Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

STUDENT COMPANION: EARLY INTERVENTION SYSTEM FOR STUDENTS MENTAL HEALTH

[1]Khadija Muskan, [*2] Dr. Mahawish Fatima, [3]Dr.MuhammadAshraf, [1]Fatima Majeed, [1]Asad Ali, [4]Dr. Muhammad Hassan Nasir, [5]Shanila Azhar, [2]Dr. Bushra Fazal

[1]Student, Department of Software Engineering Bahria University Karachi Campus, Pakistan
[*, 2]Assistant Professor, Department of Software Engineering Bahria University Karachi Campus, Pakistan
[3]Associate Professor, Department of Computer Engineering, BUITEMS, Quetta, Pakistan
[4]Department of Computer Science & IT, NED University of Engineering & Technology, Pakistan
[5]Lecturer, Department of Computer Engineering, BUITEMS, Quetta, Pakistan

[1]mallahkhadija97@gmail.com [*2]mahwishfatima.bukc@bahria.edu.pk
[3]Muhammad.ashraf@buitms.edu.pk [1]fatimamajeed854@gmail.com
[1]asadalichksdhr2003@gmail.com [5]shanila.azhar@buitms.edu.pk [2]bushrafazal.bukc@bahria.edu.pk

DOI:-https://doi.org/10.5281/zenodo.20346869

Article History

Received: 17 April 2026

Accepted: 08 May 2026

Published: 24 May 2026

Corresponding Author: *

Dr. Mahawish Fatima

Abstract

Pakistani students' mental health is in crisis. Students are under academic stress, Competition, expectations from family and society, and all this, while going through a cultural image. In countries where mental health care is highly stigmatized. The results are: untreated distress has led to student suicides across provinces. Current support is based on students They don't many of them, but they seek help themselves. Our system recommends an AI based early intervention program, Student Companion, which can detect signs of mental health issues, challenges students face prior to entering crisis without the student being in crisis to make an initial move. It operates on the basis of a two parallel data source, using the multimodal approach. First, students fill in the completed and validated questionnaire. Second, the system analyzes live facial expressions to capture emotional cues that may not be captured in self-report, through their device camera. Combining both inputs creates a more complete picture of the student's psychological state. Finally, a customized report is created and assessed against psychologist criteria to determine the level of severity of the student's condition and to introduce the student to a virtual therapy assistant that can offer unique coping strategies and guidance. Student Companion is not meant to substitute clinical care. Rather, it plugs the gap between quiet misery and professional assistance creating a less intrusive, earlier mental health assistance. Technology is not a replacement to human care in this case, but rather a tool of making sure that more students receive the necessary assistance in a timely fashion before their situation worsens.

Keywords: Mental health, AI, Facial expression recognition, Early intervention, Pakistan, NLP

1. Introduction

Mental health issues have increased in recent years around the world and suicide is a very alarming situation when mental suffering is left untreated. [1] According to global evaluations, more than 300 million individuals currently keep on living their lives with depression. [2] This issue is not unique to underdeveloped countries, but has become alarming in the developing world as well, where the support of mental health is not sufficient and social stigma still prevents listeners from openly discussing with victims. [3]

Students in Pakistan are psychologically vulnerable because of the increasing stress at school, competitive exams, burnout in educational settings, and high expectations from society, during their formative years. [4] This is worsened by limited mental health awareness and under reported cases of suicide. [5] [6] In 2019, a cluster of suicide cases among undergraduate students in Pakistan was reported in an opinion article in a short span of time, which was a worrying trend among the overall student population in Pakistan. [7]

Current studies and other evidence indicate that depression, anxiety, panic and suicide are not uncommon occurrences to be seen in Pakistan. [8] The overall estimates are in excess of 800,000 deaths per year from suicide, and adolescents and young adults represent an unusually high proportion, especially in low and middle income countries such as ours. [9]

There is no centralized system for reporting suicides and depression, which means that the actual rate of depression and suicide is likely to be higher than is currently reported. There are alarming reports and observations of undergraduate and medical students committing suicide, often linked to examination failure and untreated depression, and academic stress. In 2019, an opinion article reported at least five student suicides in various provinces within a period of a few months, and a subsequent content analysis of newspaper reports revealed 289 cases of child and adolescent suicide over two years, with the highest number of cases reported in late adolescence. All these findings point to the fact that student suicide is not an accident, but a mental health problem that is systemic in Pakistan and needs to be addressed. [10]

2. Backgroubd

With growing awareness, a lot of mental health support does exist in Pakistan, but present mental health screening mechanism primarily focuses on students who voluntarily ask for help [11], [12]. This self-initiation approach can sometimes be ineffective because a person who is suffering from anxiety and panic symptoms often feels frightened and confused when they try to recognize their own anxiety and panic or when they go to authorities. [13]

Furthermore, Current support systems approaches are based on clinical consultations, multiple face-to-face visits, completing paperwork, and medication-focused treatment plans [14]. This is required for serious cases, but for students already emotionally stressed these patterns can be overwhelming and discouraging, resulting in a delayed follow up and lack of engagement with health. All these situations make one feel the absence of unobtrusive and student friendly mechanisms for early mental well-being assessment; therefore, represent a critical gap, allowing psychological issues to intensify even before any timely support can be provided.

These gaps can be addressed by a proactive and student-centered approach that can detect early signs of psychological distress before they escalate into serious consequences. Recent developments in AI, especially in NLP and emotion recognition, present us with the opportunity to think and move beyond self initiated help seeking models, to continuous and discreet mental well-being assessment. [15][16]

Intelligent multi modal frameworks can also analyze it, both verbal expressions and non-verbal emotional signs, in addition to clinical visits, to get a more comprehensive picture of a student's mental state [17]. This attitude is especially linked in contexts like Pakistan where there is cultural stigma and fear of disclosure of identity that can make it challenging for students to express their emotions openly.

3. PROPOSED solution

To help students in an early stage, the study therefore proposes a mental wellbeing detection framework based on Artificial Intelligence, which centers on the early detection of mental health symptoms thanks to the fusion of textual emotion analysis and facial emotion analysis. To address these challenges, the purpose of this research is to explore the role of intelligent and technology assisted systems to help early mental well-being assessment of university students in Pakistan. The study aims to look into a precautionary approach to the detection of students' emotions, in addition to the detection of crisis situations, and to develop an approach that considers students' emotional realities, cultural sensitivities, and privacy concerns.

The proposed approach combines multi-modal emotional analysis to determine early psychological signs and symptoms in an non-intrusive way, to support the student prior to the onset of serious mental health issues. In this direction, the study contributes to closing the gap between and existing support structures for mental health and the need for early, accessible and student-friendly mental well-being monitoring in educational environments.

