IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

Student Mental Health Detection Using Machine Learning

Swapnil Shivasharan, Suyash Shinde, Tanmay Sharma, Arya Chavan, Prof. R.A. Patil
Department of Information Technology
Smt. Kashibai Navale College of Engineering, Pune, India.

Abstract: *Mental health has become a critical concern globally, especially among students, who face various pressures such as academic performance, peer relationships, and future career prospects. The increasing prevalence of mental health challenges among this demographic underscores the urgent need for proactive detection and intervention methods. Conventional approaches to identifying mental health issues are often limited by delayed reporting, stigma, and lack of awareness, leading to insufficient or late interventions. This project explores the application of machine learning to address these challenges by developing a system for early detection of mental health conditions among students. The proposed approach utilizes a datasets comprising behavioral, academic, and demographic features. After preprocessing and feature engineering, machine learning algorithms such as Logistic Regression, Support Vector Machines (SVM), and Random Forest are applied to classify students based on their mental health risk levels. The models are evaluated using performance metrics, including accuracy, precision, recall, and F1-score, to determine their reliability and effectiveness. Cross-validation techniques ensure the robustness of the developed system, while hyperparameter tuning optimizes model performance. This project contributes to the broader field of mental health technology by showcasing the intersection of data science and psychology. By leveraging predictive analytic , it provides a scalable, data-driven solution to a pressing issue, paving the way for smarter and more efficient mental health management strategies in educational institutions*

Keywords: Mental health

I. INTRODUCTION

The domain of this project lies at the intersection of mental health analysis, machine learning, and web-based information systems, with a specific focus on student stress detection. In the modern digital education era, students are increasingly exposed to high academic pressure, social expectations, lifestyle imbalances, and often, digital overload. These factors contribute to elevated levels of stress and deteriorating mental well-being, which can negatively impact academic performance, social interaction, and personal development. Therefore, monitoring and addressing mental health in educational settings has become a critical need.

Traditionally, stress levels and mental health conditions are monitored using manual surveys, counselor interactions, and periodic evaluations. However, these techniques are often subjective, non-continuous, and not scalable for large student populations. They also fail to provide timely intervention due to delays in data collection, analysis, and reporting.

The advancement of Artificial Intelligence (AI) and Machine Learning (ML) offers new opportunities in this domain. By leveraging data-driven algorithms and automated pipelines, it is now possible to create systems that analyze behavioral patterns, academic performance metrics, physiological data, and even social media activity to predict mental health trends more accurately and objectively. This gives rise to the domain of computational mental health assessment, where AI models like Logistic Regression, Support Vector Machines (SVMs), and Gradient Boosting Classifiers can analyze complex input variables and classify stress levels efficiently.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, June 2025 Impact Factor: 7.67

II. AIMS AND GOALS

The project aims to develop a system capable of early detection of student mental health issues by analyzing multiple data sources. The main **aims** are:

To create an integrated system that collects and processes multimodal data, including academic performance, behavioral patterns, and physiological metrics.

To implement machine learning models that analyze these data sources and generate insights into a student's mental well-being.

To offer personalized recommendations based on detected patterns, supporting both students and educators.

The **goals** include:

Developing a working application that allows for real-time data analysis and personalized mental health interventions.

