AI for Mental Wellness: A Supportive Tool for Emotional Analysis and Personalized Interactions

Priyansh Rajusth
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
https://orcid.org/0009-0003-9877-5006

Nandhini Ramesh
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
https://orcid.org/0009-0006-8459-3553

Harsh Prabhu
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
harsh.prabhu@somaiya.edu

Shubh Radia
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
shubh.radia@somaiya.edu

Tejas Pundlik
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
https://orcid.org/0009-0009-3764-3425

Nirmala Baloorkar
Computer Engineering
K.J. Somaiya College of Engineering
Mumbai, India
https://orcid.org/0000-0001-5277-7936

***Abstract*—Access to mental health support remains a global challenge due to stigma, cost, and limited availability of resources. This paper presents an AI-based mental wellness tool that provides real-time emotional support and personalized insights. The system combines a conversational AI module with advanced facial emotion detection, enabling adaptive interactions based on users' emotional states. Psychological assessments are recommended using the Apriori algorithm, ensuring relevance through data-driven insights derived from user interaction history. Deployed as a secure web application, this tool empowers users to explore mental health resources, monitor emotional well-being, and engage in meaningful, supportive conversations. By leveraging artificial intelligence, the system aims to complement traditional mental health care approaches, improving accessibility and personalization.**

Keywords—AI mental health, emotion detection, psychological assessments, conversational AI, Apriori algorithm, mental wellness, accessible mental health.

I. Introduction

Mental health issues are a major concern globally, and India is no exception. As per the National Mental Health Survey of India, 2015-16, carried out by the National Institute of Mental Health and Neuro Sciences (NIMHANS), 10.6% of Indian adults experienced mental disorders. The treatment gap for these conditions varied from 70% to 92% across different disorders. Furthermore, the occurrence of mental disorders was greater in urban metro regions (13.5%) than in rural areas (6.9%) and urban non-metro regions (4.3%). [1] Additionally, the Mental Health and Well-being of School Students: A Survey, 2022, carried out by the National Council of Educational Research and Training (NCERT), emphasized a rising occurrence of poor mental health in adolescents, worsened by the COVID-19 pandemic. The survey indicated that 11% of students reported feeling anxious, 14% went through intense emotions, and 43% experienced mood fluctuations. [2]

While traditional therapy remains vital, there is a growing demand for accessible, scalable, and personalized mental health tools to bridge these gaps. Recent advancements in artificial intelligence (AI), particularly large language models (LLMs), have shown promise in this domain. Studies indicate that LLMs can effectively detect mental health conditions and provide accessible, destigmatized eHealth services. [3]

Nonetheless, it is essential to verify that mental health LLMs are specifically fine-tuned for mental health uses, promote mental health equity, and comply with ethical guidelines.

Motivated by these developments, we have developed an online platform that leverages LLM-based therapeutic sessions as a tool for accessible mental health support. Our system incorporates an emotion recognition model to enhance context awareness by assessing users' emotional states. Additionally, the platform offers open-source psychological tests, with an Apriori algorithm-based recommendation system to suggest relevant assessments. The results, along with their timestamps, are integrated into the model to tailor responses more effectively. This multifaceted approach aims to provide a comprehensive, user-centric mental health support system that complements traditional therapeutic methods.

II. Literature Survey

Recent advances in Large Language Models (LLMs) have demonstrated significant potential in revolutionizing mental health support and therapeutic interventions. Cho et al.'s [4] evaluation of LLMs for therapeutic applications with high-functioning autistic adolescents revealed promising capabilities in maintaining coherent conversations and building initial rapport, with the system successfully adapting to changing conversation dynamics and maintaining age-appropriate language. However, their study identified critical limitations in personalization capabilities and inconsistent empathy in complex emotional scenarios, emphasizing the need for more sophisticated emotional cue detection. De Choudhury et al. [5] conducted a comprehensive analysis of LLMs in digital mental health, highlighting their potential to provide 24/7 support through AI-powered chatbots, particularly beneficial for underserved areas. Their research emphasized LLMs' capability to analyze patient data for tailored treatment plans and support clinical decision-making with evidence-based recommendations, while raising concerns about data privacy, ethical implications, and the potential for misinformation in mental health advice.

