



AI-Driven Multimodal Mental Health Monitoring System Using Emotional, Physiological and Self-Reported Data

Rohini Thimmapura Venkatesh^{1*}, Chaitra Mrutyunjaya², Sridevi Gereen Murthy³,
Pavitra Bai Srinivasa Rao⁴, Ankitha Ravindranath¹, Naman Chopra¹

¹ Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Visvesvaraya Technological University, Bangalore 560111, India

² Department of CSE, BNM Institute of Technology, Bangalore 560070, India

³ Department of Information Science and Engineering, Jyothy Institute of Technology, Bangalore 560082, India

⁴ Department of Information Science and Engineering, SJB Institute of Technology, Bangalore 560060, India

Corresponding Author Email: rohinitv@gmail.com

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300801>

ABSTRACT

Received: 24 June 2025

Revised: 6 August 2025

Accepted: 15 August 2025

Available online: 31 August 2025

Keywords:

facial expression recognition, infrared temperature, respiration monitoring, psychological questionnaires, real-time assessment, non-invasive

Mental health challenges are on the rise, driven by modern lifestyles, workplace pressures, and social stressors. Conventional assessments often rely on self-reported questionnaires and occasional clinical visits, which can miss real-time changes and sometimes lack objectivity. To address this gap, we propose an AI-driven mental health monitoring system that uniquely integrates multiple non-invasive modalities: facial expression recognition, respiration analysis, infrared body temperature sensing, and the Questionnaire-based Model. This multimodal approach allows for continuous and holistic assessment of both emotional and physiological states. Facial expression analysis helps identify emotions such as happiness, sadness, anger, and stress; respiration monitoring captures irregular breathing patterns linked to anxiety; and temperature sensing highlights stress-induced variations. These signals are further complemented by a psychological questionnaire, which achieved the highest predictive accuracy of 93%, underscoring its effectiveness when combined with physiological cues. By uniting these different perspectives, our system not only improves detection accuracy but also reduces bias, offering a more reliable tool for early mental health intervention compared to traditional single-method approaches.

1. INTRODUCTION

Mental health is a fundamental aspect of overall well-being, encompassing emotional, psychological, and social dimensions. It influences how individuals think, feel, and behave, playing a crucial role in decision-making, interpersonal relationships, and the ability to cope with stress. Good mental health is not merely the absence of mental illness but also involves emotional resilience, cognitive stability, and an overall balanced state of mind.

In today's fast-paced world, mental health disorders are becoming increasingly prevalent due to factors such as work pressure, financial stress, social expectations, and digital overexposure. According to the World Health Organization (WHO), mental health disorders affect millions of people worldwide, significantly reducing their quality of life, productivity, and social functioning. Despite their widespread impact, many mental health conditions go untreated due to stigma, lack of awareness, and limited access to professional care. Early intervention is crucial, but current mental health assessment techniques do not provide real-time monitoring, making timely intervention challenging.

Traditional methods for diagnosing mental health conditions primarily rely on self-reported assessments,

structured clinical interviews, and periodic psychological evaluations. While these methods have been widely used, they present several limitations: they are subjective, invasive, resource-intensive, and infrequent. Self-reporting can be influenced by memory biases, reluctance to disclose symptoms, and social desirability, leading to inaccurate assessments. Clinical interviews require trained professionals and significant time, making large-scale mental health monitoring impractical. Furthermore, invasive physiological tests such as electroencephalography (EEG) or electrocardiography (ECG) require specialized equipment and medical supervision, limiting their accessibility [1]. These shortcomings highlight the urgent need for innovative, real-time mental health monitoring solutions that are non-invasive, scalable, and objective.

Recent advancements in artificial intelligence (AI), data science, and sensor-based monitoring systems have paved the way for real-time, non-invasive mental health assessment techniques. Unlike traditional methods, these modern approaches leverage physiological and behavioral indicators such as facial expressions, voice patterns, respiration rates, and body temperature to detect emotional states and mental health anomalies. Wearable devices, smart homes, and mobile applications now integrate mental health monitoring systems

to provide users with real-time feedback on their emotional and psychological well-being. Wireless vision-based and acoustic sensors track subtle changes in behaviour and physiological responses, enabling early detection of stress, anxiety, and depressive symptoms [1]. AI-driven predictive models further enhance these capabilities by analyzing patterns in multi modal data sources, offering personalized recommendations and early interventions.

This study presents a comprehensive, real-time mental health monitoring system that integrates multiple non-invasive technologies to assess emotional and physiological well-being. The proposed system combines facial expression recognition, respiration analysis, infrared temperature sensing, and psychological questionnaires to provide a holistic evaluation of an individual's mental state. By leveraging AI-driven analysis and multimodal data fusion, our approach enhances accuracy in detecting stress, anxiety, and emotional fluctuations. Unlike traditional self-reported methods, our system enables continuous, objective monitoring, making mental health assessment more accessible and proactive. Advances in technology for mental health are motivated by the growing prevalence of mental health challenges, limited access to traditional care, and the need for discreet, cost-effective, and scalable solutions. Real-time monitoring through wearable, AI, and mobile apps enables early detection, personalized care, and self-management. These tools empower individuals, reduce stigma, and address global care gaps. Integration with innovations like virtual reality (VR) and block chain enhances therapy and data security.

Acute mental health episodes, such as panic attacks or severe anxiety, often occur unpredictably, making timely intervention difficult. Current methods lack the capability for real-time physiological tracking, leading to delays in diagnosis and treatment. This work contributes to the field by integrating emerging technologies such as machine learning, IoT-enabled sensors, and secure data management frameworks to ensure reliability, privacy, and scalability [2]. By bridging the gap between technology and mental healthcare, this system offers a scalable and proactive solution for early mental health assessment and intervention. The insights derived from this system can facilitate early interventions, personalized recommendations, and improved mental health management, ultimately helping individuals take control of their mental well-being. The main contributions of this work are as follows.

- We analyze breathing patterns to detect stress and anxiety levels using machine learning.
- We assess mental health through a structured questionnaire, mapping responses to psychological conditions.
- We track body temperature variations to identify physiological stress indicators.
- We use computer vision to analyze facial expressions and detect signs of mental distress.
- We evaluate model performance using accuracy, F1-score, AUC-ROC, precision and recall.
- We develop a user-friendly interface for automated mental health assessment.

The structure of this paper is as follows: Section 2 examines relevant studies and identifies gaps in the body of knowledge. Section 3 details the system architecture, including data collection, pre-processing, implementation, and model development. Section 4 presents the proposed algorithms, including feature extraction and classification models. Section 5 outlines the experimental setup and discusses the performance evaluation metrics. Section 6 provides an in-

depth analysis of the results, comparing different models and assessing the system's effectiveness. Finally, Section 7 concludes the paper by summarizing contributions and outlining potential future research directions.

2. RELATED WORK

This part reviews previous research in the topic, with an emphasis on pertinent studies and developments that serve as the framework for our investigation. The field of mental health monitoring has seen significant advancements due to the integration of artificial intelligence (AI), machine learning (ML), and sensor-based technologies. Traditional mental health assessments, which rely on clinical interviews and self-reported questionnaires, often suffer from subjectivity, infrequency, and accessibility limitations. To overcome these challenges, researchers have explored contactless sensing methods, predictive models, and AI-driven chatbots to provide real-time, non-invasive, and scalable solutions for mental health assessment and intervention.

