PROJECT SYNOPSIS

Group ID: 10

Project Title: *"From Detection to Intervention: A Multimodal AI System for Mental Health Using Facial, Text, and Speech Analysis."*

---

1. Project Definition

Mental health issues such as depression and anxiety are becoming increasingly prevalent, affecting millions of people worldwide. Traditional methods for assessing mental health typically rely on self-reports or clinical interviews, which may not capture the full scope of an individual's mental and emotional state. These methods often lack real-time monitoring and do not integrate multiple data sources, limiting their effectiveness in early detection and intervention. To address these challenges, we propose the development of a multi-modal AI system that combines facial emotion detection, text analysis, and speech analysis to provide a comprehensive and accurate assessment of mental health conditions. The system will also deliver real-time feedback and personalized interventions to support mental well-being.[1]-[3]

2. Scope

1. Existing mental health assessment methods are limited in providing a comprehensive understanding of an individual's mental state.

2. The system will automate the integration of facial expressions, text inputs, and speech recordings to offer real-time analysis and intervention.

3. The proposed system aims to enhance accessibility and provide proactive detection and support for mental health issues.

3. Objectives

1. Develop a multimodal AI system that integrates facial emotion detection, text analysis, and speech analysis for comprehensive mental health detection.

2. Enable real-time detection of mental health.

3. Enhance accessibility to mental health support by making the system available on mobile platforms.

## 4. Review of Conference/Journal Papers and Relevant Theory

Recent advancements in AI and machine learning have enabled more sophisticated approaches to mental health assessment. For example, an Electronic Document Tracking System (EDTS) was

developed to improve the efficiency of retrieving documents online and tracking their movement. This system, although designed for document management, showcases the importance of integrating multiple data sources for real-time analysis. Similarly, current research in e-government systems emphasizes the need for document tracking and approval systems, highlighting the benefits of process automation. In the context of mental health, systems that incorporate multimodal data such as facial expressions, text, and speech can provide a more accurate and holistic understanding of an individual's mental state, leading to more effective interventions.[3]

5. System Architecture

The architecture of the proposed system includes the following stages:

- **Data Collection:** Input data is collected from three main sources—facial expressions (via camera), text (from written inputs), and speech (via microphone).

- **Preprocessing:** The collected data is cleaned, normalized, and prepared for analysis.

- **Feature Extraction:** Relevant features are extracted from each data source, such as emotion scores from facial expressions, sentiment analysis from text, and speech characteristics.

- **Model Integration:** The extracted features are combined using a fusion model, which may include machine learning or deep learning techniques, to predict mental health conditions.