Several innovative systems have emerged, each addressing specific aspects of mental health support. Nie et al.'s [6] CaiTI system represents a significant advancement, implementing a comprehensive approach that evaluates users across 37 dimensions of daily functioning. The system uniquely incorporates both Cognitive Behavioral Therapy (CBT) and Motivational Interviewing (MI) techniques, using reinforcement learning to prioritize topics based on user history and interaction context. Their 24-week longitudinal study demonstrated significant improvements in user engagement and mental well-being scores, though highlighting challenges with LLM consistency in certain tasks. Tin-Lai et al.'s [7] Psy-LLM framework, built on Chinese LLMs like PanGu and WenZhong, introduced a novel approach to scaling mental health services. Their system, trained on over 22,000 Q&A pairs and extensive psychological articles, demonstrated professional-quality responses and effective screening capabilities for urgent mental health cases. However, the study identified limitations in processing bidirectional context and interpreting non-verbal cues, crucial components in psychological counseling.

Recent evaluations have provided detailed insights into the practical implementation and effectiveness of these systems. Lamichhane et al.'s [8] assessment of ChatGPT focused on zero-shot classification of mental health states from social media text, utilizing three distinct datasets for stress, depression, and suicidality detection. While their study demonstrated ChatGPT's flexibility across multiple datasets, it highlighted the need for more sophisticated prompting strategies and larger, more rigorously annotated datasets. Liu et al.'s [9] ChatCounselor introduced a significant advancement through their Psych8k dataset, comprising real counseling conversations between licensed psychologists and clients. Their evaluation framework, utilizing seven metrics assessed via GPT-4, demonstrated performance approaching ChatGPT's capabilities in counseling-specific tasks. Bolpagni et al.'s [10] research on AI-powered digital therapeutics introduced a novel approach integrating wearable devices for real-time stress monitoring and feedback. Their system emphasized the importance of transparent feedback mechanisms and blended approaches combining human facilitation with self-guided digital tools, though noting significant challenges in accessibility due to the high cost of wearable technology and privacy concerns in data processing and storage.

The development trajectory of these systems reveals a consistent pattern of advancement in therapeutic capabilities while highlighting persistent challenges in privacy protection, ethical implementation, and the need for human oversight. Recent studies increasingly emphasize the importance of integrating these systems within existing healthcare frameworks rather than viewing them as standalone solutions, suggesting a future where AI-based therapeutic tools complement rather than replace traditional mental health services.

III. Methodology

A. Dataset Acquisition and Preprocessing

The FER2013 (Face Expression Recognition 2013) dataset is a well-known collection for recognizing facial expressions. It features grayscale images of human faces, each identified by one of seven fundamental facial expressions. The FER+ dataset expands on this by categorizing the original images into eight groups: happiness, surprise, neutral, sadness, anger, fear, disgust, and contempt (with contempt as the new category). The FER+ dataset offers a broader collection of emotion labels, enhancing the detail of emotion recognition activities [11].

B. Training the Convolutional Neural Network

1) Model

The proposed model, named DeeperEmotionCNN, is a convolutional neural network (CNN) specifically designed for emotion recognition from grayscale facial images. The network follows a deep architecture with the following layers:

a) Convolutional Layers: Nine convolutional layers are employed to extract hierarchical features. These layers are interspersed with batch normalization for stable training and ReLU activations for non-linearity. The first two layers extract low-level features (32 and 64 filters). The subsequent blocks increase the depth with 128 and 256 filters for mid- and high-level feature representation. The model also includes dilated convolutions and padding to capture global and local context.

b) Pooling and Dropout: Four max-pooling layers are used to downsample the spatial dimensions progressively. Dropout layers (with rates of 0.25 and 0.5) are included after convolutional blocks to mitigate overfitting.

c) Fully Connected Layers: The flattened feature map is passed through two fully connected layers. The first has 256 neurons and uses dropout (0.5), followed by batch normalization. The final layer consists of 8 neurons corresponding to the emotion classes (happiness, surprise, neutral, sadness, anger, fear, disgust, and contempt).

d) Output Layer: A softmax activation (via cross-entropy loss) is used to output the probability distribution over emotion classes.