Rahman et al. [2] provide an extensive review of contactless sensing methods for mental health assessment. It explores a wide range of published studies that focus on using these methods to predict and monitor mental health conditions. The applications of contactless sensing are categorized into three main areas: detection, recognition, and continuous monitoring of vital signs. A detailed comparison of recent studies is also presented, highlighting the effectiveness, reliability, and scope of these methods [3].

Mughal et al. [4], explored 19 different predictive models to analyse the dataset, with the gradient Boosting Regressor emerging as the top performing model. The best results achieved an R-squared value of 0.32, suggesting moderate predictive capability. However, most models' suboptimal performance, which may be attributed to limited sample size used in the study. This constraint raises concerns about potential over fitting and limits the generalizability of the finding.

Mahammad et al. [5] employed classification techniques, including logistic Regression, Bernoulli Naïve Bayes, K-Nearest Neighbours, Random Forest, and Decision Tree algorithms. Among these, Logistic Regression demonstrated the highest performance, achieving an accuracy of 93 percent.

Kamoji et al. [6] presented an innovative predictive framework that combines ensemble learning techniques with large language models (LLMs). The first phase involves using ensemble methods like AdaBoost, voting, and bagging to build a reliable prediction model. Among these, Random Forest proves to be the most effective. In the second phase, the framework incorporates a Large Language Model (LLM) to enhance prediction capabilities. The process begins with user input, which is analyzed using the Random Forest model to predict potential mental health concerns [7]. This prediction is then sent to a Google Gemini model via an API, which provides personalized insights and recommendations. By integrating ensemble learning with advanced language models, this framework offers a more comprehensive and user-friendly approach to identifying and addressing mental health issues. This combination not only improves prediction accuracy but also enhances the overall experience for users seeking mental health support [8].

Zhang et al. [9] developed an AI-driven mental health chatbot named Sakhi. This chatbot is designed to provide

personalized emotional support, offer tailored mental health advice, and facilitate remote monitoring of mental well-being. It leverages advanced natural language processing and machine learning algorithms to engage with users empathetically, delivering actionable insights and fostering a supportive environment. Janartanan et al. [10] presented a detection of mental health at early stages.

Kavyashree and Usha [11] explored mental health issues and introduce a personalized solution in the form of a medical chatbot aimed at supporting individuals seeking mental health assistance. The proposed chatbot is designed to address mental health and well-being by utilizing artificial intelligence (AI) and natural language processing (NLP) techniques.

Despite these advancements, certain research gaps remain. Many existing AI-driven mental health models rely on self-reported data, which can introduce biases and inaccuracies. The sample sizes in several studies have been limited, affecting the generalizability of the models. Additionally, while chatbots and predictive models provide valuable insights, they often lack real-time physiological monitoring, making it difficult to detect acute mental health episodes such as panic attacks or sudden anxiety spikes [12-14]. Current AI models also require better personalization and integration with multimodal data sources to improve accuracy and reliability [15, 16].

Sheikh et al. [17] presented that mental health and smartphones and connected devices, such as smart watches, can be used to monitor behaviour.

Wang et al. [18] presented a Health monitoring embedded with intelligence is the demand of the day. In this era of a large population with the emergence of a variety of diseases, the demand for healthcare facilities is high. Yet there is a scarcity of medical experts, technicians for providing healthcare to the people affected with some medical problem. This article presents an Internet of Things (IoT) system architecture for health monitoring and how data analytics can be applied in the health sector [19-21].

To address these gaps, this study proposes a real-time, non-invasive mental health monitoring system that integrates facial expression recognition, respiration analysis, infrared temperature sensing, and psychological questionnaires. By leveraging AI-driven analysis and multimodal data fusion, the system enhances accuracy in detecting stress, anxiety, and emotional fluctuations. The proposed approach eliminates the subjectivity of self-reported assessments and provides continuous tracking, making mental health monitoring more accessible and proactive. The integration of secure data management frameworks ensures reliability and privacy, addressing concerns about ethical considerations in AI-driven mental health assessments.

The primary research objectives of this study include early detection and diagnosis of mental health conditions using physiological and behavioural signals, continuous monitoring of mental well-being beyond periodic assessments, and personalized interventions tailored to individual needs. Additionally, this research aims to reduce the stigma associated with seeking mental health support by making monitoring more discreet and non-intrusive. By improving accessibility, this system can provide mental health monitoring solutions to individuals in remote areas or those with limited access to traditional healthcare services. Through this approach, this study aims to bridge the gap between technology and mental healthcare, enabling proactive and effective mental health management.

Building on these existing approaches, our system adopts a multi-modal framework to enhance the accuracy of mental health assessment.

3. SYSTEM ARCHITECTURE

In this section, we are describing the proposed system architecture. The layered system architecture of the proposed Mental Health Monitoring (MHM) system is designed to provide real-time, non-invasive, and multimodal assessment of an individual's emotional and physiological well-being. The architecture integrates multiple data sources, including facial expression analysis, respiration monitoring, temperature sensing, and psychological questionnaires, processed using machine learning algorithms for comprehensive mental health evaluation.

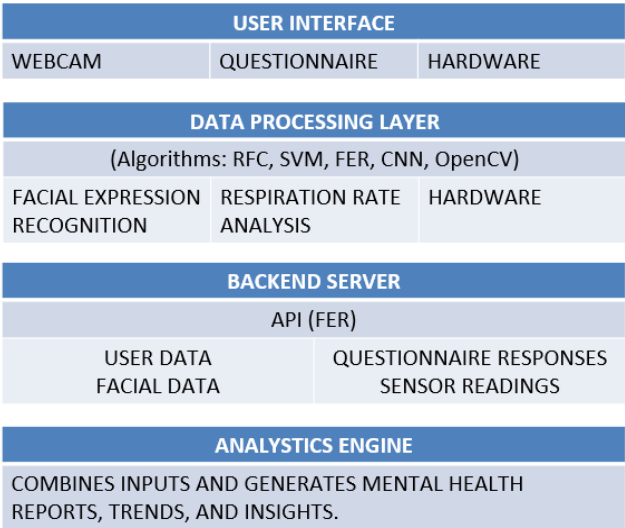


Figure 1. Level architecture diagram for MHM system

The system architecture diagram in Figure 1, illustrates how different components interact to collect, process, and analyze user data. At the core of the system is the central processing unit, which integrates data from various sensors and processes it using AI models. The system consists of a camera for facial expression recognition, sensors for respiration and body temperature monitoring, and an interactive questionnaire module for subjective assessments. Data flows through preprocessing stages before being analyzed by machine learning models, with results displayed through a user interface. The modular design of the architecture ensures scalability, allowing additional features to be integrated seamlessly in the future.

The system uses a combination of machine learning and deep learning algorithms for data analysis. The facial expression recognition module employs a CNN trained on a facial emotion recognition dataset to classify emotions such as happiness, sadness, anger, and stress. The respiration monitoring module uses signal processing techniques such as Fourier Transform for frequency-domain analysis and noise reduction to extract meaningful breathing patterns. A Random Forest classifier is used to classify breathing irregularities associated with anxiety. The questionnaire module utilizes a Logistic Regression model for analyzing user responses and identifying potential mental health risks. The final prediction is generated using a weighted decision-making algorithm that

integrates outputs from all modules.

