Transformer Models in Digital Journaling: A Review on Mood Detection and Personalization

Jenil G. Rathod[1], Saloni M. Naik[2], Aditya S. Nalla[3], Manas B. Joshi[4], Prof. Aparna V. Mote[5]
[1,2,3,4,5]*ZCOER, Pune, MH, India*

Abstract **Conventional journaling, while useful for self-reflection, tends to be a passive process, falling short of the analysis and feedback required for users to recognize emotional patterns or achieve profound self-awareness [23]. Although recent gains in Natural Language Processing (NLP) have presented solutions, the generalizability and efficacy of text-based emotion detection remain major areas of investigation [4]. This work presents Pocket Journal, a novel, mobile-oriented Android app that seeks to make journaling an engaged, perceptual activity through the use of an advanced pipeline of Transformer-style models. The framework leverages a fine-tuned RoBERTa model for sophisticated mood prediction [1], a fine-tuned BART model for abstractive summarization [14], and incorporates a robust generative LLM (Gemini) to give users on-demand, in-depth behavioral analysis. Main issues in existing text-emotion systems are limited generalizability to different models of emotion [4], reasoning challenges in handling informal or noisy text encountered in actual entries [3], and the propensity of available recommender systems to be shallow, cross-sectional analysis instead of surface-level recommendations [16], [17]. Pocket Journal resolves this through the integration of its multi-stage AI pipeline with personalized media suggestions from third-party APIs (e.g., Spotify, TMDB), rendering a proactive buddy for the improvement of mental health and the promotion of personal development.**

Index Terms **Mental Wellness, Self-Reflection, Artificial Intelligence, Transformer Models, RoBERTa, BART, Natural Language Processing (NLP), Generative LLM**

I.INTRODUCTION

The practice of journaling is an effective self-reflection technique, but conventional digital tools remain mostly passive text editors and never give analytical feedback to enable users to learn about emotional patterns [23], [24]. This work bridges this limitation by proposing Pocket Journal, which is an intelligent mobile tool that works to make journaling an active and interactive activity. Through the utilization of an advanced pipeline of Natural Language Processing (NLP) models a fine-tuned RoBERTa model for mood classification [1], a BART model for abstractive summarization [14], and a generative LLM (Gemini) for profound behavioral insights [13] the system provides users with an unprecedented level of self-awareness. Based on a cutting-edge tech stack with a Flutter front-end and Python backend, Pocket Journal also offers tailored media suggestions based on the detected mood of the user [16], [17]. This paper introduces an end-to-end framework that transforms journaling into a proactive partner for improving mental well-being and promoting personal growth.

A. The Shortcomings of Classic Journaling and the Emergence of AI-Based Analysis

The classic journaling process, an old but tried-and-tested technique of self-reflection, is still a largely passive and unguided activity. With digital tools becoming more ubiquitous, most apps are still merely text editors, offering no analysis feedback for users to discover emotional patterns or have profound self-insight. This absence of automated knowledge is a major shortcoming since the manual analysis of unstructured text for emotional information is a very intricate and usually impossible task for the ordinary user [23], [24]. To overcome these limitations, the research area is shifting toward smart journaling systems based on Artificial Intelligence (AI) and Natural Language Processing (NLP) [5], [6], [8]. These new approaches seek to turn journaling into an engaging activity, providing users with greater insight into their emotional environment. This work presents Pocket Journal, a mobile-oriented application conceived to represent this fresh paradigm.

B. Transformers and Generative Models' Role in Journaling

At the heart of this shift are sophisticated deep learning models, specifically Transformers. RoBERTa-style transformer-based architectures have shown better performance on subtle classification tasks and are thus well suited for emotion detection in informal and personal text [1]. For summarization of lengthy journal entries into actionable insights, abstractive summarization models such as BART, based on the same underlying technology, are very effective [14]. In addition to summarization and classification, the use of generative Large Language Models (LLMs) is a major advance, allowing systems to create rich, human-sounding behavioral analysis from user data [13]. Pocket Journal combines these technologies, applying RoBERTa for mood classification, BART for summarization, and a generative LLM (Gemini) for on-demand insights into a dynamic tool that actively facilitates mental health and personal development.