2) Hyperparameters

a) Input Shape: 48×48 (grayscale).

b) Batch Size: 64.

c) Learning Rate: 0.001.

d) Epochs: 50.

e) Optimizer: The model is trained using the AdamW optimizer, which combines Adam's adaptive learning rate capability with weight decay for better regularization.

3) Preprocessing: Training images undergo random rotations (±10°) and horizontal flips for augmentation. Both training and testing datasets are normalized to a range of [-1, 1].

4) Loss Function: The CrossEntropyLoss function is used to compute the classification error.

5) Early Stopping: Training includes an early stopping mechanism with a patience value of 5 to prevent overfitting. The model achieving the best validation accuracy is saved for further use.

6) Performance Metrics and Visualization

During training, accuracies for both training and validation are monitored across epochs. Initially, the model is set to be trained for 50 epochs but is early-stopped by epoch 47. The final chosen model has a training accuracy of 88.69% and a validation accuracy of 81.86%.

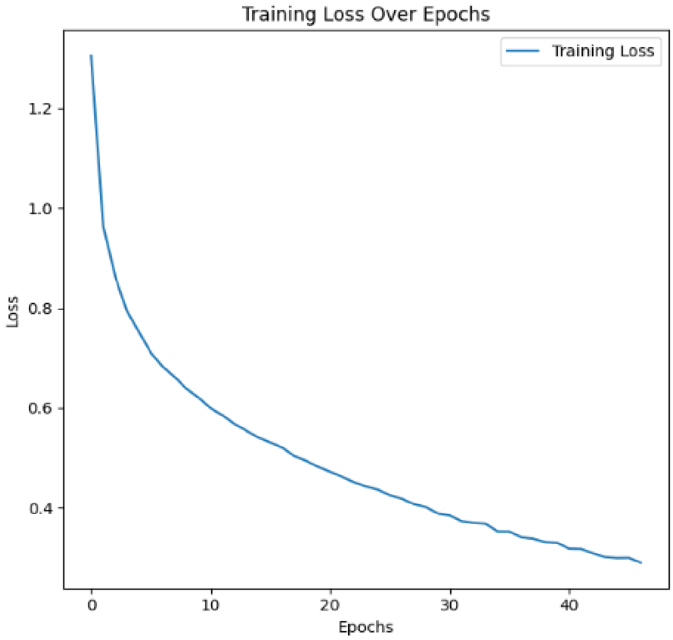

Fig 1: Training Loss Over Epochs

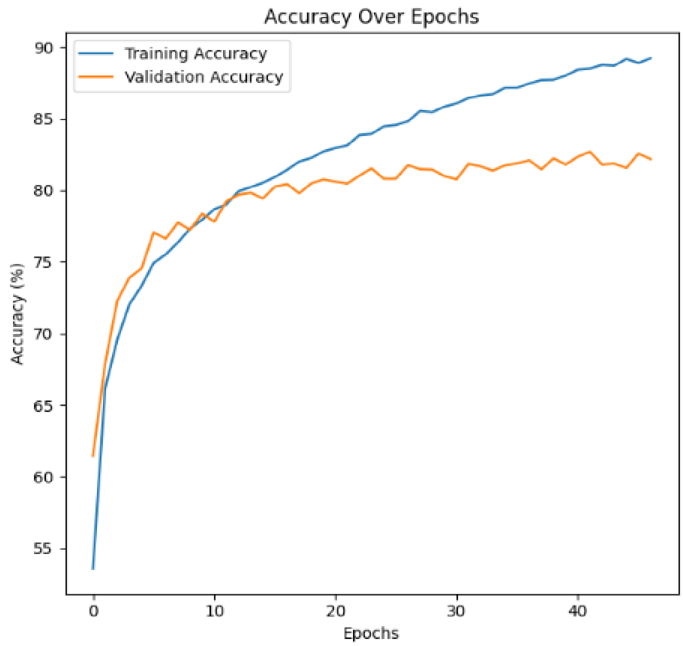

Fig 2: Accuracy Over Epochs

C. *Architecture*

The website for the AI therapist tool was developed with a clear distinction between the frontend and backend components to ensure seamless functionality and a user-friendly interface. The backend was constructed using Flask API, providing a robust and lightweight framework for handling server-side