3.1 Flowchart

The flowchart of the system in Figure 2 outlines the sequence of operations from data collection to prediction output. It starts with user input, where facial expressions, respiration patterns, and body temperature are captured. These inputs are processed by corresponding modules: the facial expression recognition module extracts emotion-based features using convolutional neural networks (CNNs), the

respiration monitoring module analyzes breathing patterns, and the temperature sensor measures stress-related physiological changes. The questionnaire module collects self-reported mental health indicators. Once the data is gathered, it is preprocessed to remove noise, normalized for consistency, and fed into machine learning models for analysis. The decision-making module aggregates the outputs from different sources to generate a final prediction of the user's mental health status. The results are then displayed to the user, with recommendations for further actions based on the detected emotional and physiological states.

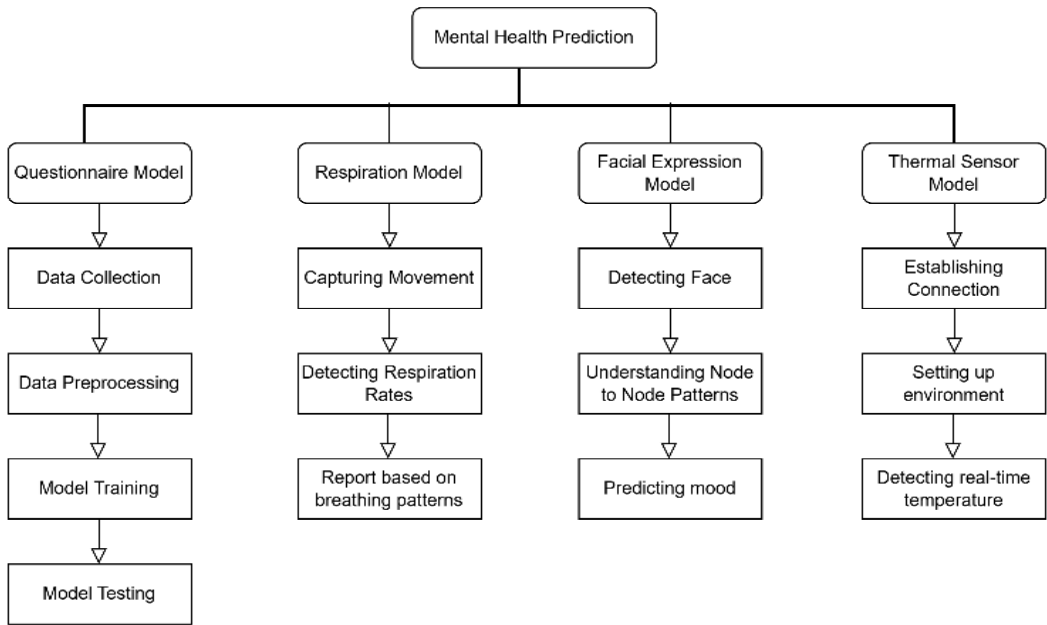


Figure 2. Flow diagram for MHM

Figure 2 is a Data flow diagram (DFD) which represents the flow of information within the Mental Health Monitoring (MHM) system, outlining how data moves through its various modules. It begins with input data collection from sensors (e.g., thermal sensor, respiration device) and user responses to the questionnaire. This data is processed by respective modules, such as facial expression analysis and physiological monitoring, before being integrated into a centralized processing unit. The system then applies machine learning algorithms to analyze and synthesize the data, generating a final prediction and insights. The DFD provides a clear, visual representation of the system's workflow, highlighting the interaction between hardware components, processing modules, and output generation. Hence, with the system architecture established, we now move on to the proposed system, detailing its components and functionality.

4. PROPOSED SYSTEM

The proposed system integrates multiple modalities, including respiration analysis, questionnaire evaluation, temperature monitoring, and facial expression recognition, to provide a comprehensive mental health monitoring (MHM).
Respiration Module – This module tracks the user’s breathing patterns using a sensor to detect irregularities linked to stress or anxiety. It analyses respiration rate variations and provides real-time feedback to help in relaxation or early stress

detection.
Questionnaire Module – This module consists of a set of psychological and behavioural questions to assess the user’s mental state. It uses predefined scoring methods to classify stress, anxiety, or depression levels. The responses help in determining further monitoring needs.
Facial Recognition Module – This module uses computer vision and machine learning to analyze facial expressions and detect emotional states. It identifies stress, sadness, or happiness based on facial landmarks and muscle movements, contributing to a comprehensive mental health assessment.
Body Temperature Sensor (Arduino) – This module utilizes an Arduino-based temperature sensor to monitor fluctuations in body temperature. Temperature changes can indicate stress responses, helping in correlation with other physiological and behavioural data for better mental health evaluation.
The Behavioural Data Analysis module utilizes a Random Forest Classifier to predict mental health conditions based on self-reported psychological questionnaire data. The Random Forest algorithm, an ensemble learning method, builds multiple decision trees and averages their outputs to reduce overfitting and improve predictive accuracy. The dataset used consists of structured mental health questionnaire responses, which are pre-processed by encoding categorical data, scaling numerical values, and removing inconsistencies before being fed into the model. The classifier detects patterns that correlate with mental health risks, enabling personalized mental health assessments. This component ensures that users receive early

and tailored interventions without requiring frequent clinical visits, enhancing accessibility for individuals in remote or resource-limited areas. Alternative algorithms, such as Logistic Regression, Support Vector Machines (SVM), and Gradient Boosting Models (GBM), can also be utilized for classification, depending on data complexity.

Contactless Respiratory Monitoring detects breathing irregularities linked to stress and anxiety, offering a passive and non-intrusive alternative to traditional assessments. This module utilizes infrared sensors or microphone-based breath detection systems to capture real-time respiratory signals. Signal processing techniques such as Fourier Transform and Wavelet Transform convert time-domain signals into frequency-domain representations, allowing the detection of irregular breathing patterns associated with heightened emotional states. The dataset for respiratory monitoring comprises pre-recorded breathing patterns from individuals with varying levels of stress and anxiety. A Random Forest model is used to classify irregularities, while alternatives like Long Short-Term Memory (LSTM) networks and Hidden Markov Models (HMMs) can be employed for time-series prediction. Pre-processing techniques such as band-pass filtering, noise removal, and normalization ensure high-quality respiratory data for analysis. This feature is particularly beneficial for individuals prone to panic attacks, PTSD, or high-stress situations, enabling early detection and real-time intervention.

Facial Expression Recognition leverages Convolution Neural Networks (CNNs) to analyse facial features and classify emotional states such as happiness, sadness, anger, or distress [9]. The system processes real-time facial images using OpenCV for face detection, extracts facial landmarks, and applies deep learning models for classification. The dataset used includes publicly available facial emotion datasets such as FER-2013 and AffectNet, which contain labelled images of various emotional expressions. Image pre-processing techniques such as gray scale conversion, histogram equalization, and image augmentation improve feature extraction and model accuracy. This technology is especially valuable in today's digital landscape, where virtual interactions and telemedicine have become commonplace. By integrating facial emotion recognition into teletherapy platforms and smart home assistants, individuals can receive instant feedback on their emotional well-being, leading to better self-awareness and proactive mental health management. Alternative techniques such as Principal Component Analysis (PCA) for dimensionality reduction and Support Vector Machines (SVMs) for classification can also enhance performance.