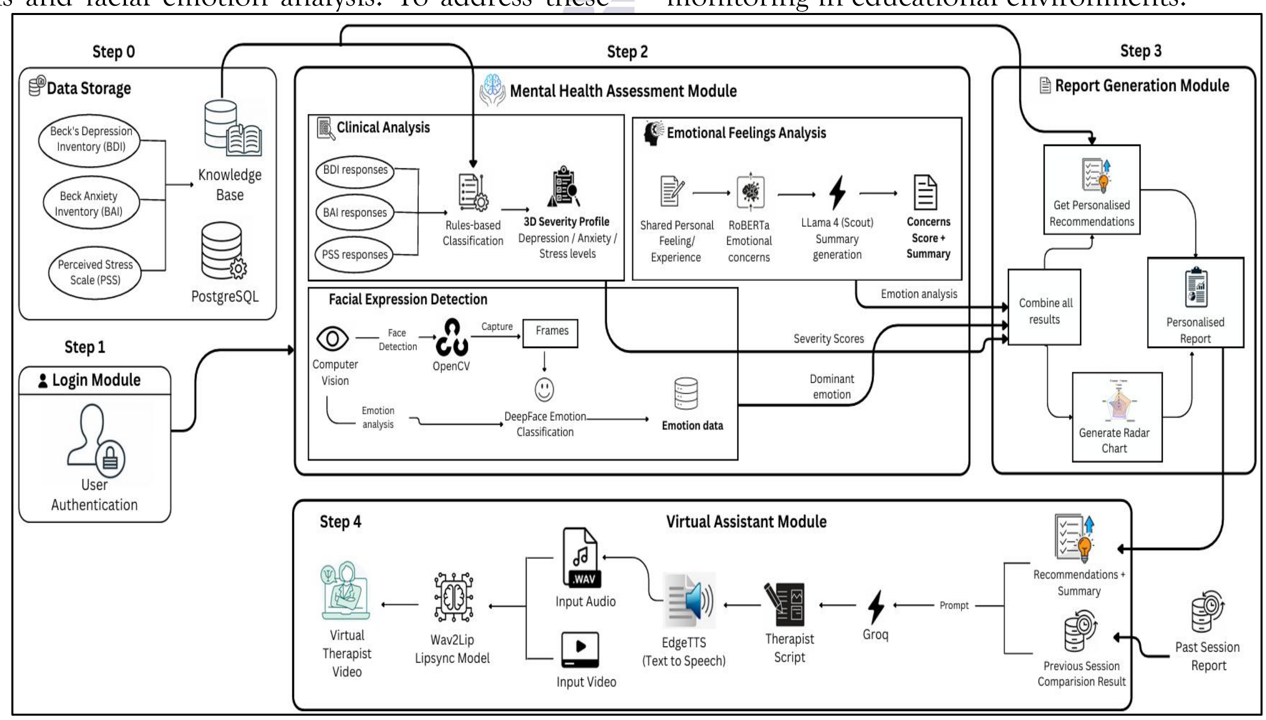

Figure 1: Proposed Methodology

Student Companion is a multimodal mental health support system that was designed in response to the increasing mental health problems among the students. It was constructed on one principle: a student who needs mental health help should be faced with a system that is nonjudgemental, easy to approach and in a way that is clear and sincere. By keeping this principle in mind, the system works in a sequential manner.

It combines two distinct sources of information about a student's mental health: the answers that a student gives to a structured questionnaire about how they have been feeling and emotional expressions that appear on their face while interacting with the system. Integrating these two streams into one, system will gain full-scale and more precise image of the mental well-being of the student and end up with the personalized report and virtual therapy session.

The sequence is smooth and predictable, which is intentional as familiarity and clarity will also play a vital role in providing comfort to the students who are experiencing any mental health issue. The system was constructed based on four fundamental steps: These include Login interface, mental health assessment module, the report generation module that combines mental health assessment result and the last one is virtual therapy assistant module that provides custom-tailored strategies to cope and recommendations. All these are detailed in the following sections.

3.1 Data Storage & Knowledge Base

Before any student logs in, the system is already operating silently.In its structure, a PostgreSQL database awaits. A large, systematic repository designed to store all knowledge the system will acquire about its students. Each account, each session, each response, every feeling documented everything is organized here, seamlessly connected and accessible, ready to be retrieved when necessary.

Also, there is a clinical knowledge base that consist of three meticulously selected clinical questionnaires, an instrument that have been evaluated, confirmed, and relied upon by psychologists for years. As a student starts their journey, they will encounter these questions:

The Beck Depression Inventory (BDI): 21 inquiries softly examining the burden a student might bear. [18]

The Beck Anxiety Inventory (BAI): 21 inquiries designed to identify the subtle presence of anxiety lurking underneath. [19]

The Perceived Stress Scale (PSS): 10 questions assessing a student's feelings of being overwhelmed or in control in their everyday life. [20]

Every questionnaire is accompanied by specific scoring guidelines and severity levels which form the basis of Rule based Classification in step 2

3.2 Login Module:

Secure Login Interface is the entry point into Student Companion system. Security in the context of the use of mental health is not only a necessity but also a source of trust. Students struggling with depression or anxiety may not utilize the system that they do not perceive to protect their privacy. For this reason, a secure authentication mechanism is implemented.

The system uses a secure authentication mechanism using Django-all auth , which is a package for Django that handles both traditional and social login, Student can create their accounts using their email address and set a password or they can sign in using their existing Google accounts using OAuth 2.0 which provide ease to the students who may get frustrated by going through a lengthy registration process. After successful login it assigns a session ID to make sure all the questionnaire responses, facial expression results, mental state reports, and therapy recommendations are stored in the database and are accessible to the dedicated student.

Once student successfully logs in, the system transports them to their personal dashboard. It fulfils two major purposes. First, it provides a clear way of starting new session with the virtual assistant. Second and, perhaps, more importantly, it serves as a reflection of the mental health journey of the student through time.

Whenever a student completes their session, it will be shown on their dashboard which includes the date the session took place, mental health problems that were identified and their severity levels in the last session, previous reports, recommendation and coping strategies summary that is suggested by therapy session module.

When the student's glance at their dashboard they are not only seeing the history of its past sessions, but they are also reading a story of the path of their own mental health which may be empowering. Students can also track improvements over time and can even recognize certain patterns which gives students a meaningful reason to continue engaging with the system. The dashboard is dynamically updating at the conclusion of each new session, to indicate the latest view of the well-being of the students. Once students click on starting a new session Mental health assessment module gets triggered to analyze the current mental health of students.

