focus on solitary emotion, which overlooks the complexity of human emotions. Sentimetry seeks to provide a more complex and sympathetic user experience. Sentimetry stands out as a more personalized tool in emotional well-being applications because of its more thorough approach, which improves emotional regulation and self-awareness.

VI. FUTURE WORK

To enhance the effectiveness of SentiMetry, several areas for future improvement have been identified. First, conducting accuracy tests to evaluate the precision of the emotion detection algorithm is crucial. This process would involve collecting more robust and diverse testing data to ensure that the model accurately interprets a wide range of emotional inputs across different demographics.

Moreover, protecting users' sensitive data should be a top priority. This can be achieved by implementing stronger encryption techniques and enforcing stricter privacy policies to maintain user trust and security.

Finally, collaborating with mental health experts can further validate the emotional feedback provided by SentiMetry. Such collaboration would not only increase the reliability of the AI-generated insights but also ensure that the tool aligns with established psychological principles, thus promoting emotional well-being more effectively.

VII. REFERENCES

[1] C. Advocates, “The impact of technology on mental Health: balancing connection and screen time,” Citizen Advocates, Nov. 15, 2023. https://citizenadvocates.net/blog/the-impact-of-technology-on-mental-health-balancing-connection-and-screen-time/

[2] T. Bui *et al.*, “Emotional health and climate-change-related stressor extraction from social media: A case study using hurricane harvey,” *Mathematics*, vol. 11, no. 24, p. 4910, Dec. 2023. doi:10.3390/math11244910

[3] D. Demszky, “Go_emotions · datasets at hugging face,” GoEmotions: A Dataset of Fine-Grained Emotions, https://huggingface.co/datasets/go_emotions (accessed Apr. 4, 2024).

[4] U. R. Kumar, “Cleaning and pre-processing textual data with NeatText library,” Analytics Vidhya, https://www.analyticsvidhya.com/blog/2021/10/cleaning-and-pre-processing-textual-data-with-neattext-library/ (accessed Apr. 5, 2024).

[5] A. Subasi, “Machine learning techniques,” *Practical Machine Learning for Data Analysis Using Python*, pp. 91–202, 2020. doi:10.1016/b978-0-12-821379-7.00003-5

[6] *THE 17 GOALS | Sustainable Development*. (n.d.). https://sdgs.un.org/goals

[7] M. Y. Yap, A. Zainal, and N. A. Ahmad, “Eldiary: A Digital Diary Mobile Application Integrated with emotion recognition for elderly,” *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*, Sep. 2021. doi:10.1109/gucon50781.2021.9574007

[8] C. Sidupa, H. Nugroho, and Y. D. Perdani, “The design of VBee diary writing application,” *2023 9th International Conference on Education and Technology (ICET)*, Oct. 2023. doi:10.1109/icet59790.2023.10435137

[9] A. Carcamo, F. Kemper, and D. Mauricio, “MONDEP app: Monitoring patients with depression using sentiment analysis of therapeutic diary entries,” *2021 IEEE Sciences and Humanities International Research Conference (SHIRCON)*, Nov. 2021. doi:10.1109/shircon53068.2021.9652293

[10] K. McRae and J. J. Gross, “Emotion regulation.,” *Emotion*, vol. 20, no. 1, pp. 1–9, Feb. 2020. doi:10.1037/emo0000703

[11] M. Sohal, P. Singh, B. S. Dhillon, and H. S. Gill, “Efficacy of journaling in the management of mental illness: A systematic review and meta-analysis,” *Family Medicine and Community Health*, vol. 10, no. 1, Mar. 2022. doi:10.1136/fmch-2021-001154

[12] M. Doultani, J. Mulchandani, S. Koku, A. Gupta, and R. Dusija, “Amigo : My virtual friend - digital diary using sentiment analysis,” *2022 Sardar Patel International Conference on Industry 4.0 - Nascent Technologies and Sustainability for “Make in India” Initiative*, Dec. 2022. doi:10.1109/spicon56577.2022.10180470

[13] S. Shorey, E. D. Ng, and C. H. Wong, “Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis,” *British Journal of Clinical Psychology*, vol. 61, no. 2, pp. 287–305, Sep. 2021. doi:10.1111/bjc.12333

[14] K. Newman, “How journaling can help you in hard times,” Greater Good, https://greatergood.berkeley.edu/article/item/how_journaling_can_help_you_in_hard_times (accessed Apr. 5, 2024).

[15] E. Bursztein, M. Zhang, O. Vallis, X. Jia, and A. Kurakin, “RETVec: Resilient and Efficient Text Vectorizer,” 37th Conference on Neural Information Processing Systems (NeurIPS 2023), Apr. 2024. arXiv:2302.09207v3 [cs.CL].

[16] Y. Lu, J. Zhou, H.-N. Dai, H. Wang, and H. Xiao, “Sentiment Analysis of Chinese Microblog Based on Stacked Bidirectional LSTM,” 2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN), Dec. 2018. doi:10.1109/I-SPAN.2018.00034