Body Temperature Sensing provides another objective physiological marker for mental health assessment. Stress-induced body temperature fluctuations have been widely studied, with increases in temperature often linked to emotional arousal and anxiety. The system utilizes an infrared temperature sensor to measure subtle temperature variations from the forehead, wrist, or other exposed skin areas. The dataset includes temperature variations recorded under different emotional conditions, allowing the model to correlate temperature fluctuations with mental states. A Decision Tree-based model classifies temperature anomalies, while alternative models such as k-Nearest Neighbours (k-NN) and Artificial Neural Networks (ANNs) can be employed for pattern recognition. Data pre-processing steps include temperature normalization, outlier removal, and feature

scaling to ensure accuracy. This feature is particularly useful in scenarios where verbal communication is limited, such as monitoring mental health in elderly individuals, non-verbal patients, and young children.

The integration of these four core components ensures that the proposed system provides a holistic and multimodal assessment of mental health, bridging the gap between technology and healthcare. Unlike conventional methods that rely solely on subjective responses, this system continuously tracks physiological and behavioural indicators, allowing for early detection of mental distress. This is especially useful for individuals who may not recognize their own mental health deterioration or those hesitant to seek professional help due to stigma or accessibility barriers.

In real-world applications, this system can be integrated into workplaces, educational institutions, and telemedicine platforms to promote mental well-being, reduce workplace stress, and enhance productivity. By offering a scalable, cost-effective, and widely deployable solution, it enables real-time intervention and data-driven insights for mental health management. The use of secure data management frameworks ensures privacy and ethical compliance, making the system reliable and trustworthy. Additionally, continuous data collection can provide researchers and healthcare professionals with valuable insights into mental health trends, enabling better treatment strategies and policy development.

By eliminating the subjectivity of traditional assessments, providing real-time physiological monitoring, and enhancing accessibility, this system represents a significant step forward in proactive mental healthcare. Through the fusion of AI, IoT-enabled sensors, and multimodal data analysis, individuals are empowered to take control of their mental well-being, leading to healthier, more resilient communities. With the proposed system defined, we now proceed to the experimental setup, detailing the data collection, model training, and evaluation process.

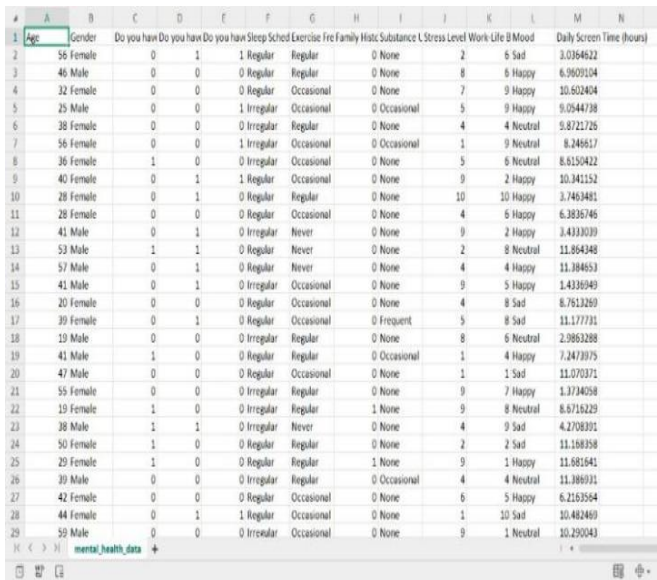
5. EXPERIMENTAL SETUP

The MHM is designed to address the growing need for continuous, real-time, and non-invasive mental health assessments. Traditional systems often rely on self-reported data or periodic clinical evaluations, limiting their ability to provide timely insights. This system integrates facial expression analysis, respiration monitoring, temperature detection, and subjective feedback through questionnaires to create a holistic framework for mental health evaluation. The goal is to combine emotional and physiological data with subjective insights to ensure accurate, actionable results that support early intervention. Key challenges include achieving seamless integration of hardware and software, ensuring data accuracy, and optimizing real-time performance. By leveraging advanced technologies and readily available datasets, the system bridges the gap between subjective and objective assessments, making mental health monitoring more accessible and reliable.

5.1 Questionnaire module

The system must provide a dynamic questionnaire to assess mental health. It must include a scoring mechanism to determine depression levels based on user responses. Results must be displayed to the user with an appropriate

recommendation (e.g., mild, moderate, or severe depression). The questionnaire must support multiple languages (if applicable). Responses must be stored securely for analysis or future reference.



#	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Age	Gender	Do you have	Do you have	Do you have	Sleep Sched	Exercise	Free Family	Hisic	Substance	L Stress Level	Work Life	B Mood	Daily Screen Time (hours)
2	56	Female	0	1	1	Regular	Regular	0	None	2	6	Sad	3.0364622	
3	46	Male	0	0	0	Regular	Regular	0	None	8	6	Happy	6.96070104	
4	32	Female	0	0	0	Regular	Occasional	0	None	7	9	Happy	10.602404	
5	25	Male	0	0	1	Irregular	Occasional	0	Occasional	5	9	Happy	9.0544738	
6	38	Female	0	0	0	Irregular	Regular	0	None	4	4	Neutral	8.87272725	
7	56	Female	0	0	1	Irregular	Occasional	0	Occasional	1	9	Neutral	8.248517	
8	36	Female	1	0	0	Irregular	Occasional	0	None	5	6	Neutral	8.6150422	
9	40	Female	0	1	1	Regular	Occasional	0	None	9	2	Happy	10.341152	
10	28	Female	0	1	0	Regular	Regular	0	None	10	10	Happy	3.7463481	
11	28	Female	0	0	0	Regular	Occasional	0	None	4	6	Happy	6.3836746	
12	41	Male	0	1	0	Irregular	Never	0	None	9	2	Happy	3.4333039	
13	53	Male	1	1	0	Regular	Never	0	None	2	8	Neutral	11.864348	
14	57	Male	0	1	0	Regular	Never	0	None	4	4	Happy	11.384653	
15	41	Male	0	1	0	Irregular	Occasional	0	None	9	5	Happy	1.4338091	
16	20	Female	0	0	0	Regular	Occasional	0	None	4	8	Sad	6.7613269	
17	39	Female	0	1	0	Regular	Occasional	0	Frequent	5	8	Sad	11.17731	
18	19	Male	0	0	0	Irregular	Regular	0	None	8	6	Neutral	2.985288	
19	41	Male	1	0	0	Regular	Regular	0	Occasional	1	4	Happy	7.3473975	
20	47	Male	0	0	0	Regular	Occasional	0	None	1	1	Sad	11.070371	
21	55	Female	0	0	0	Irregular	Regular	0	None	9	7	Happy	1.3734058	
22	19	Female	1	0	0	Regular	Regular	1	None	9	8	Neutral	6.5716229	
23	38	Male	1	1	0	Irregular	Never	0	None	4	9	Sad	4.270801	
24	50	Female	1	0	0	Regular	Regular	0	None	2	2	Sad	11.168358	
25	29	Female	1	0	0	Regular	Regular	1	None	9	1	Happy	11.681541	
26	39	Male	0	0	0	Irregular	Regular	0	Occasional	4	4	Neutral	11.388931	
27	42	Female	0	0	0	Regular	Occasional	0	None	6	5	Happy	6.2163584	
28	44	Female	0	1	1	Regular	Occasional	0	None	1	10	Sad	10.482460	
29	59	Male	0	0	0	Irregular	Occasional	0	None	9	1	Neutral	10.290343	

Figure 3. Data set for questionnaire module

The hardware and software requirements of the module are as follows.