II. LITERATURE SURVEY: APPROACHES IN AI-POWERED JOURNALING

The design of Pocket Journal is based on a solid foundation of previous research in various key areas in Natural Language Processing (NLP). Exhaustive literature review sheds light on the approaches and tools that have been used successfully in text-based emotion identification, abstractive summarization, and emotion-aware recommender systems. Overall reviews of the area point out the transition from simple sentiment analysis to subtle emotion detection and the important role played by AI and deep learning methods [23], [24]. This review combines certain methodologies under these categories in order to determine the state-of-the-art at present and the research gaps which Pocket Journal seeks to fill.

A. Progress in Text-Based Emotion Detection Models

Modern emotion detection research has tended to unite around deep learning models, and especially Transformers, due to their capacity to learn intricate contextual subtleties. Comparative studies have benchmarked a range of Transformer models, such as DistilBERT, XLNet, RoBERTa, and BigBird, to ascertain their performances in emotion classification in conversational text [1]. Aside from common Transformers, other architectures are also being investigated. For example, research on Graph Neural Networks (GNNs) has demonstrated how models constructed using semantic graph representations of text outperform models constructed using syntactic structures [2]. One of the biggest challenges in this area is how to deal with the noisy, casual quality of user-generated data. Towards this end, new input encoding methods have been specifically designed to enhance accuracy in micro-blog text [3]. The other crucial area of research is the emotion model's generalizability, as experiments reveal that the emotion categories and training sets are critical in determining a system's performance in various environments [4].

B. Text Summarization Methodologies

Abstractive text summarization is important to condense the essential concepts from very long journal entries, and ongoing research has centered around sophisticated neural frameworks to achieve this. Hybrid architectures, for example, the T5-LSTM FusionNet, have been put forward for improving summarization of intricate psychological text through the combination of the contextual depth of a Transformer (T5) with the sequential sensitivity of an LSTM [14]. Additionally, Large Language Model (LLM) usage has been optimized to summarize large datasets of feedback. These methods typically employ methods such as hierarchical chunking and retrieval-augmented generation (RAG) to build effective summarization and query-answering systems over large corpora of text [13].

C. Emotion-Aware Recommender Systems

One of the most significant applications of emotion detection is the design of more personalized and empathetic recommender systems. Recent studies present numerous prototypes that recommend media based on user mood. Recommendation systems have been developed to suggest books and music by using direct user mood input and content-based filtering algorithms [16], [17]. Additionally, movie recommenders based on emotions have been constructed using text-emotion mapping to map film recommendations with the emotional status of a user [15]. More sophisticated architectures have investigated the employment of "emotion vectors" from text to drive suggestions [18] and even multimodal systems that identify emotion from text, speech, and facial cues [19]. Nevertheless, such systems tend to suffer from issues such as dependence on small datasets, the "cold-start" issue for novel users, and a tendency to be unable to manage complicated or blended emotional states [16], [17].

III. PROBLEM STATEMENT: GAPS IN EXISTING SYSTEMS

In spite of the advances in NLP and the acknowledged advantages of journaling, a number of key issues prevent practical implementation and success in intelligent self-reflection aids.

1. Passive Nature and Lack of Insight in Traditional Journaling

Conventional digital journaling software primarily exists as a passive repository of text, not offering the analysis-based feedback required for users to recognize emotional patterns or access profound self-consciousness [23], [24]. This leaves users with a vast amount of unstructured personal data without providing any automatic means of extracting actionable insights, an increasing gap that today's AI can potentially bridge [5], [6].

2. Performance and Generalizability Bottlenecks in Emotion Detection

The performance of AI-based journaling largely relies upon how accurately its underlying models of emotion detection work, which are subject to a number of inherent limitations. Most models do not have the ability to generalize well across diverse sets of data and environments because selecting the categories of emotion greatly affects performance [4]. In addition, the systems tend to exhibit lower accuracy in handling the noisy, casual, and shortened language prevalent in actual personal logs, e.g., the kinds used in micro-blogs and text messages [3], [8]

3. Superficiality of Current Emotion-Aware Recommenders

Most existing emotion-aware recommender systems provide merely surface-level recommendations based on a one-time, point-in-time mood input [16], [17], [20]. They typically are plagued by the "cold-start" issue for new users and are limited by being based on small datasets [17]. Most importantly, they do not possess the ability to conduct deep, long-term analyses, failing to offer users dynamic insights based on long-term behavioral patterns [15], [18].