3.3 The Mental Health Assessment Module

This module operates through three sub-modules; each provides a distinct dimension of student mental state.

3.3.1 Clinical analysis

When a student starts a new session with virtual assistant, the mental health assessment module getsactive, which works by asking students a set of questions and analyzing their responses to identify symptoms of depression, anxiety and stress. These set of questions are designed using BDI (Beck's Depression Inventory), BAI (Beck Anxiety Inventory), PSS (Perceived Stress Scale) which is one of the most rigorously validated and widely used psychometric instruments in clinical psychology.

It consists of 52 questions distributed across three subscales twenty-one items for depression, twenty-one for anxiety, and ten for stress. Each questions asks student to rate on a scale from 0 – 3 about how much a particular statement applied to them. The easy language and short format enable students to understand it easily and its reliable scoring system ensures meaningful and comparable results.

The questionnaire is displayed on screen one at a time to minimize the cognitive load and promote deliberate responses. This format causes the process to appear more of a conversation than an interrogation. To avoid biased answers, the student never sees the total of his/her scores running during the session because of a phenomenon termed demand characteristics as the participants change the answers on what the system anticipates them to provide. [21]. After the student has given the answers to all the questions, his or her answers are processed through a rule-based classification engine by the system. The choice of the rule-based approach instead of a machine learning model was not taken without a reason and without thoughtful consideration.

Rule-based systems are completely open in the process of transmitting a result and each classification can be traced down to a particular set of scoring rules hence the system can always justify how it came to a certain conclusion. The classification engine works by summing the student's responses for each of the three subscales and multiplying the result by two, in accordance with the standard scoring convention for the twenty-one-item version. The resulting score in each subscale is then compared against published severity thresholds to determine whether the presentation of the student is in the normal, mild, moderate, High or extremely severe range. In the case of depression, a scale of zero to nine is regarded as normal, a scale of ten to thirteen is mild depression, twenty-one to twenty-seven is severe depression, and twenty-eight and above is extremely severe depression. There are also similar threshold ranges of anxiety and stress, the definite cut-off values of which vary among subscales provided in the original DAI, PSS manual. The output of this process is a three-dimensional severity profile for each student for example, moderate depression, mild anxiety, and normal stress which is then passed to the report module.

3.3.2 Emotional Feeling Analysis

Once the student is done responding to all the questions of the questionnaire, the session is not terminted immediately by the system. Rather, it poses one final question, which is a simple, open question, with no pre-determined answers, and it is: **What would you like to share with us?** There is a reason why this question exists.

The pre-test questionnaire was able to gauge the extent to which a student may be struggling but it was never capable of doing so in such a way that it could actually understand how that struggle felt in their own words. So this last prompt provides students with space. Some write just a line or two. There are others who have a lot more.

The system does not take lightly whatever they write. After the student provides his or her answer, the text undergoes two levels of analysis. It is first sent to **RoBERTa**, a text-reading model that is trained to comprehend textual emotions. [22] RoBERTa reads the text that the student wrote and finds out the emotional tones that it has, such as sadness, fear, frustration, or loneliness. It does not generate any one emotion label but a collection of labels, which in combination represent the emotional contents of the response.

These labels, together with the text that the student originally wrote are then given to **Llama 4 (Scout). [23]** This model has words and the emotion labels and writes a readable, brief summary of the manner in which the student is feeling. The aim of such a summary is not to diagnose anything - it is just to provide a clear and significant picture of the emotional state of the student that can be quickly perceived by a counsellor or advisor when they read the report.

Nonetheless, the use of AI models alone poses some risks. Such models as RoBERTa and Llama do not necessarily work correctly in identifying or treating sensitive or risky material. They may not notice the little hints, misunderstand motives, or have no urgency in dire circumstances.

To overcome this weakness, another rule-based safety filter is used alongside the AI models. This component directly works on raw input of the student to identify the presence of certain high-risk indicators, such assuicidal thoughts, self-harm, violent intentions, or harm to others. This filter, despite being similar in the use of AI models, does not seek to analyze or summarize the content; it merely searches through it to find some predetermined patterns and keywords. In case any of such indicators are identified, the system will have a certain mention in the final report.

The system balances the three by combining the strengths of RoBERTa to detect emotions, Llama 4 (Scout), to summarize, and a rule-based safety mechanism to detect risks. It not only captures and conveys the emotional experience of the student but also creates a high level of safety since the situations which could be dangerous are never swept under the carpet.

3.3.3 Facial Expression Recognition Analysis

While the student is answering the questionnaire, facial expression recognition is taking place in parallel with the mental health assessment module. When a new session is launched, it requests student to provide permission on using camera with OpenCV. If a student grants access, then the facial expression module begins silently and capture the frames throughout the session at a specific time interval rather than capturing every single frame to maintain efficiency.

The capture process is designed as being as natural as possible so as to ensure that the student emotional expressions remain unaffected by the process of monitoring. When the student refuses to use the camera, the system proceeds to the analysis of the questionnaire alone and places a note in the final report that the data on facial expressions was absent during the given session.

A pre-trained deep learning model (Deep Face) [24] is used to capture facial expressions of each captured frame separately and then classify them in seven discrete emotional states. These seven states are happiness, sadness, anger, fear, disgust, surprise and neutral expression. The model generates a probability value of each frame in seven emotion categories and the emotion where the probability is greatest is the dominant expression of that frame.

Once the session ends, these per-frame predictions are summed up in the system to form a session-level profile of emotions in the form of a frequency distribution of how many times each of the seven emotions was detected in all captured frames. This emotional information is forwarded to the report module where it is combined with the scores of the student questionnaire.

In order to present this data in a clinically significant manner, the system initially transforms these raw forms of emotional data into psychological insights based on existing studies. When the system identifies a consistent trend of sadness in a session, then this is interpreted as a supporting symptom of depression which displays the heavy, lingering quality of the condition. Similarly, frequent expressions of fear are associated with anxiety, and frequent expressions of anger or disgust are associated with increased stress or loss of emotional control in the student.

On the other hand, the presence of a stable level of neutral or happy expressions reveals a more positive picture, accusing that the student did not experience any serious cases of emotional shock during the evaluation.

3.3.5 Report Generation Module

The next step of the Student Companion pipeline is the Report Generation Module, at which all the findings of the mental health assessment modules are summarized and made available as a single report.