logic. For data management, SQL was employed, offering a reliable structure for storing and retrieving models and user data efficiently. The database architecture, represented in Figure 3, showcases the relational design that supports the application's requirements, ensuring scalability and easy management of psychometric tests, user interactions, and model outputs.

To provide a comprehensive view of how various components interact, the overall application architecture is depicted in Figure 4. This architecture outlines the flow of data between the front end, backend, database, and the models (Custom Trained Emotion Recognition Model, Apriori algorithm for frequent itemset mining based Recommendation System with 50 percent minimum confidence threshold, LLM chatbot) It demonstrates how the tool integrates user inputs, processes them through the trained facial emotion recognition model, and delivers personalized insights. Together, these architectural designs highlight the technical sophistication of the application, ensuring both functionality and ease of use.

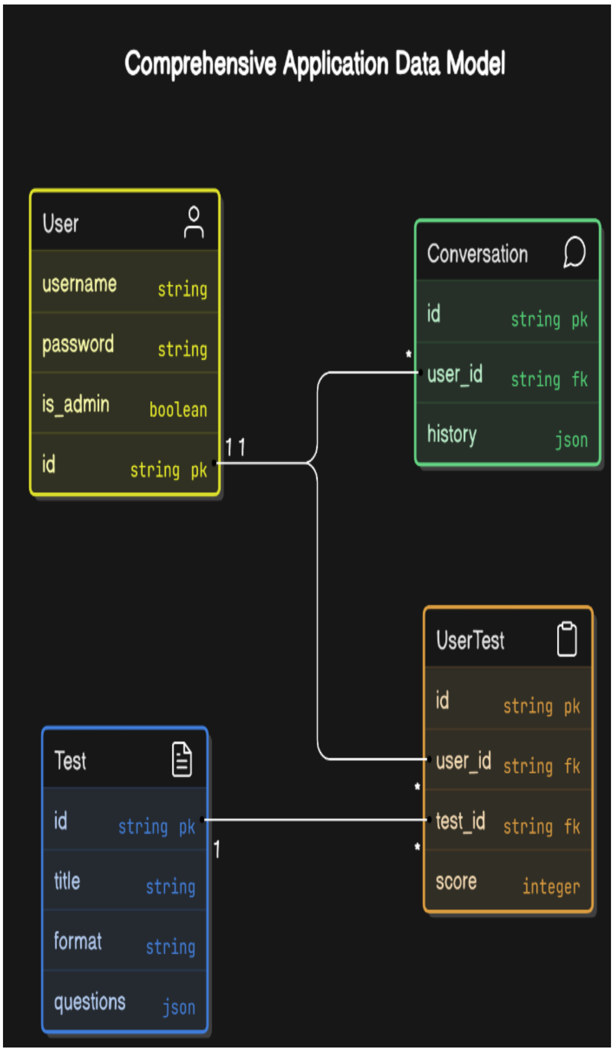

Fig 3. Database Schema

D. *Psychological Tests Integration*

The integration of psychological tests into the tool involved acquiring standardized assessments from trusted resources like Psychology Tools. [12] The selection process was carefully curated to align with the tool's primary objectives, such as assessing anxiety, depression, and other emotional well-being parameters. A significant emphasis was