4. Lack of a Holistic, Proactive Wellness Solution

Today's digital wellness ecosystem is not integrated. Users tend to have to implement individual apps for journaling, mood monitoring, and watching supportive media content. There is clearly no single, comprehensive platform that easily unites deep analysis of journaling with proactive, user-tailored content suggestions. This lack of integration disallows a synergistic experience in which self-reflection actively enhances and informs a user's mental well-being journey.

Table 1: - Comparative Analysis of Emotion Detection Approaches

Metric	Lexicon-Based Approaches	Traditional Machine Learning (ML)	Deep Learning (Transformers)
Accuracy	Low to Moderate. They work well for basic sentiment polarity but are less accurate in classifying subtle emotions, usually acting as a baseline [6]	Moderate to High. Algorithms such as SVM and Naïve Bayes can perform well but are usually beaten by deep learning models on sophisticated, large-scale data [7], [12].	High to Very High. Signifies the current state-of-the-art, showing high accuracy on sophisticated tasks. For instance, research indicates accuracies of 91-93% with models such as CNNs andTransformers [1], [8].
Semantic & Contextual Nuance	Low. Hinges on existing, hand-crafted dictionaries (lexicons) and linguistic rules, which doesn't demand large, task-specific labeled training corpora [10], [23].	Low to Moderate. Can learn some local context via statistical features such as n-grams, but not a genuine, deep knowledge of sentence structure and long-range dependencies [23].	Low to Moderate. Can learn some local context via statistical features such as n-grams, but not a genuine, deep knowledge of sentence structure and long-range dependencies [23].

Data Requirement	Low. Depends on existing, pre-curated dictionaries (lexicons) and grammar rules, which does not need huge, task-specific labeled training sets [10], [23].	Very High (pre-training), Moderate (fine-tuning). Depend on transfer learning. Although pre-training needs enormous, unlabelled sets, the model can be fine-tuned later on smaller, domain-specific labeled sets [7], [24].	High (pre-training) to Moderate (fine-tuning). Based on transfer learning, in which models pre-trained on large datasets are fine-tuned using small, labeled, domain-specific data [7], [24].
Computational Cost	Low. It entails easy, quick keyword searches and rule application, which are very efficient and ideal for resource-poor setups [23].	Moderate. It is more resource-hungry than lexicon approaches but less so than training deep learning models in large volumes from scratch.	High to Very High. Training and fine-tuning large models such as RoBERTa is costly in terms of computation and demands extensive hardware resources (GPUs/TPUs), a recognized limitation in research [1], [14].
Generalizability & Transferability	Low. Lexicons are usually domain-specific (e.g., financial, product reviews) and do not transfer to new subject matter or styles of writing (e.g., personal journaling) without human maintenance [23].	Moderate. Models are heavily dependent on the particular features and data with which they were trained. They usually need to be retrained completely in order to be transferred to a new domain.	High. The underlying strength is transfer learning. A general pre-trained model on an enormous language corpus can be successfully fine-tuned to a new, specific topic with a comparably modest amount of data [4], [24].
Interpretability / Explainability	High. The decisions of the model are clear-cut and directly traceable to given keywords and rules. One can easily see why a particular emotion was predicted.	Variable. The interpretability is model-dependent. Decision Trees, for instance, are fairly easy to understand, whereas models such as Support Vector Machines (SVMs) are less transparent.	Low. Sometimes called "black box" models since their complexity renders it hard to understand why a given prediction was made [23].

IV. PROPOSED DIRECTION: THE POCKET JOURNAL AI PIPELINE

The Pocket Journal framework aims to bridge passive journaling into actionable, multi-layered insights. It uses a hybrid AI pipeline that focuses on abstractive summarization, subtle emotion classification, and profound behavioral analysis to give users a full picture of their mental and emotional states.

A. Core Analytical Components

1. BART for Abstractive Summarization

The pipeline starts by tokenizing user input using a pre-trained BART (Bidirectional and Auto-Regressive Transformer) model. BART's encoder-decoder architecture is very ideal for abstractive summarization, so it can produce fluent, concise summaries that capture the essence of what the user wrote. This method is consistent with next-generation summarization methods utilizing advanced Transformer models to distill complicated content [14].