After the Mental Health Assessment Module has finished analysing the data, the Report Generation Module gathers three types of information: the severity scores from the clinical questionnaire, the dominant emotion profile from facial expression recognition, and the emotional feeling analysis output (RoBERTa emotion labels, Llama 4 (Scout) generated summary, and any safety filter flags). These are all then put together and presented in a structured report which is easy to read for the student and any reviewing professional.

The report is divided into two main sections.

The former section shows the results of clinical severity in the three of the subscales of depression, anxiety, and stress. Individual subscales are presented as individual cards with the severity label, the raw score, an interpretation in plain language and a list of personalized recommendations. If facial expression data was captured, the dominant emotion observed is also noted within the card. If no valid camera frames were available, this is clearly mentioned.

The second section presents the Emotional Reflection Analysis based on the student's open-ended personal statement. It displays the Llama 4 (Scout) generated summary describing how the student appears to be feeling in empathetic and readable language. The RoBERTa detected emotion labels are then visualized through a radar chart and a horizontal bar chart, both showing the intensity of each detected emotion such as despair, anxiety, vulnerability, or humiliation etc on a scale of zero to one.

The report ends with a standard Medical Disclaimer to remind the reader that it is an AI aided screening device and is not a clinical diagnosis and the findings must always be reviewed by a qualified mental health professional. After the session is done, all report-related data is stored to the account of its student and displayed on their customized dashboard as a continuation of their mental health history.

3.4 Virtual Therapy Session Module

Once the report is created, the student has a choice of starting a virtual therapy session. When the student clicks the “Talk toTherapist” button, the Virtual Therapy Session Module is launched, and this follows the following sequential pipeline.

3.4.1 Facial Expression Mapping

The module has already set the criteria for mapping the seven facial emotions to the three core psychological dimensions of depression, anxiety, and stress. It maps the dominant facial emotion with the defined criteria. Emotions that do not correspond meaningfully to any of these dimensions are categorized as none.

Table 1: Facial emotion and mapped dimension

Facial Emotion	Mapped Dimension
Sadness	Depression
Fear	Anxiety
Anger	Stress
Disgust	Stress
Surprise	None
Happiness	None
Neutral	None

3.4.2 Severity Level Adjustment

The level of severity in relation to any of the three dimensions is increased one step higher where the facial expression mapping supports any of the three dimensions. Examples include low to moderate, moderate to high or severe. This modification will see to it that non-verbal communication will not be ignored when developing the therapy response.

3.4.3 Recommendation Extraction and Past Session Comparison

The module then derives specific recommendations depending upon the adjusted level of severity of the dimension. At the same time, it retrieves the two latest previous sessions of the student to determine whether he or she has been improving, stable or worsens over time.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

3.4.4 Therapist Script and Audio Generation

The past session data, emotional summary and recommendation is forwarded to Groq [25] with a formatted prompt to produce a custom therapist script. This script is then sent to the EdgeTTSwhich translates it to a natural sounding audio file.

3.4.5 Virtual Therapist Video Generation and Display

The audio and a therapist demo video are pre-recorded and handed over to Wav2Lip[26], a locally integrated lip-synchronization model, which renders a synchronized video of a virtual therapist. After generation is done, the video is shown to the student and a graphshowsthe time taken to generate the video. Recommendation logic, generative AI, speech synthesis, and video processing are unified in this module to create a single experience that by the time a student meets with their virtual therapist, the experience they get is not a generalized one but one that responds to who they are and what they are experiencing.

4. RESULTS

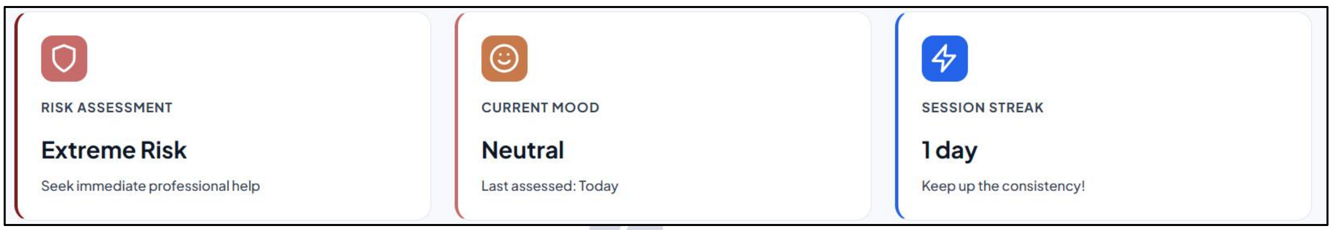

Figure 2: User Interface

4.1 Risk Assessment and Current Mood:

Among the top cards in the dashboard are Risk Assessment, Current Mood, and Session Streak. Risk Assessment as well as Current Mood represent an essential moment in the psychological condition of the person at the time when the evaluation took place. In the recent evaluation of the situation, a Severe Risk was noted, and this means that a prompt medical treatment is needed. The risk level mentioned above refers to the Angry mood state, and this means that there was some emotional distress involved. While Session Streak implies the beginning of using the program, the top cards imply that emotional instability goes beyond the scope of the usual academic stress.

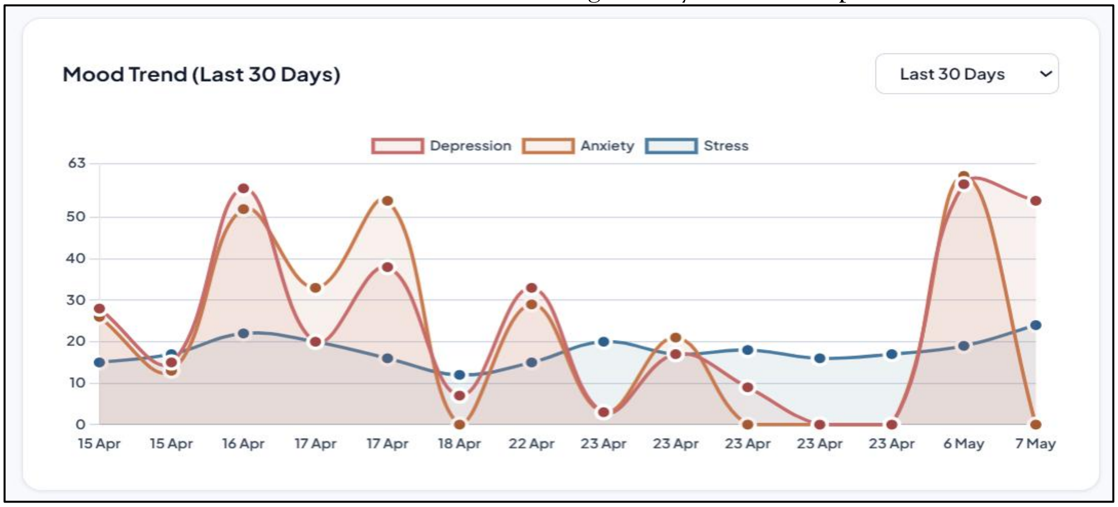

Figure 3: Assessment group

4.2 Mood Trend

The Mood Trends line graph illustrates a longitudinal study of the changes of emotions of the participant where we observe notable spikes of depression and anxiety measures from April 16th to April 17th. During this period, we can see that measures for anxiety and depression spiked far beyond the baseline and were very close to the maximum level. These observations are connected with the high risk assessment. As for the odds, measures for stress stayed fairly stable during the whole month but still higher than average. In such a way, the user experiences some kind of permanent stress, while depressive and anxious moods come suddenly.