The questionnaire module uses data from Figure 3 and runs on tablets, smartphones, or computers with a Full HD resolution (1920x1080) for clear question display. It requires minimal storage (1 GB) for locally saving responses if not cloud-based. The software includes Python for system integration, with data processing handled by Pandas and NumPy.

Data-Source:
<https://www.kaggle.com/code/imtkaggleteam/mental-health-eda-prediction>

5.2 Respiration monitoring

The system must capture real-time respiration data using a sensor. Data should be processed to identify abnormalities in breathing patterns. Alerts must be triggered if respiration rates fall outside the normal range. The system must store respiration data for visualization and reporting. Support for exporting data in standard formats (e.g., CSV, PDF).

The hardware and software requirements of the module are as follows.

This module utilizes a piezoelectric or IR-based sensor with ± 0.1 breaths per minute accuracy. It features embedded memory (256 KB) for temporary data buffering and supports Bluetooth 5.0, Wi-Fi, or USB for connectivity. The software integrates respiratory monitoring tools, including camera-based or radar-based solutions for analyzing breathing patterns.

5.3 Body temperature sensor

The system must measure body temperature accurately using a sensor. It must classify temperature readings as normal, low, or high based on standard medical guidelines. Temperature data must be logged with timestamps for trend

analysis.
Provide alerts if body temperature exceeds critical thresholds. Allow users to view historical data via a graphical interface.

The hardware and software requirements of the module are as follows.

A digital IR-based or thermistor-based sensor ensures ± 0.1°C accuracy with a 0.01°C resolution. It includes a built-in memory (64 KB) for buffering data and supports USB, Bluetooth, or direct microcontroller connections. The software is programmed using Arduino IDE, enabling seamless integration with microcontrollers.

5.4 Facial expression recognition

The system must capture and process facial images in real-time or from uploaded photos. It must analyze facial expressions to detect emotions such as happiness, sadness, anger, or stress. Results must be displayed with confidence levels for each detected emotion. The module must support integration with other data (e.g., depression questionnaire results) for comprehensive analysis. The system must ensure data privacy and avoid storing images unnecessarily unless required.

Objective: An image used to identify emotional states based on facial expressions.

Implementation: A camera captures real-time facial data, which is processed using OpenCV and pre-trained models for emotion recognition.

The hardware and software requirements of the module are as follows.

This module employs a camera with at least 1280x720 resolution and 30 FPS for real-time analysis. An NVIDIA Jetson Nano (or equivalent) handles processing with 4 GB RAM and 16 GB storage. It supports USB 3.0 and Wi-Fi for data transmission. The software uses OpenCV for image processing and machine learning libraries like Scikit-learn for emotion detection.

5.4.1 Respiration detection

Objective: image used to monitor breathing patterns, a critical indicator of mental well-being.

Implementation: The system uses radar-based sensors or microphones to detect respiration rates and irregularities. Data is processed to identify shallow or rapid breathing patterns that may indicate stress or anxiety.

5.4.2 Temperature detection

Objective: It is used to track body temperature as a physiological indicator of stress or discomfort.

Implementation: An MLX90614 infrared sensor, controlled by the Arduino Uno, measures body temperature without contact. Temperature variations are correlated with other data streams to assess stress levels.

5.4.3 Questionnaire module

Objective: It is used to collect subjective feedback on the user’s mental health.

Implementation: The system prompts the user to complete a questionnaire designed with psychological parameters. The responses are analyzed alongside physiological data for a holistic evaluation.

The findings of integrating multiple models (Facial Expression Analysis, Respiration Monitoring, and

Questionnaire-based Prediction) to detect mental health conditions are detailed in this section. The dataset used for training the models was curated using a combination of publicly available resources and synthetic data generation techniques, improving both size and quality. Each model utilized unique preprocessing and data enhancement methods to ensure accuracy and robustness in their specific domains.

5.4.4 Facial expression analysis

For the Facial Expression Analysis Model, preprocessing involved facial landmark detection and normalization, while feature extraction was performed using Convolution Neural Networks (CNNs). The model was trained using the Adam optimizer with an initial learning rate of 0.001, a minimum learning rate of 0.0001, and a batch size of 64. The categorical cross entropy loss function was applied, and the Soft max classifier was utilized for emotion classification.

The Respiration Monitoring Model involved signal preprocessing techniques, including noise reduction and normalization, followed by feature extraction using statistical and frequency domain measures. A Random Forest model was trained with hyper-parameters optimized using Grid Search, yielding improved performance. The training data was split with 80% for training and 20% for validation as seen in Figure 4.

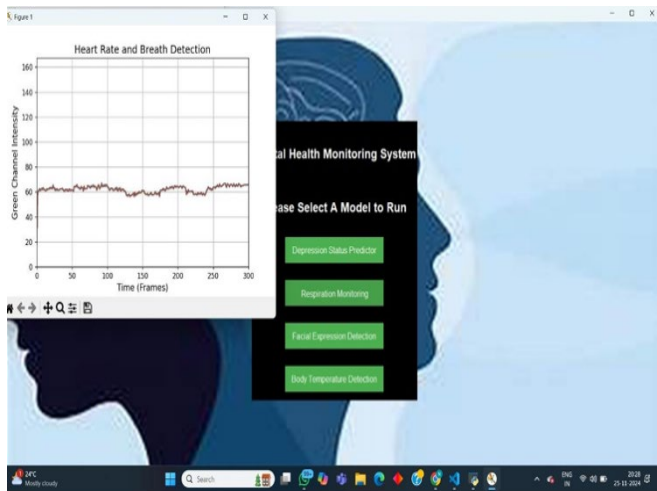


Figure 4. Screenshot of respiration module

For the Questionnaire-based Model, feature engineering included encoding categorical variables and scaling numerical data. Logistic Regression was used as the classifier, optimized with a learning rate of 0.01 and regularization parameters tuned to prevent over fitting. All three models were integrated into a cohesive system, enabling simultaneous evaluation of multiple parameters. Data was passed through the respective models, and their outputs were aggregated using a weighted decision-making mechanism to determine the final mental health score.

5.5 Clinical relevance

Our system combines four models—respiration monitoring, body temperature detection, facial expression recognition, and a Questionnaire-based classifier—to provide a multimodal approach to mental health assessment. Each modality adds value: breathing irregularities are linked to anxiety, temperature fluctuations may indicate stress, facial

expressions capture non-verbal cues, and questionnaires validate findings through self-reports. Together, they create a more reliable framework than single-model approaches.

•Limitations and Mitigation: In real-world deployment, respiration monitoring may be affected by noise or movement, body temperature by environmental conditions, and facial recognition by poor lighting or occlusions. Questionnaires remain prone to subjectivity. These challenges can be mitigated through filtering and sensor fusion (respiration), baseline calibration (temperature), diverse datasets with preprocessing (facial recognition), and adaptive questioning with sensor integration (questionnaire).

•Ethical Considerations: The system handles sensitive health data; hence informed consent, data privacy, and protection against misuse are essential.