2. RoBERTa for Subtle Mood Classification

The abstract produced by BART is then fed into a fine-tuned RoBERTa (Robustly Optimized BERT Pretraining Approach) model for mood classification. RoBERTa was chosen as it has an improved pre-training algorithm, which has proven to perform better on subtle classification problems than other variants of the Transformer [1]. Our model was fine-tuned on a dataset of around 14,000 instances spanning seven emotions (happy, angry, sad, fear, disgust, surprise, and neutral) with a validation accuracy of 91.46%.

3. Generative LLM for Deep Insight Generation

To transcend mere categorization, Pocket Journal incorporates a strong generative LLM (Gemini) to perform on-demand analysis. Relevant entries are structured into a prompt, choreographed through LangChain, and the LLM is told to find behavioral trends and report back in a structured JSON response. This application of an LLM to produce rich, queryable insight from large collections of user-created text is an emerging strategy for data summarization and analysis [13].

B. New Feature: On-Demand Deep Insight Generation
Pocket Journal includes an innovative analytical feature beyond the common, automatic mood classification in most wellness apps:

Feature: In the case where day summary and mood score are not sufficient, the user can choose to ask for a deep, longitudinal analysis of their entries within the selected time range.

Mechanism: Once the user's explicit consent (i.e., selecting a date range and pressing the "Generate Insights" button), the system retrieves the relevant journal entries. They are then formatted into a well-crafted prompt for the Gemini Large Language Model (LLM) API, with LangChain used to coordinate the process. The LLM is instructed to analyze the text for themes, behavioral patterns, and progress toward goals, submitting its findings in a structured JSON structure. The JSON completes a detailed report on appreciations, internal conflict, and practical suggestions.

Purpose: This aspect creates a strong vehicle for profound self-examination, converting the journal from a basic mood monitor into an active self-improvement aid. By displaying profound, individual feedback in an organized manner, it enables people to quickly isolate and take action on observations from their own work, actively facilitating their path to self-development.