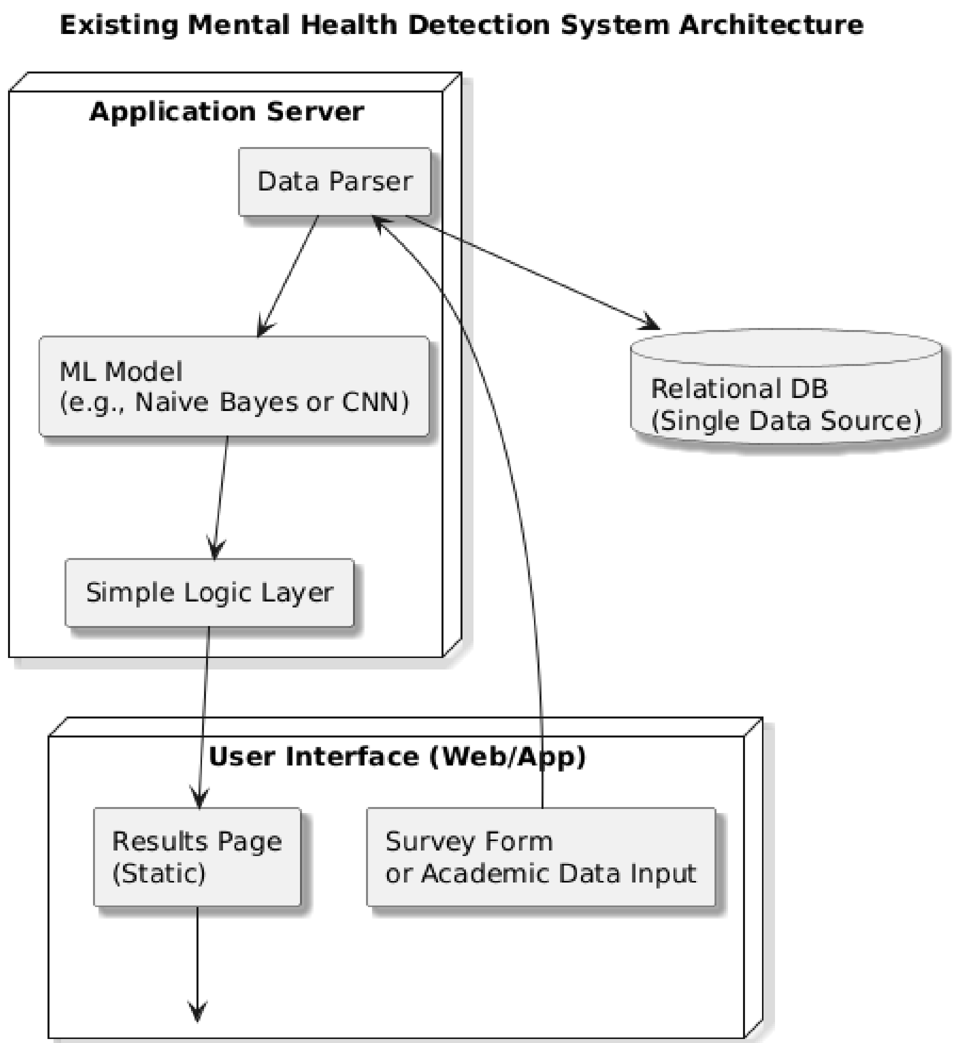
Ensuring the system is scalable and secure, with seamless integration into existing educational frameworks.

Achieving high accuracy in mental health detection, with an expected accuracy rate of over 90%.

III. GAP ANALYSIS

Feature	Existing Systems	Proposed "Student Mental Health Detection" System	Gap Highlighted
Technology Stack	Use traditional backend and frontend technologies.	MERN stack for frontend, Python for backend model design.	Modern and efficient tech stack (MERN).
Data Processing	Basic data processing with predefined fields.	Advanced word processing and tokenization techniques.	Enhanced accuracy through tokenization.
Dataset Source	Limited to static datasets from specific institutions or surveys.	Combination of Kaggle and real-time data collection.	Real-time data collection for better insights
Model Accuracy	Varies but generally ranges from 70% to 85%.	Targeting more than 90% accuracy.	Significant improvement in prediction accuracy
Notification System	Limited or no integration of notification systems.	Email API integration for real-time notifications.	Real-time alerts for timely interventions
Input Method	Structured forms or predefined survey responses.	Free-text input for more flexible data collection.	Increased flexibility in data input
Personalization	Generalized models with little personalization.	Potential for personalized detection based on individual inputs.	Tailored mental health detection.
User Interface (UI)	Often less interactive or outdated.	Modern, responsive UI built with MERN stack.	Enhanced user experience with modern UI.
Scalability	Limited scalability due to older tech and static datasets.	Scalable to accommodate real-time data and cloud-based architecture.	Scalability for broader adoption

IV. EXISTING SYSTEM ARCHITECTURE



IV. SYSTEM ARCHITECTURE

Existing mental health detection systems typically rely on singular data sources like academic performance, survey data, or occasional behavioral observations. While these systems provide some insights, they often fail to offer a holistic picture of a student's mental health. Systems such as **Heart Disease Prediction with Naive Bayes** or **Stress Detection using CNN** serve as benchmarks for how health predictions can be made through data analysis. However, their applications are limited due to their singular focus on a specific health issue or data source.