•Applications and Future Work: This multimodal system can serve as an early screening tool in universities, workplaces, or telemedicine. Future work will focus on larger and more diverse datasets, mobile or wearable integration, and advanced deep learning models to capture temporal variations in signals.

6. RESULT ANALYSIS

The results analysis evaluates the performance of our models based on key metrics such as accuracy, F1-score, precision and recall.

The models were trained and validated on separate datasets to ensure generalization. For the Facial Expression Analysis Model, accuracy improved steadily over training epochs, as depicted in Figure 5. The loss curves showed a decline in both training and validation loss, confirming the effectiveness of the chosen hyper-parameters. The Respiration Monitoring Model achieved high accuracy and robustness against noisy data due to the preprocessing steps and Random Forest's ensemble learning capability. The feature importance analysis highlighted key respiratory metrics that contributed significantly to predictions.

For the Questionnaire-based Model, the precision, recall, and F1-score demonstrated its ability to distinguish between different mental health conditions effectively. Validation accuracy was consistent, confirming the model's stability.

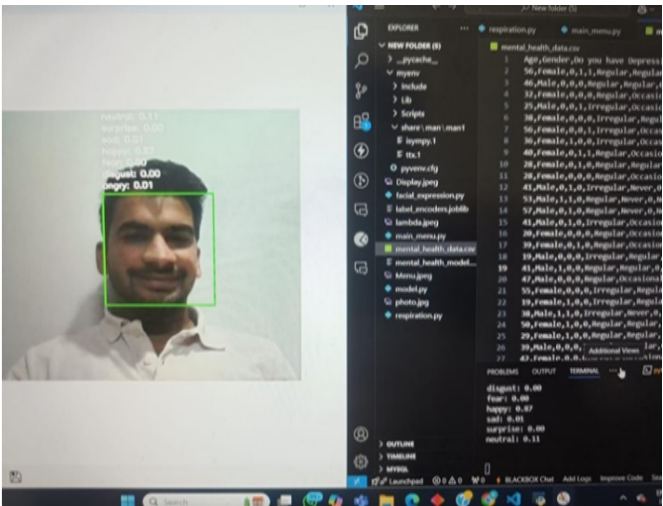


Figure 5. Screenshot of facial recognition module

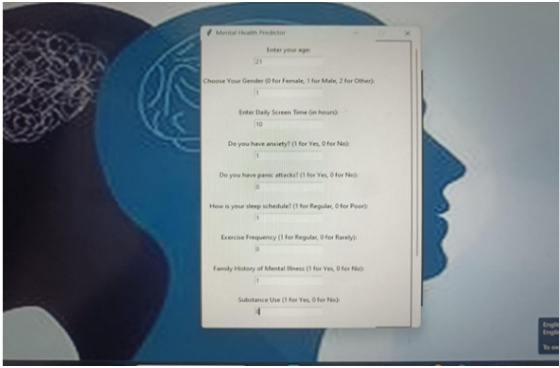


Figure 6. Screenshot of questionnaire model

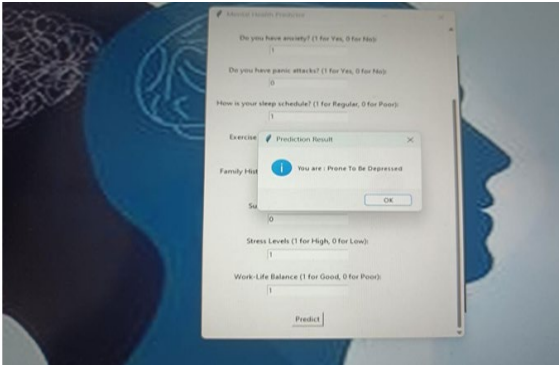


Figure 7. Screenshot of the final prediction

6.1 Final system outcomes

When tested on real-world data, the integrated system

displayed high accuracy and reliability. The Facial Expression Model accurately classified emotions like anxiety and sadness, the Respiration Monitoring Model identified abnormal breathing patterns, and the Questionnaire Model added context with subjective responses. Together, these models offered a comprehensive mental health assessment. By using advanced techniques, optimized hyper-parameters, and robust validation processes, the proposed system provides a practical and scalable solution for early mental health detection.

Figure 5 illustrates the working of the respiration monitoring model which traces the change in the skin color due to blood flow on the forehead area of the person through a web cam and shows the breathing pattern of the person which aids with detection of the mental state.

Figure 6 shows how the emotions of a person are detected using the facial expression recognition module, which utilizes CNN, which will help in the analysis of how the person is feeling and what is his/her emotional state currently.

Figure 7 analyzes if a person is prone to depression / is under depression / not depressed by assessing the answers provided by the person to the questions specified in the model. Multiple questions regarding sleep patterns, work hours, substance consumption, and other necessary information are asked, which could potentially influence a person’s mental well-being.

6.2 Comparison with pre-trained models

When compared to pre-trained models, our customized model achieved competitive results. Table 1 presents the comparison of the performance of this multi-modal approach with other approaches. Table 2 summarizes the performance metrics of the proposed model and the pre-trained models.

Table 1. Performance comparison between unimodal and the proposed multimodal approach

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Facial Expression Model	81	78	80	79
Respiration Monitoring	84	82	81	81.5
Questionnaire-Based Model	93	90	92	91
Body Temperature Model	76	74	73	73.5
Proposed Multimodal Fusion	96	94	95	94.5

Table 2. Comparison table

Model	Batch Size	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
Facial Expression Model (CNN)	64	High	Stable	Decreasing	Decreasing
Respiration Monitoring (Random Forest)	80-20 split	High	Robust	Low	Low
Questionnaire-Based Model	-	High	Consistent	-	-

Using the Facial Expression Model, Accuracy improves steadily over epochs and loss curves show a decline in both training and validation, confirming effective hyper-parameters.

Respiration Monitoring Model shows High accuracy and robustness against noisy data. On using the Questionnaire-Based Model, Precision, recall, and F1-score metrics validate its ability to distinguish mental health conditions effectively. The performance of the proposed modified model on the test set is displayed in the confusion matrix below, which contrasts the true labels with the predicted labels.

The mlx90614 infrared temperature sensor is used to measure body temperature without direct contact. This non-invasive method captures accurate thermal readings, which are correlated with physiological responses to stress or emotional

changes. By integrating these measurements into the system, the module helps in identifying early signs of mental health conditions such as anxiety or stress, supporting timely interventions. The performance of the proposed modified model on the test set is displayed in the confusion matrix, which contrasts the true labels with the predicted labels. Figures 8-10 provide Confusion matrix for facial expression, respiration and questionnaire models respectively.

The confusion matrix for the Facial Expression Analysis Model (CNN) was used to evaluate classification accuracy for different emotional states. The model showed high true positive rates for well-defined emotions like happy, sad, and angry. Some misclassifications between neutral and sad expressions, likely due to similar facial features.