C. Pocket Journal System Diagrams

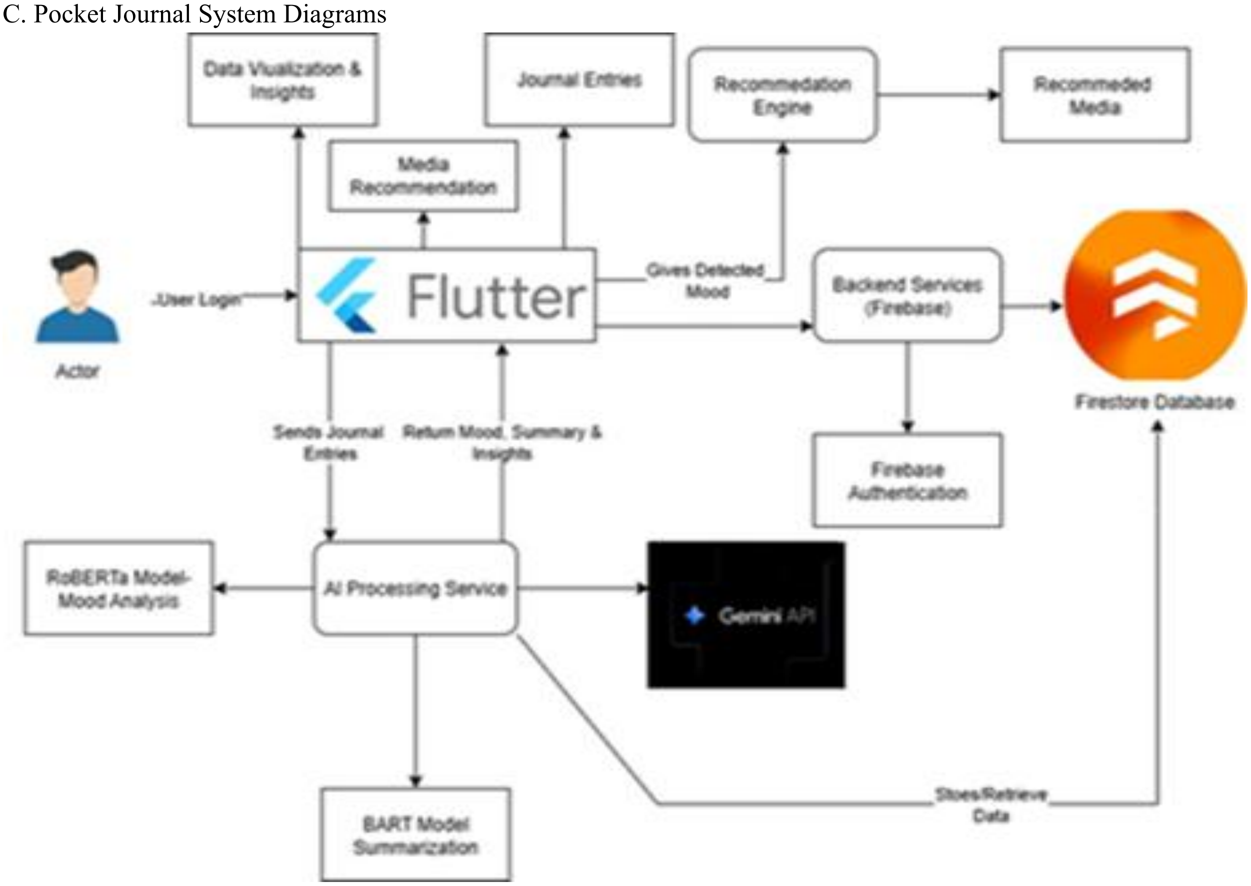

Figure 1 System Architecture

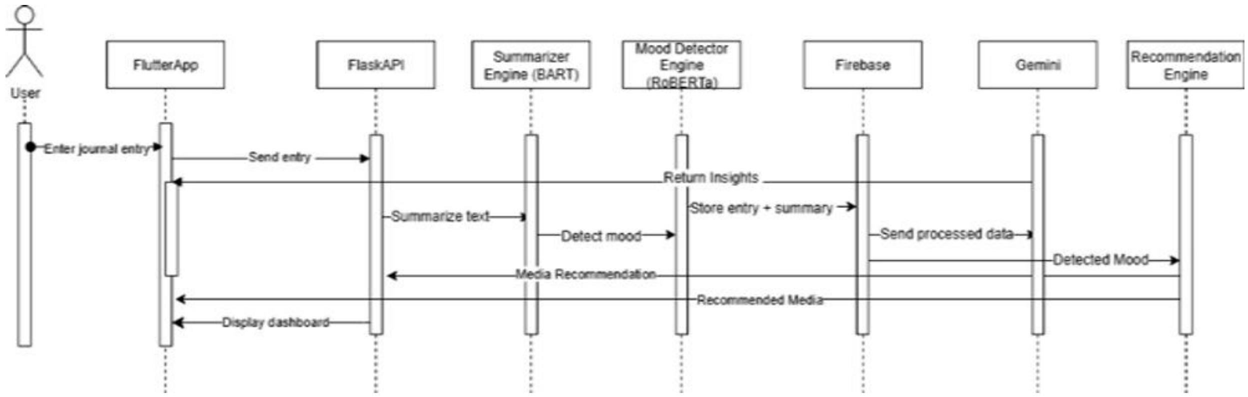

Figure 2 Sequence Diagram

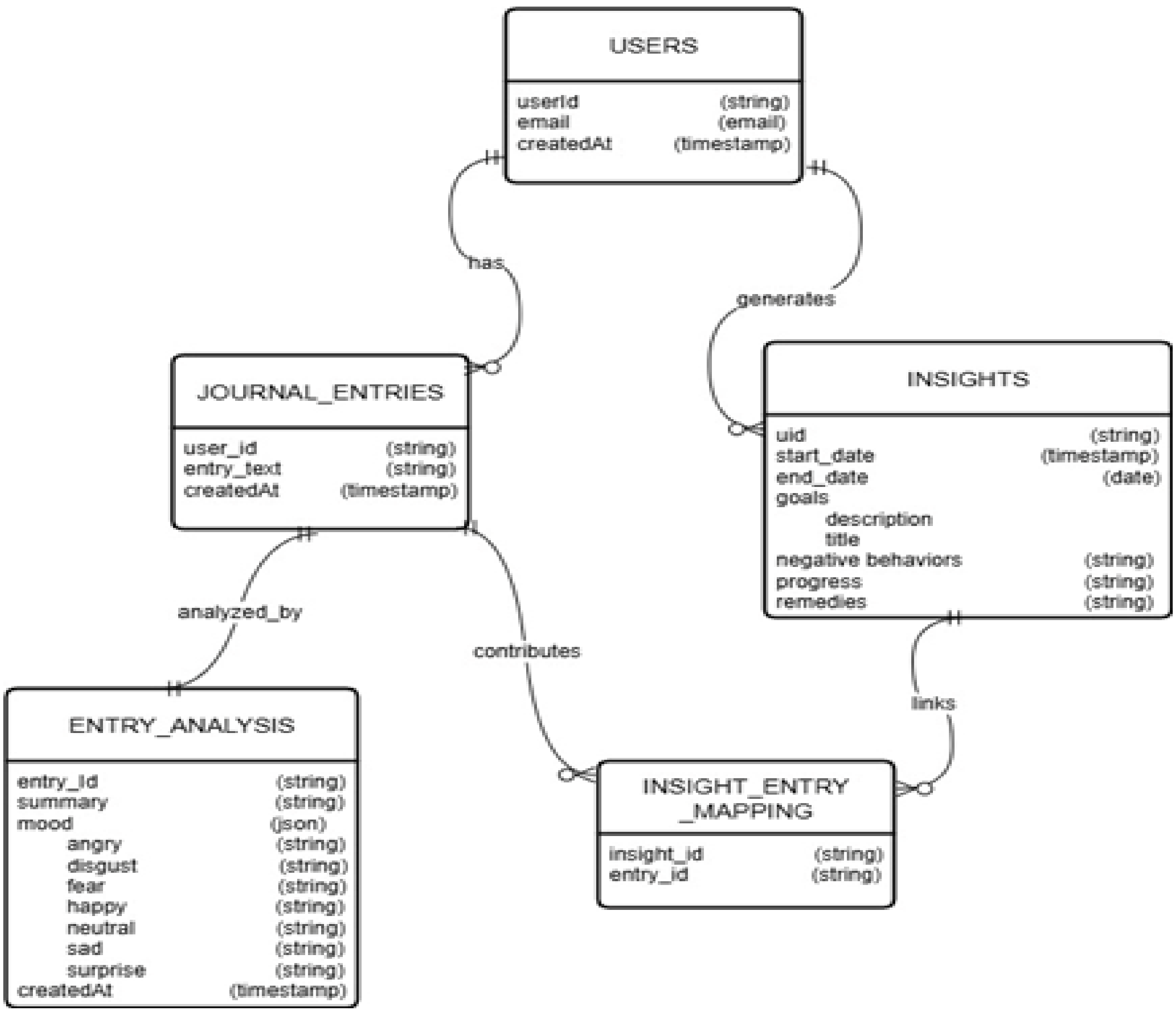

Fig 3 Data Model

Figure 4 UML Diagram

V. DISCUSSION: HOW POCKET JOURNAL OVERCOMES EXISTING LIMITATIONS

The Pocket Journal framework is designed specifically to address the primary shortfalls observed in both standard journaling programs and the present emotion-aware systems:

1. Transforming Passive Journaling into an Active Dialogue

Pocket Journal defies the passive nature of conventional journaling by creating an interactive feedback loop [23], [24]. Instead of passively filing text, the multi-step AI pipeline of the system actively engages with users' content. Having utilized a BART model for abstractive summarization, the process summarizes long, unstructured entries into concise, actionable points to enable users to quickly grasp the essence of what they are thinking and identify key themes [14].

2. Enhancing Accuracy and Sensitivity in Emotion Recognition

In order to address the performance and generalizability problems faced by a majority of emotion recognition models [4], Pocket Journal leverages a cutting-edge, fine-tuned RoBERTa model. This model was chosen because it has a superior ability to cope with the informal, noisy, and subtle text characteristic of individual journal texts [1], [3]. By refining the model on a huge, domain-specific data set, the system achieves high accuracy and high relevance, overcoming the issue of generalizability that is common in heterogeneous data-trained models [4].