- **Mental health detection:** Based on the predictions, and the model integration, the system gives appropriate mental health condition results.[1],[4]
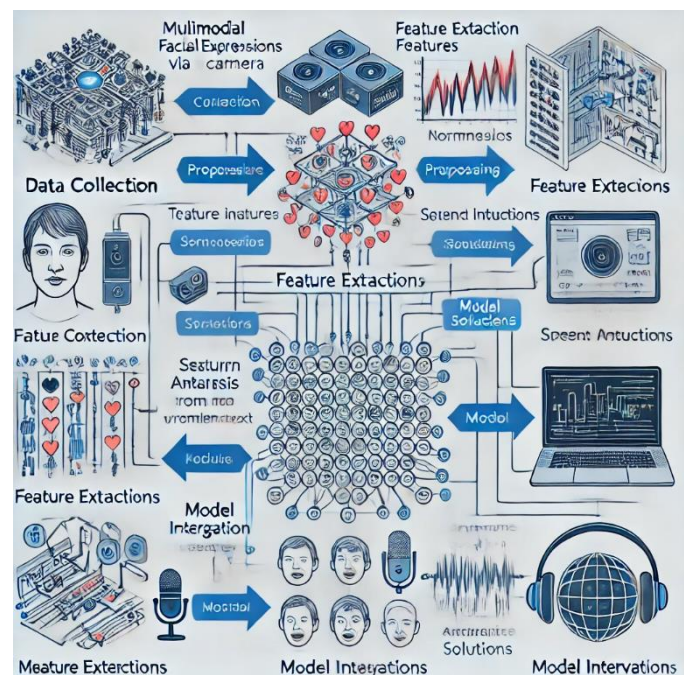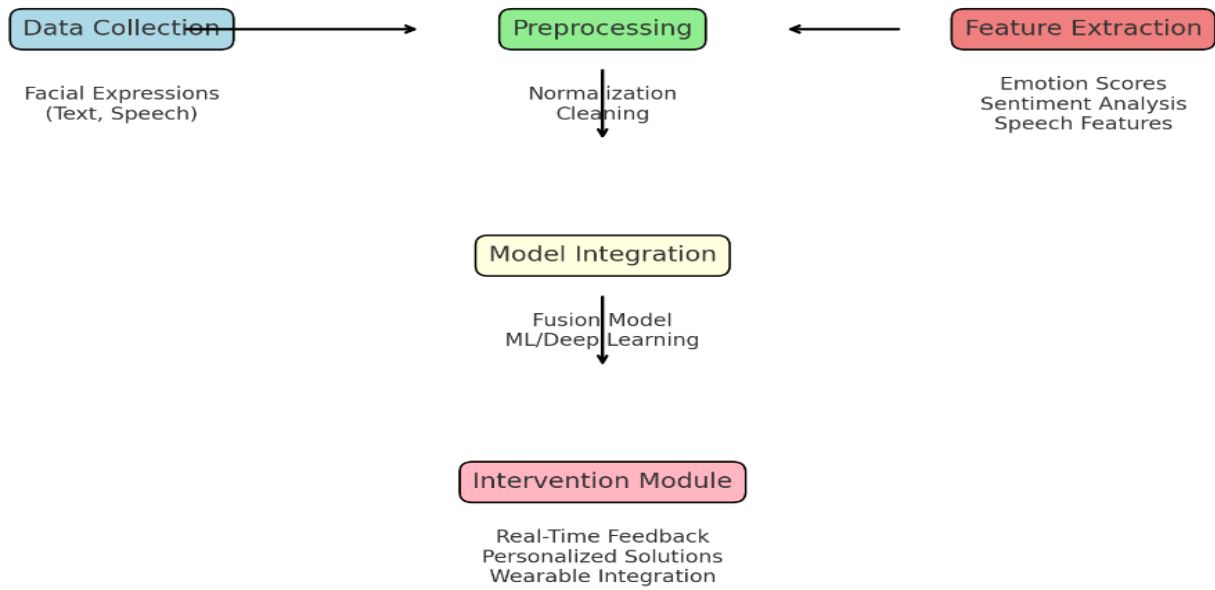
6. System Requirements

Hardware Requirements:

1. Processor: Intel Core i5 or Ryzen 5

2. Hard Disk: 256 GB

3. RAM: 8 GB

Software Requirements:

1. Operating System: Windows 10/11, Linux, MacOS

2. Browser: Mozilla, Opera, Chrome

3. Database: MySQL

4. Programming Languages: HTML, CSS, JavaScript, Python

5. Framework: Django

6. Tools for Analysis: OpenCV, SpaCy or NLTK, speech processing libraries

The proposed system will take input information from the user, in various forms and using various models, it will identify the mental health condition of the user independently by each module, and then the results will be combined and will give the integrated output.

7. Conclusion

Our project aims to improve mental health detection by integrating multiple data sources and providing real-time interventions. By transitioning from detection to intervention, the system has the potential to offer timely support, enhance mental well-being, and reduce the burden on traditional mental health care systems.

8. References

1. https://ieeexplore.ieee.org/document/9810926
2. https://ieeexplore.ieee.org/document/10360126

3. https://www.mdpi.com/1424-8220/24/2/348
4. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9914523/
5. **Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017).** "A Review of Affective Computing: From Multimodal Analysis to Affective Understanding." *IEEE Transactions on Affective Computing*
6. **Tzirakis, P., Trigeorgis, G., Nicolaou, M. A., Schuller, B. W., & Zafeiriou, S. (2017).** "End-to-End Multimodal Emotion Recognition Using Deep Neural Networks." *IEEE Journal of Selected Topics in Signal Processing*

Group Members

Sr. No.	Roll No	Student Name	Signature
1	56	Gund Krushna Dilip	
2	60	Jadhav Adesh Arun	
3	64	Jadhav Shravani Anil	
4	65	Jadhav Swapnil Jalindar	

Prof. P. M. Dhanrao   Dr. S.R. Deshmukh   Dr. D.B Kshirsagar

Project Guide   **Project Coordinator**   **HOD Computer Department**