Figure 4: Wellness Factors

4.3 Wellness Factor

Wellness Factors merely becomes the opposite of the obstacles that a participant is going through. This breaks down the increasing influence of these emotions in the sense that stress and anxiety are the main reasons for the current mental health state of a participant. Stress is the greatest wellness factor among others with a percentage of 62.5%, while anxiety accounts for 54% and depression comes in third at 47.6%. The percentages are derived from the ratio between the scores of the assessment and their worst possible values. It shows how there is a complicated relationship wherein high levels of stress act as a trigger for anxiety, which then becomes depression.

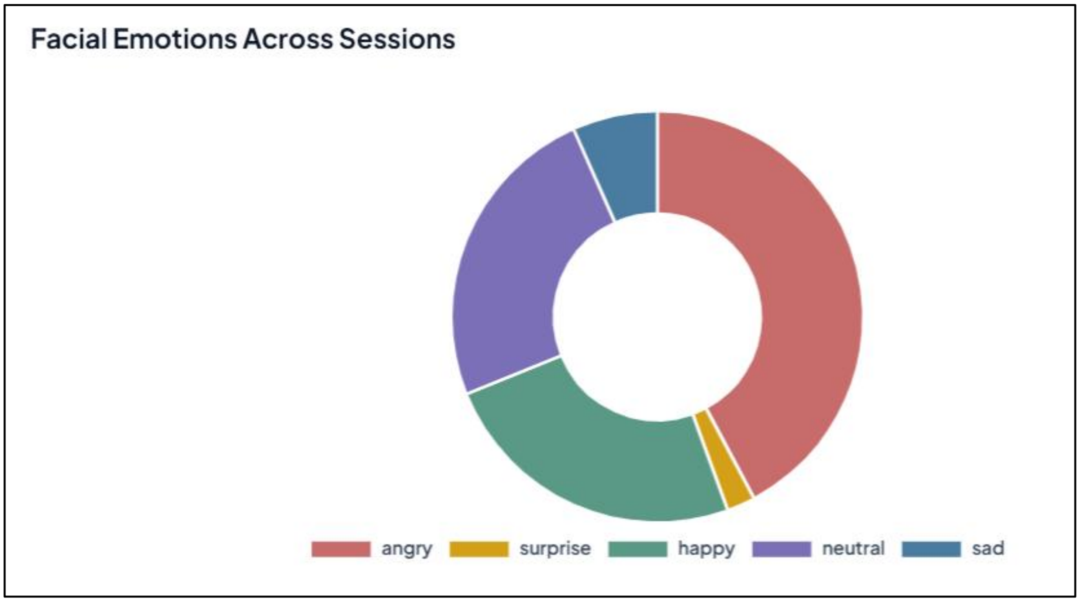

Figure 5: Access Session

4.4 Facial Emotions Across Sessions

The Facial Emotions Across Sessions is depicted using a donut graph, which shows a distribution of mixed emotions observed while the user is interacting, with "Angry" and "Happy" dominating the larger portion of the overall dataset. The prominence of the negative emotions, such as anger, and to some extent, surprise, implies that there are many times when the emotions were highly aroused. Nevertheless, the large amount of "Happy" and "Neutral" emotions depicted in this donut graph implies that the user still has some time where he feels emotionally balanced despite his increasing level of distress observed in the other modules.

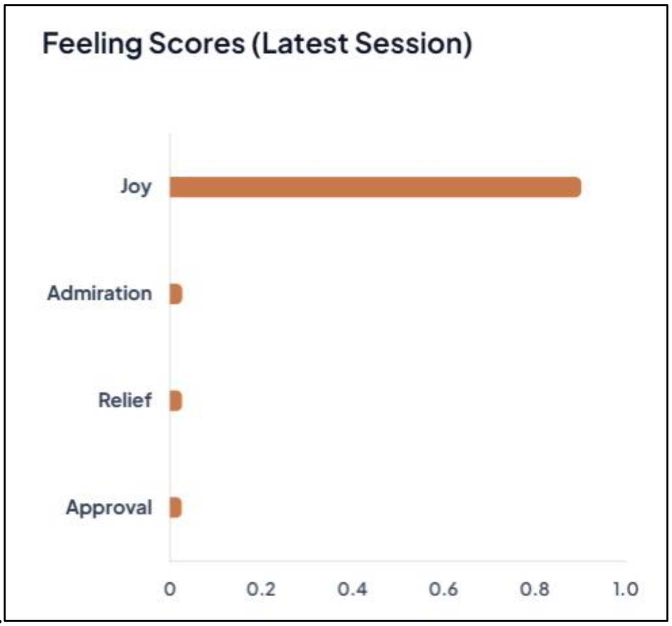

Figure 5: Feeling Score

3.5 Feeling Score

Another bar chart is the one illustrating the Feeling Scores, giving us an insight into the specific negative impacts induced by the most recent assessment. Anger is seen to be the prevailing emotion with a score of about 0.7, which is significantly higher than both Annoyance or Disapproval. The prevalence of such a highly intense feeling can be explained by the Severe Risk rating that has been highlighted in the overview. The user's current state of dissatisfaction goes beyond that of being annoyed. It seems to be something much more powerful than that. These disparities between scores show that the user is prone to having emotional breakdowns.

ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

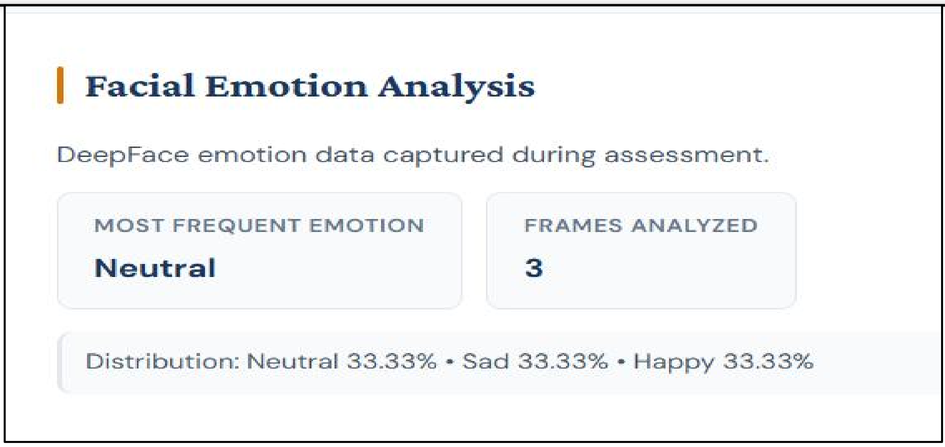

Figure 6: Emotion Analysis

3.6 Facial Expression Analysis

This figure shows the results of the facial emotion analysis via DeepFace, which occurs in the background while the student is taking the questionnaire. Over 3 session frames, the emotions were spread equally across Neutral, Sad and Happy - each at 33.33% - with Neutral being the most frequent emotion.

The distribution of three emotions, with equal weight, suggests the student's facial expressions changed during the session rather than being uniformly reflective of one emotion. Although the most frequent emotion will always be Neutral, it's important to note that Sadness is present with an equal weight here - and when sent to the Report and Virtual Therapy modules, this information is cross-mapped with the questionnaire scores to give a more comprehensive understanding of the student's mental health than either data source can offer on its own.

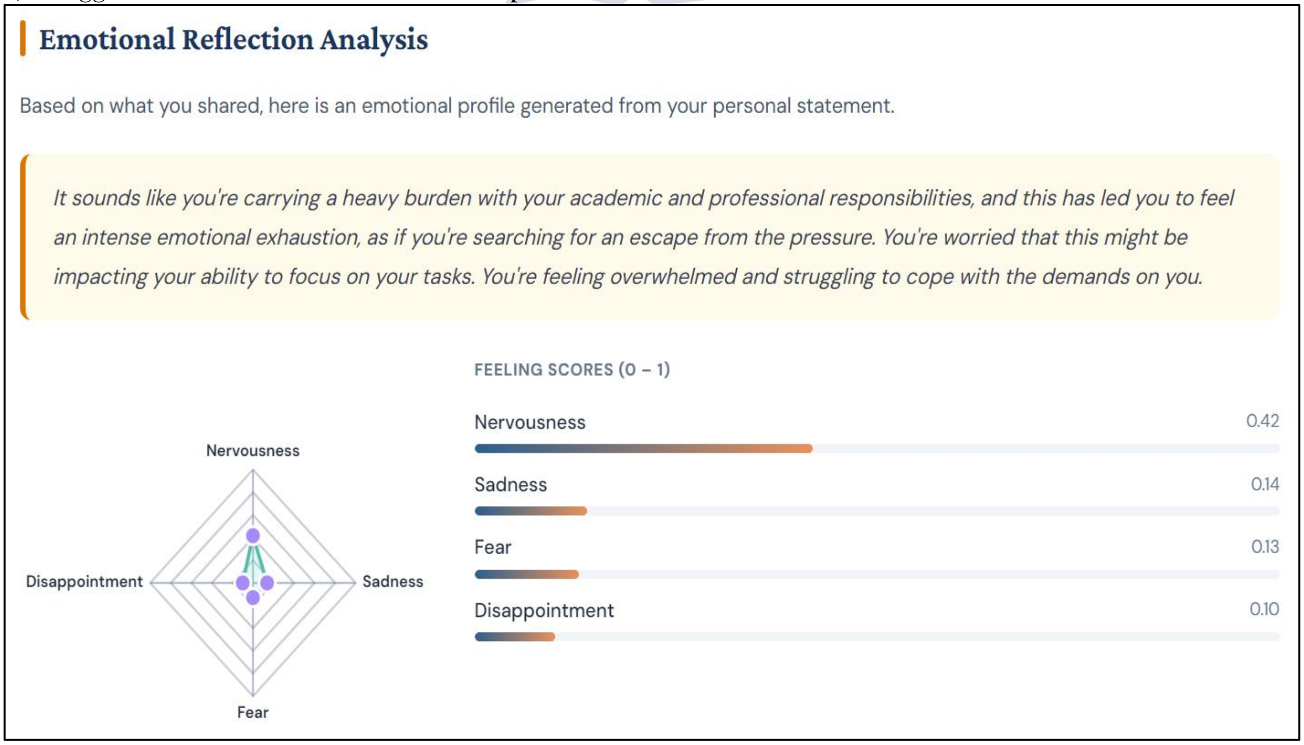

Figure 7: Reflection Analysis

3.7 Emotional Reflection Analysis

Once the student has completed the questionnaire, the final question is an open prompt - "Is there anything else you would like to tell us?" - a place for the student to express themselves freely. This passes through the Emotional Reflection Analysis, where RoBERTa identifies the emotional tones, and Llama 4 (Scout) translates them into a narrative, while a rule-based safety filter is running in parallel to look for any red flags.

In this case, the safety filter was triggered. The students' comments contained explicit indicators of suicidal thoughts (scored at the highest level of 1.00), sadness (0.58), and traces of disappointment, nervousness and remorse. The narrative reads like this student was a lonely, guilty and emotionally drained student who is suffering.

The radar and bar charts give emotional profile its visual shape, making severity clearly visible to any professional eye. The section concludes with a medical disclaimer, emphasizing that these observations are a screening tool, not a diagnosis, and the call to action for human intervention.

3.8 Virtual Therapy Session Module

This figure represents the last step in the Student Companion: the moment the system stops analysing and starts speaking. Following the report's generation, the student is presented with their Virtual Therapist, an AI-generated video with lip-synchronized speech created with Wav2Lip, that delivers a generated therapist script by Groq and voiced by EdgeTTS.

In the video above, the student has received a report with Depression, Stress and Anxiety levels at a Low severity. As such, the therapist script - displayed at the bottom - begins with a positive affirmation, noting depressive symptoms and assuring the student that anxiety and stress are low.

After the video, the Personalized Recommendations section on the right side expands - showing the same coping strategies and insights that were provided verbally throughout the session. The rationale for this order is that the student first consumes the information verbally and then is presented with recommendations in written form for them to read, consider and commit to. The three aspects of Depression, Stress and Anxiety are presented in collapsible cards, so the student can access specific advice for their needs.