Our proposed system enhances these traditional approaches by integrating **multimodal data fusion**, bringing in academic records, social media activity, and physiological metrics to offer a more comprehensive mental health assessment. Existing literature such as **Multimodal Educational Data Fusion** (2022) demonstrates the efficacy of combining multiple data streams, while machine learning algorithms such as **LightGBM** and **Deep Neural Networks** have shown potential in improving predictive accuracy.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

VI. PROPOSED SYSTEM ARCHITECTURE

The system architecture of the Student Mental Health Detection project integrates various data sources, processes the data, and applies machine learning algorithms to detect mental health issues. The system is composed of several layers:

Data Collection Layer: Gathers multimodal data from academic records, behavioral patterns, physiological data (from wearable devices), social media activity, and self-reported surveys.

Data Preprocessing Layer: Prepares the data for analysis by cleaning, tokenizing text, normalizing values, and extracting relevant features for machine learning.

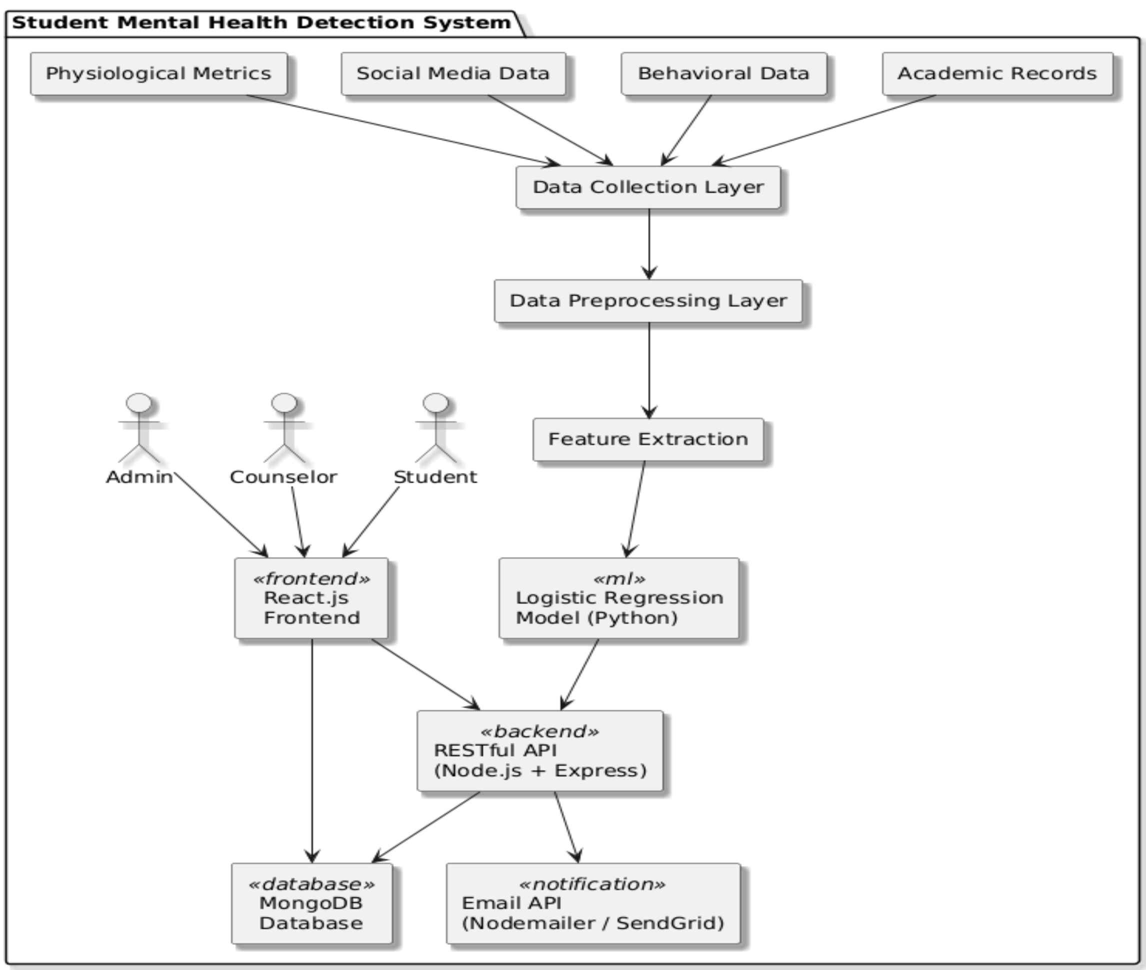
Data Fusion Layer: Combines the processed data using feature-level and decision-level integration techniques to generate a comprehensive dataset for analysis.