3. From Superficial Advice to Deep Insights

Contrary to the existing recommender systems that have a tendency to make recommendations solely based on a point-in-time snapshot input [16], [17], Pocket Journal does much deeper analysis. A new feature is having a generative LLM (Gemini) on demand for insight. This allows users to carry out longitudinal analysis of their entries and identify repeated patterns of behavior, inner conflict, and improvement over time. This is more than simple suggestions and provides dynamic, data-based self-knowledge, an element lacking in current emotion-aware systems [13], [18].

4. Providing an Integrated and Holistic Wellness Tool

The fundamental value of the framework is its harmonious integration of a number of functionalities into a single, whole-wellness tool. It consolidates an individualized journaling experience, a deep analysis engine (BART, RoBERTa, Gemini), and a customized media suggestion service. It generates a balanced experience that promotes mental well-being from multiple angles: self-insight, emotional awareness, and supportive content watching, addressing the fragmentation that exists in the current situation of digital well-being apps.

VI. FUTURE RESEARCH DIRECTIONS

The Pocket Journal approach keeps several doors ajar for future research and development to broaden its scope, capabilities, and interactivity.

1. Extension to Multimodal and Multilingual Analysis

Future work will focus on combining voice and multimedia blogging features, in which the user may leave audio comments or add images. This would entail combining speech emotion recognition models and visual sentiment analysis models to towards a complete multimodal analytical model [19], [21]. To make the app accessible to the global public, we will also extend the NLP capability to process many regional and global languages, an everyday issue in the region [23].

2. Integration of Next-Gen AI to Enhance Longitudinal Insights

The short-term goal is to introduce more advanced large language models to provide richer, more personalized feedback [13]. Long-term studies will include algorithms for longitudinal analysis, so the system can track emotional and behavioral trends over months or years. This would transform the application from a short-term analysis tool to a long-term personal self-improvement assistant that can identify subtle, long-term trends that are not apparent in short-term data [7].

3. Connecting Self-Help with Professional Support Systems

To provide a more supportive system, future releases could include an optional and secure element that allows users to connect with certified mental health professionals. This would connect self-help technology with professional services so users can share their insights from AI with a trusted professional in order to enable more educated and effective sessions of therapy.

VII. CONCLUSION

The necessity for intelligent tools that turn self-reflection into an active process of improvement of mental wellness defines the future generation of digital diaries, but today's systems are plagued by significant gaps, such as the poor generalizability of emotional detection models [4], struggling to handle informal text [3], and the superficial, non-longitudinal character of current recommender systems [16], [17].

The Pocket Journal approach solves these challenges via a multi-stage pipeline of AI with RoBERTa for rich, high-accuracy mood classification [1] and BART for abstractive summarization [14]; a generative LLM for on-demand, rich behavioral insights that go beyond mere recommendations [13]; and an integrated, holistic architecture blending deep analysis with personalized, supportive content. Through the synergy of these cutting-edge NLP models, Pocket Journal obtains a balance of analytical depth, personalization, and user interaction and creates a strong, user-focused basis for future intelligent self-reflection and digital mental well-being tools.

REFERENCE

[1] A. Ahmed, Barkha, T. Kanjwani, and S. Nasim, "Comparative Analysis of DistilBERT, XLNET, ROBERTa & BigBird for Emotion Detection in Conversational Text," ILMA Journal of Technology & Software Management, vol. 6, no. 1, 2025.

[2] I. Ameer, N. Bölücü, G. Sidorov, and B. Can, "Emotion Classification in Texts Over Graph Neural Networks: Semantic Representation is Better Than Syntactic," IEEE Access, vol. 11, pp. 56921-56934, 2023.

[3] F. Anzum and M. L. Gavrilova, "Emotion Detection From Micro-Blogs Using Novel Input Representation," IEEE Access, vol. 11, pp. 19512-19522, 2023.

[4] A. de León Languré and M. Zareei, "Evaluating the Effect of Emotion Models on the Generalizability of Text Emotion Detection Systems," IEEE Access, vol. 12, pp. 70489–70500, 2024.