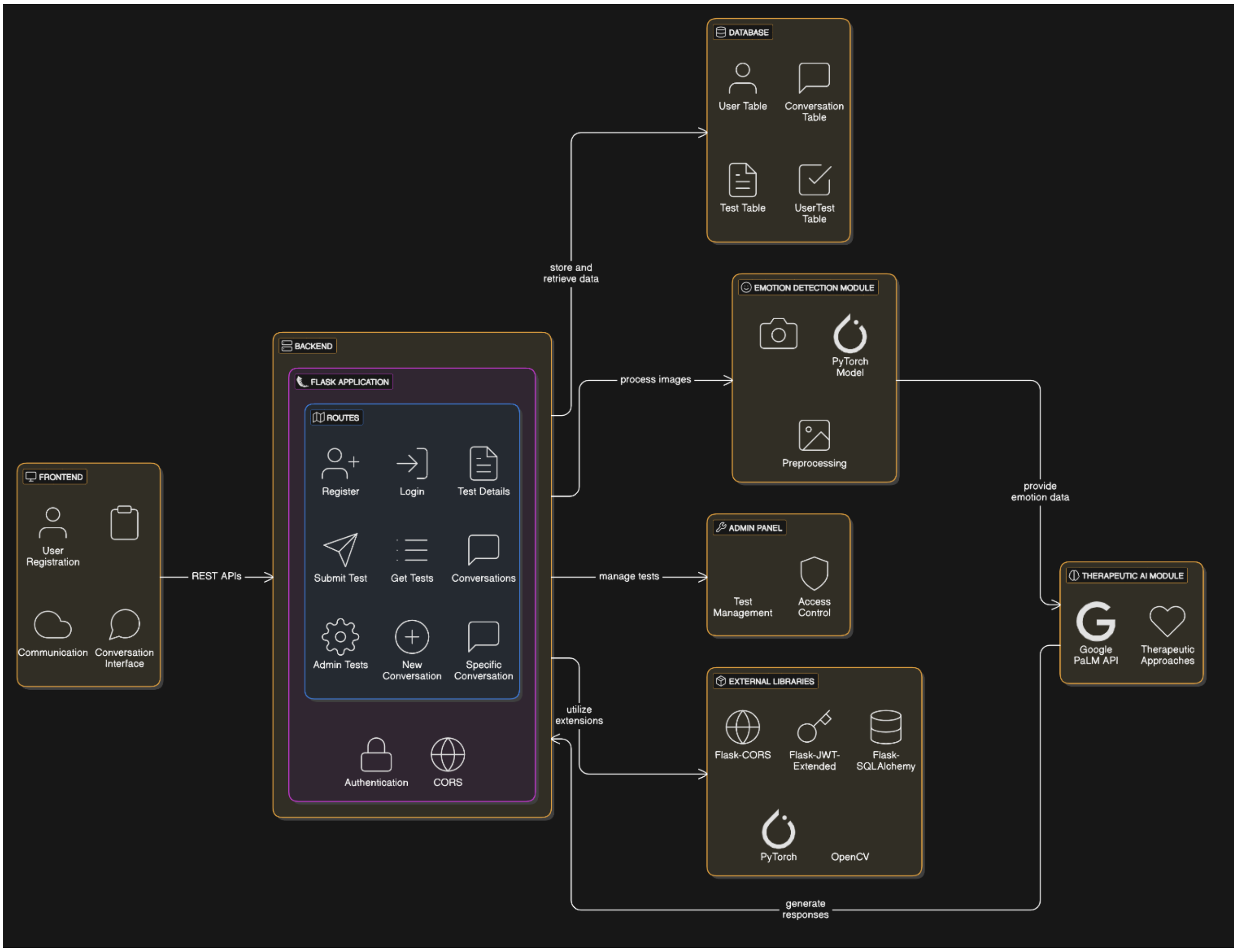

Fig 4. Architecture of the application

placed on ensuring that the interface for completing and scoring these tests is intuitive and accessible to users, allowing seamless navigation and understanding of their emotional health. The platform is designed to accommodate any number of psychometric tests in MCQ format, offering scalability and adaptability for future expansions. Currently, the tool includes a diverse range of assessments: the Generalized Anxiety Disorder (GAD-7) [13] scale, the Patient Health Questionnaire (PHQ-9) [14], the Perceived Stress Scale (PSS), the Borderline Symptom List 23 [15], the Severity Measure for Agoraphobia—Adult [16], the Chronic Pain Acceptance Questionnaire – Revised (CPAQ-R) [17], the Camouflaging Autistic Traits Questionnaire (CAT-Q) [18], the Leeds Dependence Questionnaire for drug addiction [19], the Brief Fear of Negative Evaluation Scale [20], the CUDOS Scale for depression [21], the Interpersonal Needs Questionnaire (INQ) for suicidal ideation [22], and the Buss & Perry Aggression Questionnaire [23]. These tests collectively enhance the tool's ability to deliver a comprehensive and personalized evaluation of a user's psychological state.