The student is not engaging with a chatbot or report, but with a professional and calm person speaking directly about their results - making the experience feel less like listening to a computer and more like listening to a person

.

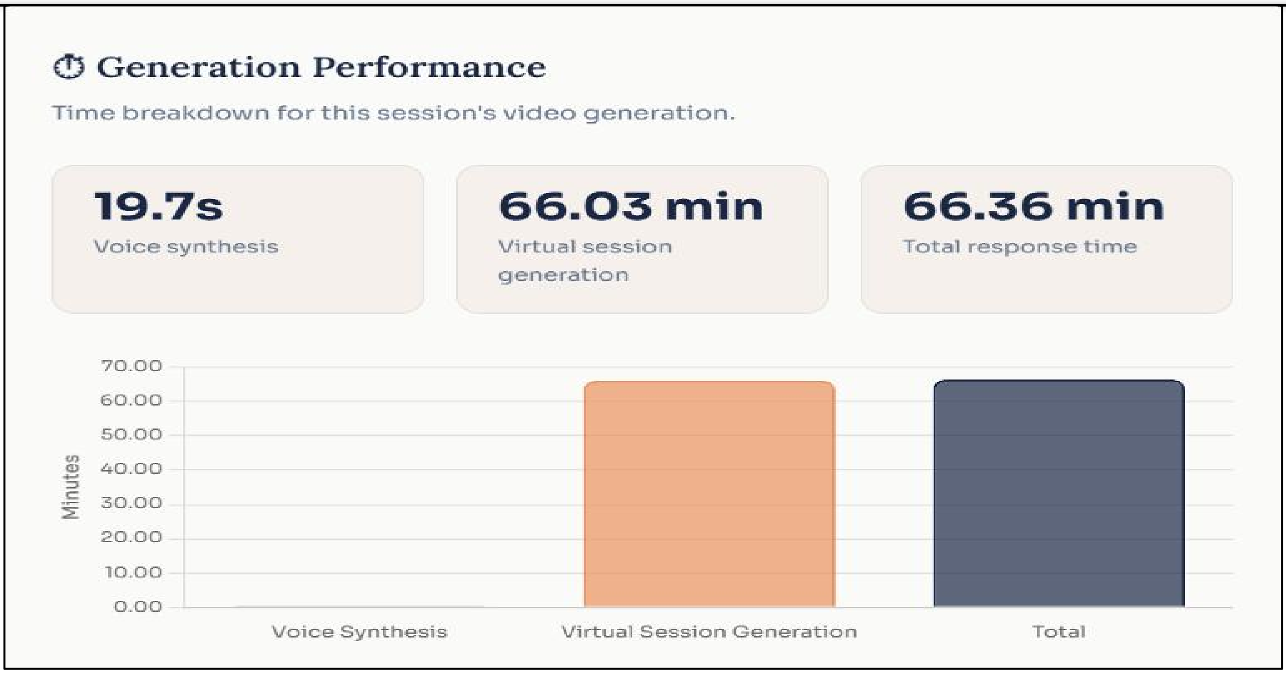

Figure 8: Performance

3.9 Generation Performance — Virtual Therapy Session

This figure illustrates the time required to produce the virtual therapist. The virtual therapist's voice is generated (using EdgeTTS) in 19.7 seconds, but the more computationally intensive lip-synchronization process (using Wav2Lip) - mapping the voice onto a prerecorded video of the therapist - takes 66.03 minutes, resulting in a response time of 66.36 minutes.

This is clearly shown in the bar chart, with video processing taking up most of the time. This is a trade-off: Wav2Lip is run locally and is more accurate than fast. The response time is still a weakness of the system and priority for improvement.

4. COMPARIsON WITH EXISTING SYSTEM

Table X is a comparison of Student Companion with other mental health platforms in use. Current tools are based on self-reported data or only on text-based interaction with text-chatbots. Student Companion is unique in its implementation of multiple measures of assessment, such as the recognition of facial emotion, clinical questionnaires, and analysis of the National Language (NLP), and the fact that it does not require the student to take the initiative to seek help, which in Pakistan's cultural context is the biggest hurdle.

Table 2: Feature Comparison with Existing Mental Health Platforms

Feature	Wysa	Woebot	Umang (PK)	Humraaz (PK)	Student Companion
Clinically validated questionnaire	Partial	Partial	X	X	✓
Facial emotion recognition	X	X	X	X	✓
NLP-based emotion analysis	X	Partial	X	X	✓

Suicidal ideation safety filter	Partial	✓	Partial	Partial	✓
Virtual therapy session	X	X	X	X	✓
Proactive detection	X	X	X	X	✓
Multimodal assessment	X	X	X	X	✓
Session history tracking	X	Partial	X	X	✓
Pakistan-specific design	X	X	✓	✓	✓

5. Limitations and Future Work

5.1 Limitations

- **Video Generation Speed:** Although Wav2Lip can work in real-time, it requires about 66 minutes to create one therapy video locally which is not yet feasible in its current state.
- **No Clinical Validation:** Not tested with actual students in a supervised clinical setting. Before use, severity classifications must be validated by licensed psychologists.
- **Camera Dependency:** You must grant camera permissions to be able to use facial recognition. The assessment is performed on the basis of the questionnaire in the event of its refusal, which results in less completeness.
- **English Only:** Currently the system is English only. Some Pakistani students may be at ease in Urdu and this can lead to less authenticity of open-ended answers.
- **General-Purpose AI Models:** RoBERTa and Llama 4 are not trained on student data from Pakistan, and thus might misinterpret culturally specific or code-switched expressions of distress.

5.2 Future Work

- **Clinical Trial:** A formal study to validate the accuracy of the system with licensed psychologists of a Pakistani university.
- **Urdu Language Support:** Increase accessibility and authenticity by adding Urdu to the interface, questionnaire and therapy scripts.

Faster Video Pipeline: To cut down on the generation time by replacing the Wav2Lip model with a faster one or pre-rendered video segments.

Mobile Application: A mobile version to make it easier for students to access and for regular checks during the day.

Counsellor Dashboard: A protected dashboard for university mental health personnel to track the trend of risk and get alerts of students who are flagged as severe risk.

6. conclusion

This paper suggests Student Companion, which is an AI based multimodal system to address the mental health crisis of university students especially in Pakistan. The system will offer the efficient adaptation of technology to actively detect psychological distress early without waiting for students to take the initiative and seek help.