Machine Learning Analysis Layer: Applies algorithms like Support Vector Machines (SVMs), Recurrent Neural Networks (RNNs), and Logistic Regression to detect patterns indicative of mental health risks.

Decision-Making Layer: Generates insights and predictions about the student's mental health and triggers appropriate alerts and recommendations.

User Interaction Layer: Provides interfaces for students, educators, and counselors to interact with the system, submit data, and view reports.

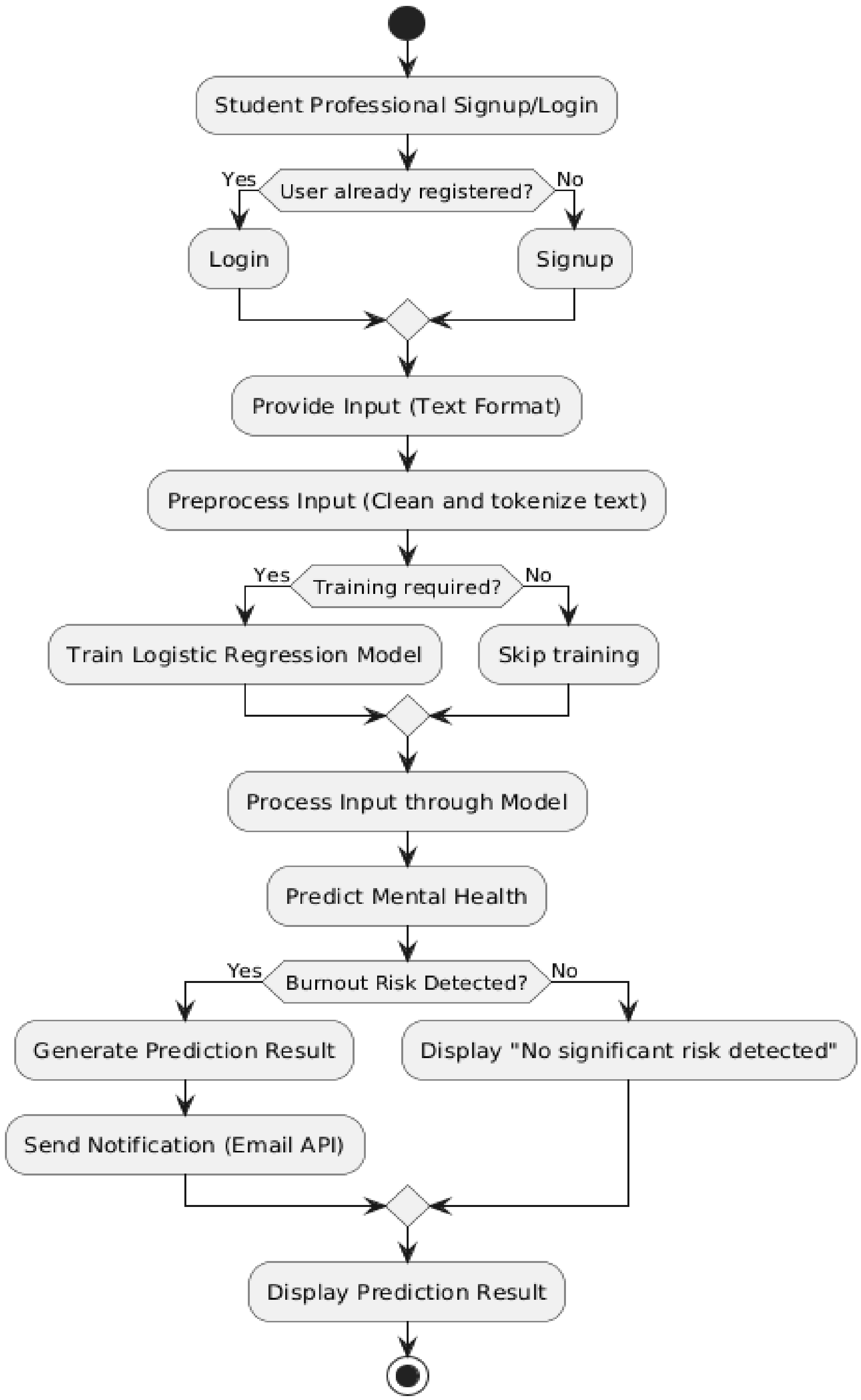
Notification and Alert System: Sends real-time alerts to students, counselors, or educators when mental health risks are detected.