E. LLM and Prompt Engineering

In developing our therapeutic application, we have integrated a Large Language Model to serve as the virtual therapist, with a focus on delivering text-based interactions. While several LLMs are capable of performing this role effectively, we have selected Google's Gemini model due to its specific advantages in handling extensive textual data and maintaining coherent, contextually relevant conversations. One of the primary reasons for choosing Gemini is its exceptionally long context window. [24] The Gemini 1.5 Pro model supports a context window of up to 2 million tokens, allowing it to process and retain large amounts of conversational history. This capability is crucial in therapeutic settings, where understanding the user's previous inputs and maintaining continuity across sessions enhances the quality of interactions.

Additionally, Gemini's architecture is designed to handle complex language tasks with high efficiency. Its ability to process and generate human-like text makes it well-suited for applications requiring nuanced understanding and response generation, such as virtual therapy. [25] The model's design facilitates the delivery of empathetic and contextually appropriate responses, which are essential in therapeutic communications.

Furthermore, Gemini's integration capabilities with various platforms and its support for multimodal inputs provide flexibility for future expansions of our application.

[26] While our current implementation focuses on text-based interactions, Gemini's design allows for potential incorporation of other data forms, enhancing the scope and effectiveness of the therapeutic interventions we can offer.

F. *Website*

The website consists of 3 main pages. 'Tests' pages consists of all the tests a person can take along with recommendations based on the Apriori algorithm. 'My Tests' displays all the tests previously taken by the user along with the marks. 'Conversations' page consists of all the conversations the user has, and upon clicking any one of them, it displays the chats between the user and the AI Therapist LLM. The LLM also explains the psychology-based techniques (such as Cognitive Behaviour Therapy, Acceptance and Commitment Theory, etc.) used to generate responses. Figures 5-8 consist of the pages as implemented in the website.