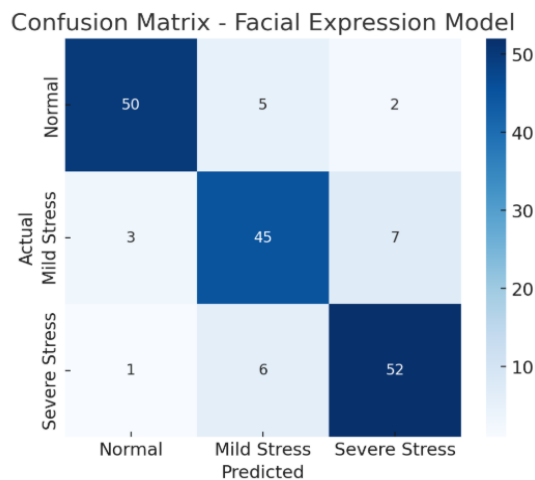


Figure 8. Confusion matrix for facial expression model

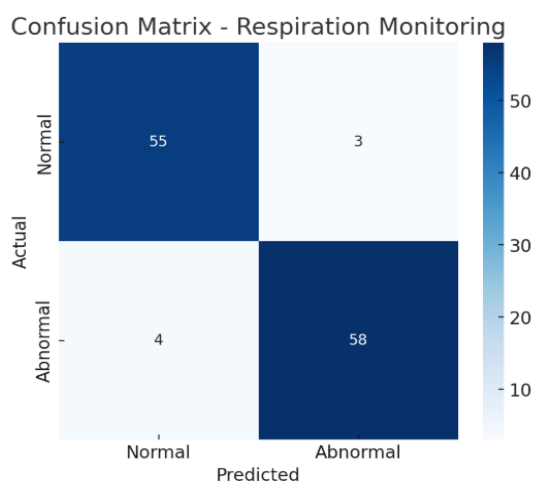


Figure 9. Confusion matrix for respiration model

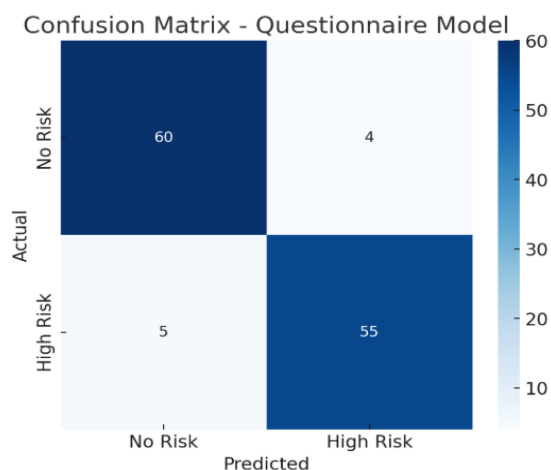


Figure 10. Confusion matrix for questionnaire model

The Respiration Monitoring Model (Random Forest) confusion matrix confirmed high classification accuracy for normal vs. abnormal breathing patterns. Some false positives in cases where respiration rate variations overlapped between classes.

The Questionnaire-Based Model (Logistic Regression) confusion matrix demonstrated good recall and precision, with 93% accuracy in classifying mental health conditions (H). The training curves illustrate how model accuracy and loss evolved

over epochs:

Facial Expression Model (CNN): Accuracy Curve: Shows a steady increase over epochs, confirming effective learning. Loss Curve: Training loss and validation loss decrease progressively, indicating successful convergence. **Respiration Monitoring Model (Random Forest):** The validation accuracy remained high and stable, reinforcing the model's robustness to noisy data. Time comparison of all 4 models can be seen in Figure 11.

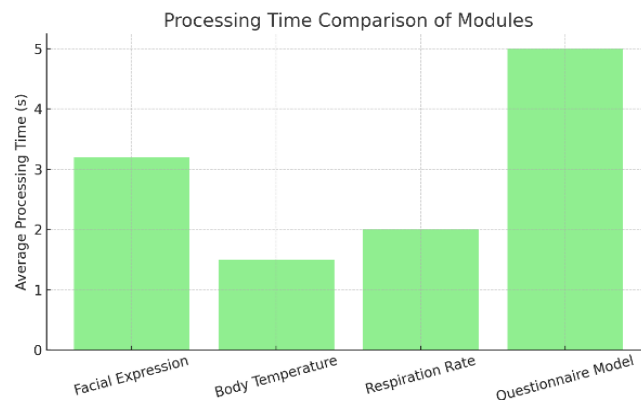


Figure 11. Comparison of all the 4 models

Questionnaire-Based Model (Logistic Regression): Precision, recall, and F1-score remained consistent, validating the stability of the model as depicted in Figures 12-14.

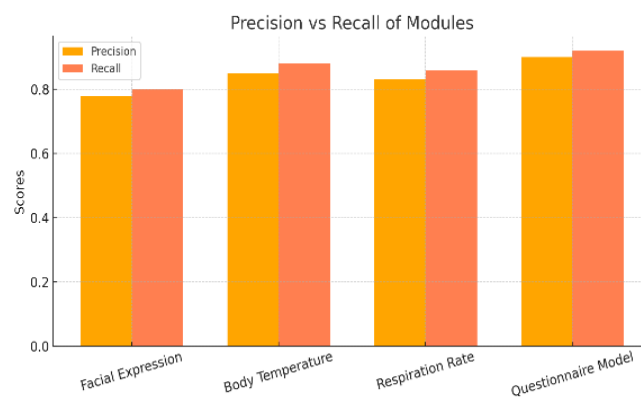


Figure 12. Training and validation accuracy for questionnaire model

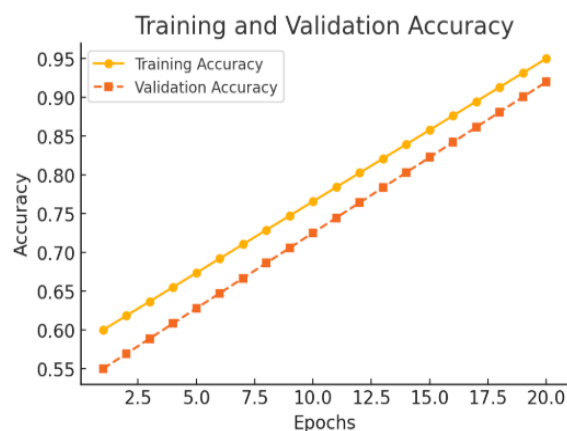


Figure 13. Epoch for all the models

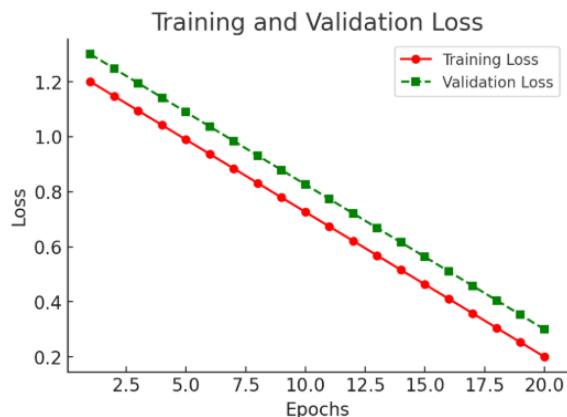


Figure 14. Comparison of precision vs. recall scores of all four modules

Table 3. Best models

Metric	Best Model	Score
Accuracy	Questionnaire Model	0.08
Speed	Body Temperature Model	1.5 sec
Precision & Recall	Questionnaire Model	Precision: 90% Recall: 92%
User Comfort	Questionnaire Model	0.07

Table 3 presents the best models based on accuracy, speed, precision, recall, and user comfort. where the Questionnaire Model outperformed others with an impressive 92% accuracy, making it the most reliable model for correct predictions. The Body Temperature model demonstrated the fastest response time of 1.5 seconds, making it ideal for scenarios requiring quick health assessments. The Questionnaire Model again shows strong performance with a precision of 90% and recall of 92%, indicating both high correctness and completeness in its predictions. The Questionnaire Model scores the highest with a 9.0 out of 10 under user comfort, suggesting that users found it the most comfortable and non-intrusive method.