Proposed System Architecture

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

VII. ACTIVITY DIAGRAM



Activity Diagram

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

The activity diagram represents the workflow of mental health detection:
Start: The student or educator logs into the system.
Submit Input: The student provides input data (text, behavioral metrics).
Preprocessing: The system cleans and tokenizes the data.
Model Training (if required): The machine learning model is trained on the dataset.
Prediction: The system processes the data through the model to predict mental health risks.
Decision Point: If a mental health risk is detected, the system generates a recommendation and triggers an alert.
End: The results are displayed, and notifications are sent if necessary

VIII. PROPOSED ALGORITHMS

1. Logistic Regression:
Used for binary classification to assess the likelihood of mental health issues based on input features.

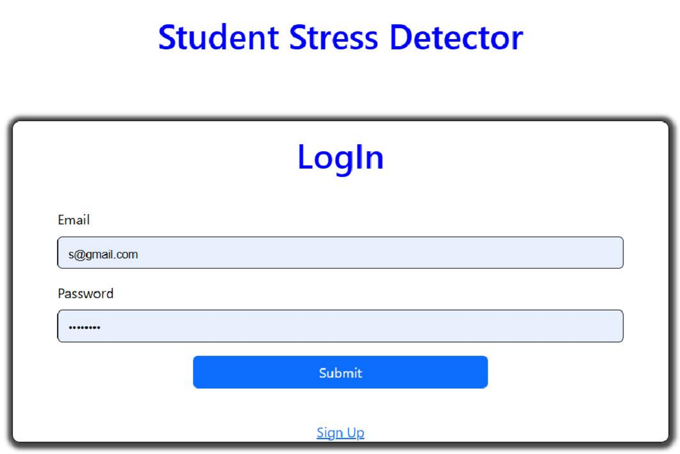
2. MERN Stack Integration:
Combines MongoDB, Express.js, React, and Node.js for a seamless, full-stack application experience, enabling real-time data processing and user interaction.

3. Python Module Design:
Implements data preprocessing, feature extraction, and model training for accurate predictions, leveraging libraries such as Pandas and Scikit-learn.

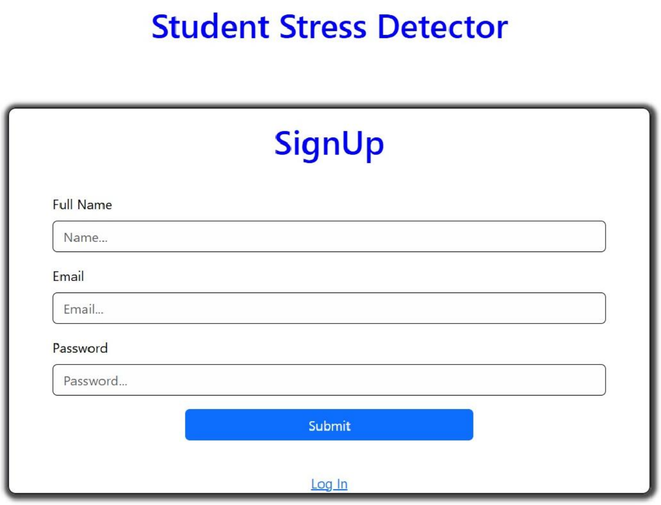
4. Data Visualization Tools:
Utilizes libraries like Matplotlib and Seaborn in Python for visualizing data trends and model outputs, aiding in interpretability.