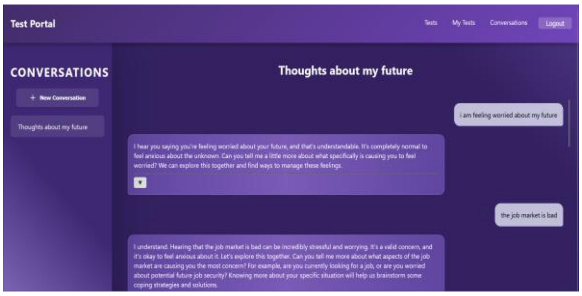

Fig 5. UI of the Conversations Page

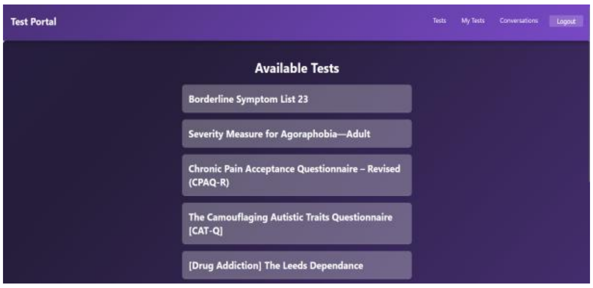

Fig 6. UI of the Tests Page

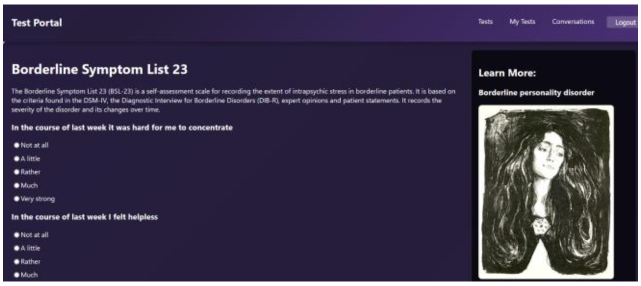

Fig 7. Opening a particular test to attempt it

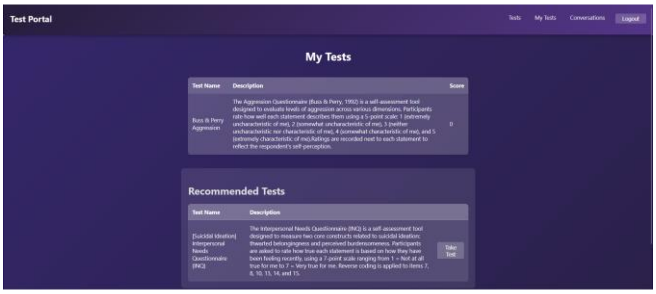

Fig 8. User's Test Page with recommendation

IV. Conclusion

The development and implementation of our AI-based mental wellness tool represents a significant step toward making mental health support more accessible and personalized. By combining emotion recognition capabilities, psychological assessments, and LLM-based therapeutic interactions, our system provides a comprehensive approach to mental health support. The integration of the Apriori algorithm for test recommendations ensures that users receive relevant assessments based on their interaction patterns and needs. The emotion recognition model, achieving a validation accuracy of 81.86%, demonstrates the system's capability to understand and respond to users' emotional states effectively. While this tool is not intended to replace traditional mental health care, it serves as a valuable complement to existing services, particularly in addressing the substantial treatment gap identified in regions like India, where treatment gaps range from 70% to 92% for different conditions. The system's ability to provide 24/7 support, maintain conversation context through Gemini's extensive context window, and offer standardized psychological assessments makes it a practical solution for initial mental health support and ongoing emotional well-being monitoring. The incorporation of diverse psychological tests comprehensive assessment capabilities across various mental health dimensions. Future work could focus on expanding the system's capabilities through additional assessment tools, enhanced emotion recognition accuracy, and deeper integration with traditional mental health services to create a more robust and comprehensive mental health support ecosystem. Additionally, continued research into the effectiveness of AI-powered therapeutic interventions and their impact on mental health outcomes will be crucial for refining and improving such systems, ultimately working toward the goal of making quality mental health support accessible to all who need it.

V. References

[1] National Institute of Mental Health and Neuro Sciences, "National Mental Health Survey of India, 2015-16: Summary," Bengaluru, 2016. https://mohfw.gov.in/sites/default/files/National%20Mental%20Health%20Survey%2C%202015-16%20-%20Summary%20Report_0.pdf (accessed Jan. 21, 2025).

[2] National Council of Educational Research and Training, "Mental Health and Well-being of School Students: A Survey, 2022," New Delhi, 2022. https://ncert.nic.in/pdf/Mental_Health_WSS_A_Survey_new.pdf (accessed Jan. 21, 2025).

[3] Lawrence, H. R., Schneider, R. A., Rubin, S. B., Matarić, M. J., McDuff, D. J., & Jones Bell, M. (2024). The Opportunities and Risks of Large Language Models in Mental Health. JMIR mental health, 11, e59479. https://doi.org/10.2196/59479

[4] Cho, Yujin, Mingeon Kim, Seojin Kim, Oyun Kwon, Ryan Donghan Kwon, Yoonha Lee, and Dohyun Lim. "Evaluating the efficacy of interactive language therapy based on LLM for high-functioning autistic adolescent psychological counseling." *arXiv preprint arXiv:2311.09243* (2023).

[5] De Choudhury, Munmun, Sachin R. Pendse, and Neha Kumar. "Benefits and harms of large language models in digital mental health." *arXiv preprint arXiv:2311.14693* (2023).

[6] Nie, Jingping, Hanya Shao, Yuang Fan, Qijia Shao, Haoxuan You, Matthias Preindl, and Xiaofan Jiang. "LLM-based Conversational AI Therapist for Daily Functioning Screening and Psychotherapeutic Intervention via Everyday Smart Devices." *arXiv preprint arXiv:2403.10779* (2024).