The multimodal method uses self-reporting explicitly and implicitly via emotional cues, which are not readily accessible or welcome by the students. This dual-channel measurement tool is especially relevant in culturally sensitive settings where the stigma around mental well-being hinders information exchange. The rule-based classification system guarantees clinical validity due to severity measures that can be mapped to specific psychometric measures, whereas machine learning makes such measures untraceable. The combination of safety filters and artificial intelligence to produce an emotional summary is something that must receive immediate attention from humans, that is, high-risk factors like suicidal tendencies that need immediate attention from humans. Virtual Therapy Session Module has a coping strategy tailored to the severity.

6.1 Concluding

Solutions for addressing the problem of student mental well-being must include technological advances in artificial intelligence as well as the clinical standards that will allow the identification of those who have problems and might pose a danger to themselves. Despite the continuing challenges of computing efficiency and empirical validity, the proposed framework serves as a basis on which to implement culturally responsive mental well-being services, changing the nature of the services provided from reactive to preventive ones. It is not the level of technical complexity that defines the success of the endeavor, but rather its ability to connect with students who currently are not being served by the existing system.

7. REferences

[1] World Health Organization, "Suicide." Accessed: May 02, 2026. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/suicide

[2] World Health Organization, "Mental disorders." Accessed: May 02, 2026. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders

[3] K. A. Ganasen, S. Parker, C. J. Hugo, D. J. Stein, R. A. Emsley, and S. Seedat, "Mental health literacy developing countries : review article," *Afr. J. Psychiatry*, vol. 11, no. 1, pp. 23–28, Feb. 2008, doi: 10.10520/EJC72655.

[4] M. U. Rehman, "Academic Stress and Mental Health Issues among University Students in Pakistan," *ApexMed J. Health Sci.*, vol. 1, no. 3, pp. 09–16, 2025.

[5] W. Husain and F. A. Faize, "Public awareness of psychological problems in Pakistan," *Ment. Health Rev. J.*, vol. 25, no. 1, pp. 35–45, 2020.

[6] M. M. Khan, "Suicidal behaviours in Pakistan," in *Suicidal behavior in Muslim majority countries: Epidemiology, risk factors, and prevention*, Springer, 2024, pp. 223–243.

[7] M. Shakil, "A qualitative analysis of suicides committed by the students in Pakistan," *Pak. J. Med. Res.*, vol. 58, no. 1, pp. 35–40, 2019.

[8] B. Ahmed, S. F. Enam, Z. Iqbal, G. Murtaza, and S. Bashir, "Depression and anxiety: a snapshot of the situation in Pakistan," *Int. J. Neurosci. Behav. Sci.*, vol. 4, no. 2, p. 32, 2016.

[9] A. Fleischmann, "Suicide in the world," *Suicide Unnecessary Death*, p. 1, 2016.

[10] H. Shahzain and A. Munsif, "A sudden rise in suicide in students in Pakistan," J. Neurosci. Neurol. Surg., vol. 4, no. 4, 2019.

[11] R. A. Mahesar, S. A.-Z. Ali, S. Shoib, M. Babar Khan, and A. Ventriglio, "Mental health services in Pakistan," *Int. Rev. Psychiatry*, vol. 37, no. 3–4, pp. 378–383, May 2025, doi: 10.1080/09540261.2024.2391796.

[12] F. R. Choudhry, N. Khan, and K. Munawar, "Barriers and facilitators to mental health care: A systematic review in Pakistan," *Int. J. Ment. Health*, vol. 52, no. 2, pp. 124–162, Apr. 2023, doi: 10.1080/00207411.2021.1941563.

[13] A. J. Morgan and A. F. Jorm, "Outcomes of self-help efforts in anxiety disorders," *Expert Rev. Pharmacoecon. Outcomes Res.*, vol. 9, no. 5, pp. 445–459, Oct. 2009, doi: 10.1586/erp.09.47.

[14] P. O'Leary, N. Carroll, and I. Richardson, "The Practitioner's Perspective on Clinical Pathway Support Systems," in *2014 IEEE International Conference on Healthcare Informatics*, Sep. 2014, pp. 194–201. doi: 10.1109/ICHI.2014.33.

[15] S. Shen *et al.*, “Passive sensing for mental health monitoring using machine learning with wearables and smartphones: Scoping review,” *Journal of Medical Internet Research*, vol. 27, p. e77066, 2025. doi: 10.2196/77066.

[16] X. Sun, A. Jahangir, A. A. Alghamdi, and F. Liu, “Use of AI-based mental health tools and psychological well-being among Chinese university students: A parallel mediation model of emotional self-efficacy and perceived autonomy,” *Scientific Reports*, vol. 15, no. 1, p. 40305, Nov. 2025, doi: 10.1038/s41598-025-24013-8.

[17] N. Aggarwal et al., "Advancing artificial intelligence in health settings outside the hospital and clinic," NAM Perspectives, National Academy of Medicine, Washington, DC, 2020, doi: 10.31478/202011f.

[18] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, and J. Erbaugh, "An inventory for measuring depression," Arch. Gen. Psychiatry, vol. 4, no. 6, pp. 561–571, 1961.

[19] A. T. Beck, N. Epstein, G. Brown, and R. A. Steer, "An inventory for measuring clinical anxiety: Psychometric properties," J. Consult. Clin. Psychol., vol. 56, no. 6, pp. 893–897, 1988.

[20] S. Cohen, T. Kamarck, and R. Mermelstein, "A global measure of perceived stress," J. Health Soc. Behav., vol. 24, no. 4, pp. 385–396, 1983.

[21] A. L. Nichols and J. K. Maner, “The Good-Subject Effect: Investigating Participant Demand Characteristics,” *J. Gen. Psychol.*, vol. 135, no. 2, pp. 151–166, Apr. 2008, doi: 10.3200/GENP.135.2.151-166.

[22] Y. Liu et al., "RoBERTa: A robustly optimized BERT pretraining approach," arXiv preprint arXiv:1907.11692, Jul. 2019.

[23] A. Shukla, "Llama 4: Meta's multimodal leap in AI — architecture, capabilities, and comparative analysis," arXiv preprint, 2025.

[24] S. I. Serengil and A. Ozpinar, "LightFace: A hybrid deep face recognition framework," in Proc. 2020 Innovations in Intelligent Systems and Applications Conf. (ASYU), Oct. 2020, pp. 23–27, doi: 10.1109/ASYU50717.2020.9259802.

[25] L. Gwennap, "Groq rocks neural networks," Microprocessor Rep., vol. 1, 2020.

[26] K. Prajwal, R. Mukhopadhyay, V. P. Namboodiri, and C. Jawahar, “A lip sync expert is all you need for speech to lip generation in the wild,” presented at the Proceedings of the 28th ACM international conference on multimedia, 2020, pp. 484–492.