[5] G. Singh, D. Singh, R. Sharma, and K. Bhardwaj, "Potential Approach for Text-Based Emotion Detection Using NLP Coupled With Deep Learning of Sentiment Analysis," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2024.

[6] K. Machová, M. Szabóová, J. Paralič, and J. Mičko, "Detection of emotion by text analysis using machine learning," Frontiers in Psychology, vol. 14, 2023.

[7] R. Goyal, N. Chaudhry, and M. Singh, "Personalized Emotion Detection from Text using Machine Learning," in 2022 3rd International Conference on Computing, Analytics and Networks (ICAN), 2022.

[8] A. W. J, A. Julian, S. K, and K. R, "Applying Artificial Intelligence for Emotion Detection from Text Messages," in 2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), 2024.

[9] S. L. Jemina and P. PRV, "Decoding Feelings: A Novel Approach To Emotion Recognition From

Textual Data," IOSR Journal of Computer Engineering (IOSR-JCE), vol. 27, no. 2, pp. 58-64, 2025.

[10] S. Yang, "Emotion Detection and Analysis Techniques Based on NLP," in Proceedings of the 3rd International Conference on Mechatronics and Smart Systems, 2025.

[11] R. Padate, A. Gupta, P. Chakrabarti, and A. Sharma, "Emotion Recognition from WhatsApp Text Messages Using Unsupervised Machine Learning," in 2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2024.

[12] A. Al Maruf, Z. M. Ziyad, M. M. Haque, and F. Khanam, "Emotion Detection from Text and Sentiment Analysis of Ukraine-Russia War using Machine Learning Technique," International Journal of Advanced Computer Science and Applications, vol. 13, no. 12, 2022.

[13] B. Praneeth, et al., "Optimization of Customer Feedback Summarization Using Large Language Models (LLM) and Advanced Retrieval-Augmented Generation," IEEE Access, vol. 13, pp. 124319–124332, 2025.

[14] B. Khan, M. Usman, I. Khan, J. Khan, D. Hussain, and Y. H. Gu, "Next-Generation Text Summarization: A T5-LSTM FusionNet Hybrid Approach for Psychological Data," IEEE Access, vol. 13, pp. 37557-37571, 2025.

[15] N. Tennakoon, O. Senaweera, and H. A. S. G. Dharmarathne, "Emotion-Based Movie Recommendation System," International Journal on Advances in ICT for Emerging Regions, vol. 17, no. 1, 2024.

[16] N. Krishnakumar, M. K. K, and S. E. T, "MOOD-BASED RECOMMENDER: Music and Book Recommendations," International Journal of Novel Research and Development, vol. 9, no. 5, May 2024.

[17] S. Unnisa and A. N. S, "Personalized Mood-Centric Book Recommendation Integrating Machine Learning with Content-Based Filtering," International Journal of Information Technology, Research and Applications, vol. 3, no. 3, pp. 15-22, 2024.

[18] J. K. Leung, I. Griva, and W. G. Kennedy, "TEXT-BASED EMOTION AWARE RECOMMENDER," in CS & IT - CSCP, 2020, pp. 101-114.

[19] S. Challa, N. P., G. M. V., D. D. C. C.A, and R. M.S, "Content recommendation based on recognised Emotion," in 2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), 2023.

[20] H. Gunjal, T. Kharde, A. Nair, T. Sase, and Y. Sisodia, "EMOTION-BASED MUSIC AND MOVIE RECOMMENDATION SYSTEM USING PYTHON," International Journal of Current Science, vol. 13, no. 2, April 2023.

[21] T.-Y. Kim, H. Ko, S.-H. Kim, and H.-D. Kim, "Modeling of Recommendation System Based on Emotional Information and Collaborative Filtering," Sensors, vol. 21, no. 6, p. 1997, 2021.

[22] A. Karve, A. Humnabadkar, B. Shivbhakta, and S. Naik, "Machine Learning based Mood Prediction and Recommendation System," International Journal of Scientific Research in Engineering and Management (IJSREM), vol. 8, no. 9, Sept. 2024.

[23] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," Social Network Analysis and Mining, vol. 11, no. 1, 2021.

[24] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Varadarajan, "AI Based Emotion Detection for Textual Big Data: Techniques and Contribution," Big Data and Cognitive Computing, vol. 5, no. 3, p. 43, 2021.