7. CONCLUSIONS

The proposed mental health monitoring system aims to assess an individual's mental well-being through a multi-module framework. It integrates four key modules: a questionnaire for self-assessment, a respiration monitoring system, facial recognition for emotion detection, and a body temperature sensor using Arduino. The framework combines hardware and software components, enabling real-time data collection and analysis. The proposed method utilizes machine learning algorithms to interpret facial expressions and vital signs, providing early detection of stress and anxiety levels. The system successfully delivers preliminary results with reasonable accuracy, demonstrating its potential as a supportive tool in mental health monitoring. However, the system has certain limitations, such as the need for a stable environment for accurate facial recognition and dependency on sensor calibration for precise data collection. Additionally, the dataset size and diversity can impact the prediction accuracy. In the future, enhancements could include integrating wearable devices for continuous tracking, developing a user-friendly mobile application, improving

machine learning models with larger datasets, and incorporating voice analysis for better mental health assessment.

REFERENCES

- [1] Liu, J., Chen, Y., Wang, Y., Chen, X., Cheng, J., Yang, J. (2018). Monitoring vital signs and postures during sleep using WiFi signals. *IEEE Internet of Things Journal*, 5(3): 2071-2084. <https://doi.org/10.1109/IIOT.2018.2822818>
- [2] Rahman, M., NaghshvarianJahromi, M., Mirjavadi, S.S., Hamouda, A.M. (2018). Resonator based switching technique between ultra wide band (UWB) and single/dual continuously tunable-notch behaviors in UWB radar for wireless vital signs monitoring. *Sensors*, 18(10): 3330. <https://doi.org/10.3390/s18103330>
- [3] Rajanna, S., Jayaramaiah, C., Sridhar, R., Chandrappa, P.H., Venkatesh, R.T. (2023). Fuzzy inference with enhanced convolutional neural network based classification framework for predicting heart attack using sensor data. *Revue d'Intelligence Artificielle*, 37(1): 93-99. <https://doi.org/10.18280/ria.370112>
- [4] Mughal, F., Raffae, W., Stubbs, P., Garcia, J. (2022). Towards depression monitoring and prevention in older populations using smart wearables: quantitative Findings. In *2022 IEEE 10th International Conference on Serious Games and Applications for Health (SeGAH)*, Sydney, Australia, pp. 1-8. <https://doi.org/10.1109/SEGAH54908.2022.9978305>
- [5] Mahammad, A.B., Kumar, R., Singh, R.K. (2024). Comparative analysis of Naïve Bayes, nearest neighbor, and logistic regression classifiers for the prediction of diabetes. In *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Tashkent, Uzbekistan, pp. 1492-1497. <https://doi.org/10.1109/ICTACS62700.2024.10840851>
- [6] Kamoji, S., Rozario, S., Almeida, S., Patil, S., Patankar, S., Pendhari, H. (2024). Mental health prediction using machine learning models and large language model. In *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*, Bangalore, India, pp. 185-190. <https://doi.org/10.1109/ICICI62254.2024.00040>
- [7] Ferrara, E. (2024). Large language models for wearable sensor-based human activity recognition, health monitoring, and behavioral modeling: A survey of early trends, datasets, and challenges. *Sensors*, 24(15): 5045. <https://doi.org/10.3390/s24155045>
- [8] Abdullah, M., Negied, N. (2024). Detection and prediction of future mental disorder from social media data using machine learning, ensemble learning, and large language models. *IEEE Access*, 12: 120553-120569. <https://doi.org/10.1109/ACCESS.2024.3406469>
- [9] Zhang, L., Cui, W., Zhao, L. (2025). Characteristic analysis of occurrence probability for vortex-induced vibration of long-span bridge based on health monitoring system. *Reliability Engineering & System Safety*, 265(B): 111516. <https://doi.org/10.1016/j.res.2025.111516>
- [10] Janartanan, V.D., Venkatesh, R.T., Rao, P.B.S.,

- Vaithilingam, S.D., Srinivasan, A. (2025). Stress and depression classification in social media using contextual knowledge attention based gated recurrent network. *International Journal of Intelligent Engineering & Systems*, 18(3): 420-431. <https://doi.org/10.22266/ijies2025.0430.29>
- [11] Kavyashree, N., Usha, J. (2023). MediBot: Healthcare assistant on mental health and well being. In 2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, pp. 1-5. <https://doi.org/10.1109/CSITSS60515.2023.10334083>
- [12] Sun, M., Li, P., Qin, H., Liu, N., et al. (2023). Liquid metal/CNTs hydrogel-based transparent strain sensor for wireless health monitoring of aquatic animals. *Chemical Engineering Journal*, 454: 140459. <https://doi.org/10.1016/j.cej.2022.140459>
- [13] Venkatnarayan, R. H., Page, G., Shahzad, M. (2018). Multi-user gesture recognition using WiFi. In Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services, Munich, Germany, pp. 401-413. <https://doi.org/10.1145/3210240.3210335>
- [14] Guo, L., Wang, L., Liu, J., Zhou, W., Lu, B. (2018). HuAc: Human activity recognition using crowdsourced WiFi signals and skeleton data. *Wireless Communications and Mobile Computing*, 2018(1): 6163475. <https://doi.org/10.1155/2018/6163475>
- [15] Wang, Y., Wu, K., Ni, L.M. (2016). Wifall: Device-free fall detection by wireless networks. *IEEE Transactions on Mobile Computing*, 16(2): 581-594. <https://doi.org/10.1109/TMC.2016.2557792>
- [16] Le, M. (2020). Heart rate extraction based on eigenvalues using UWB impulse radar remote sensing. *Sensors and Actuators A: Physical*, 303: 111689. <https://doi.org/10.1016/j.sna.2019.111689>
- [17] Sheikh, M., Qassem, M., Kyriacou, P.A. (2021). Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. *Frontiers in Digital Health*, 3: 662811. <https://doi.org/10.3389/fdgth.2021.662811>
- [18] Wang, T., Li, D., Li, B., Zhang, J. (2025). Lightweight structural health monitoring method for bridges based on cloud-edge collaborative optimization. *Structures*, 80: 109954. <https://doi.org/10.1016/j.istruc.2025.109954>
- [19] Dwivedi, R., Mittal, H., Agarwal, M., Dwivedi, S. (2022). Application of IoT in patient health monitoring system. In *Recent Advances in IoT and Blockchain Technology*, pp. 96-112. <https://doi.org/10.2174/97898150516051220401>
- [20] Dwivedi, R.K., Kumar, R., Buyya, R. (2022). Secure healthcare monitoring sensor cloud with attribute-based elliptical curve cryptography. In *Research Anthology on Securing Medical Systems and Records*, pp. 922-941. <https://doi.org/10.4018/978-1-6684-6311-6.ch044>
- [21] Vahdati, M., Gholizadeh HamlAbadi, K., Saghiri, A.M. (2020). IoT-Based healthcare monitoring using blockchain. In *Applications of Blockchain in Healthcare*, pp. 141-170. https://doi.org/10.1007/978-981-15-9547-9_6