[7] Lai, T., Y. Shi, Z. Du, J. Wu, K. Fu, Y. Dou, and Z. Wang. "Psy-LLM: scaling up global mental health psychological services with AI-based large language models. arXiv." *arXiv preprint arXiv:2307.11991* (2023).

[8] Lamichhane, Bishal. "Evaluation of chatgpt for nlp-based mental health applications." *arXiv preprint arXiv:2303.15727* (2023).

[9] Liu, June M., Donghao Li, He Cao, Tianhe Ren, Zeyi Liao, and Jiamin Wu. "Chatcounselor: A large language models for mental health support." *arXiv preprint arXiv:2309.15461* (2023).

[10] Bolpagni, Marco, Susanna Pardini, and Silvia Gabrielli. "Human centered design of AI-powered Digital Therapeutics for stress prevention: Perspectives from multi-stakeholders' workshops about the SHIVA solution." *Internet Interventions* 38 (2024): 100775.

[11] Microsoft, "FERPlus Dataset," GitHub Repository. [Online]. Available: https://github.com/microsoft/FERPlus. (accessed Jan. 21, 2025).

[12] Psychology Tools, "Download Scales and Measures," [Online]. Available: https://www.psychologytools.com/download-scales-and-measures. (accessed Jan. 21, 2025).

[13] Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of internal medicine, 166(10), 1092-1097.

[14] Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ‐9: validity of a brief depression severity measure. Journal of general internal medicine, 16(9), 606-613.

[15] Bohus, M., Limberger, M. F., Frank, U., Chapman, A. L., Kühler, T., & Stieglitz, R.-D. (2007). Psychometric properties of the Borderline Symptom List (BSL). Psychopathology, 40(2), 126–132. https://doi.org/10.1159/000098493

[16] Craske, M., Wittchen, U., Bogels, S., Stein, M., Andrews, G., & Lebeu, R. (2013). Severity Measure for Agoraphobia–Adult [Measurement instrument].

[17] McCraken, L. M., Vowles, K. E. & Eccleston, C. (2004). Acceptance of chronic pain: component analysis and a revised assessment method. Pain, 107, 159-166.

[18] Hull, L., & Mandy, W. (2021). Camouflaging Autistic Traits Questionnaire (CAT-Q). Encyclopedia of Autism Spectrum Disorders, 795-797.

[19] Raistrick, D., Bradshaw, J., Tober, G., Weiner, J., Allison, J., & Healey, C. (1994). Development of the Leeds Dependence Questionnaire (LDQ): a questionnaire to measure alcohol and opiate dependence in the context of a treatment evaluation package. Addiction, 89(5), 563-572.

[20] Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. Personality and social psychology bulletin, 9(3), 371-375.

[21] Zimmerman, M., Chelminski, I., McGlinchey, J. B., & Posternak, M. A. (2008). A clinically useful depression outcome scale. Comprehensive psychiatry, 49(2), 131-140.

[22] Van Orden, K. A., Cukrowicz, K. C., Witte, T. K., & Joiner Jr, T. E. (2012). Thwarted belongingness and perceived burdensomeness: construct validity and psychometric properties of the Interpersonal Needs Questionnaire. Psychological assessment, 24(1), 197.

[23] Buss, A. H., & Perry, M. (1992). The aggression questionnaire. Journal of personality and social psychology, 63(3), 452.

[24] Google. "New Features for the Gemini API and Google AI Studio." [Online]. Available: https://developers.googleblog.com/en/new-features-for-the-gemini-api-and-google-ai-studio/ (accessed Jan. 21, 2025).

[25] AI Scaleup. "Google Gemini AI." [Online]. Available: https://www.ai-scaleup.com/articles/ai-tools/google-gemini-ai/ (accessed Jan. 21, 2025).

[26] Zapier. "Google Gemini: Everything You Need to Know." [Online]. Available: https://zapier.com/blog/google-gemini/ (accessed Jan. 21, 2025).