5. RESTful API Development:
Facilitates communication between the front end and back end, allowing for efficient data retrieval and submission

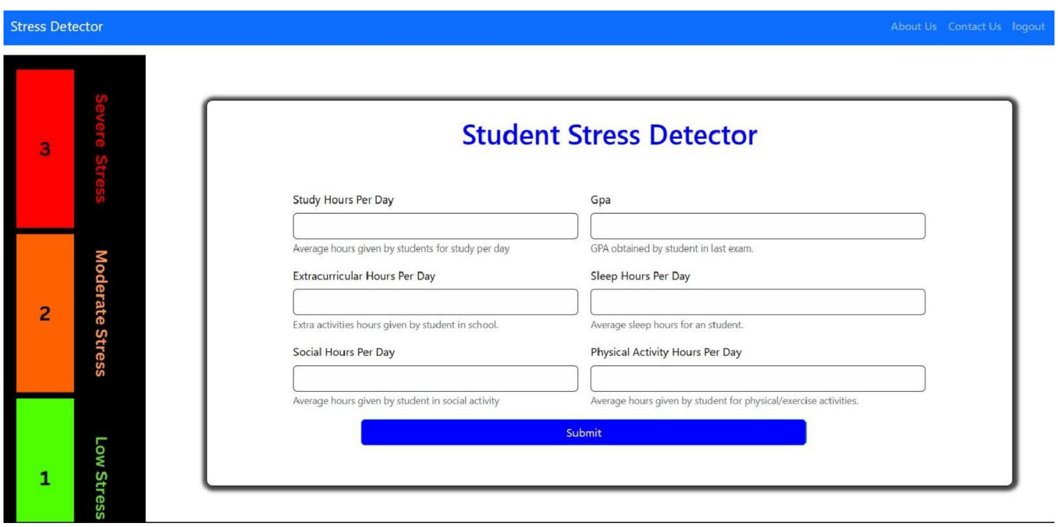
9. Snap Shot
Login



Sign up

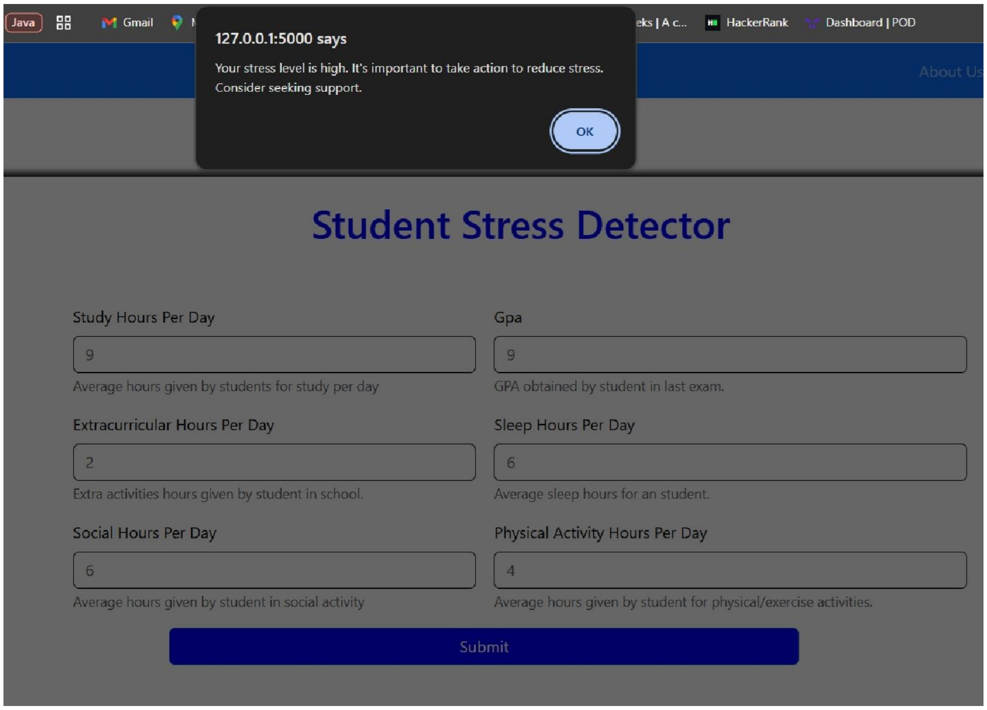


Home Page

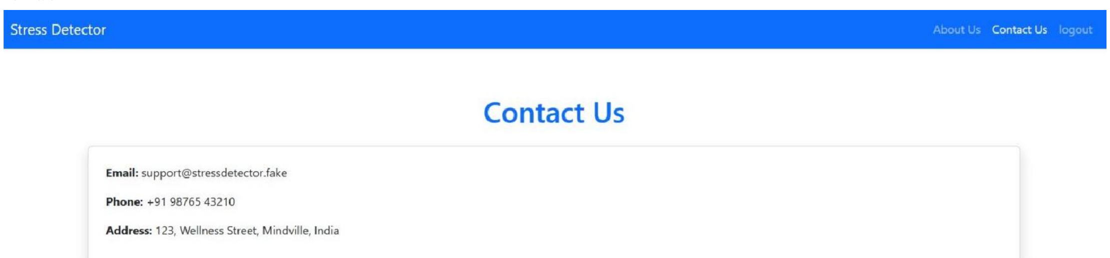


IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

Pop up message

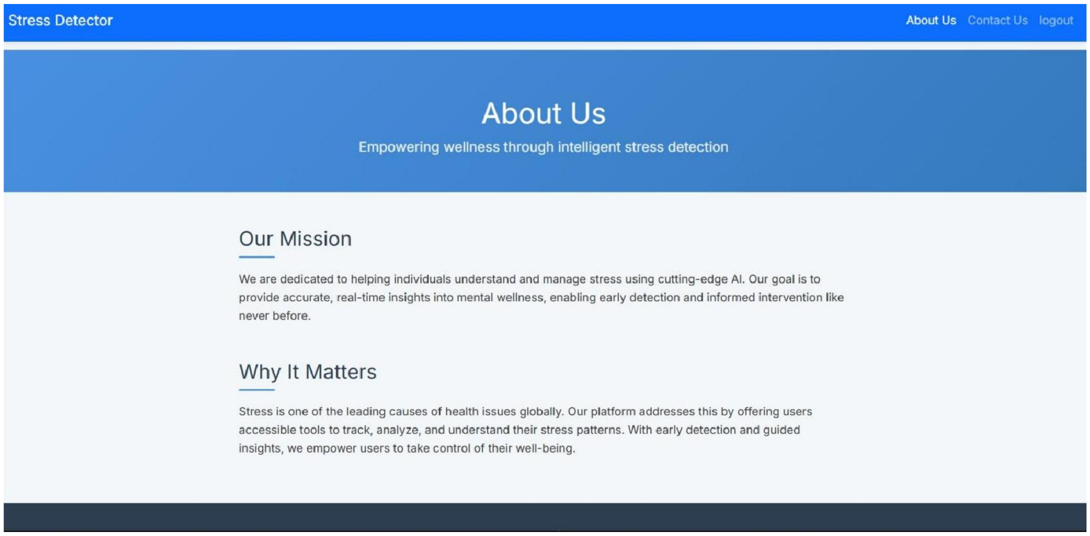


Contact us



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

About us



X. LIMITATIONS AND FUTURE WORK

A. Limitations

1. Dependence on data quality for accuracy.
2. Challenges in scaling to larger populations beyond students.
3. Privacy concerns require robust encryption methods.

B. Future Work

1. **Real-time Data Integration**: Expand data sources to include biometric data and advanced sentiment analysis.
2. **Personalization**: Tailor recommendations based on individual profiles.
3. **Global Adaptation**: Incorporate multilingual and region-specific features for broaderapplicability.
4. **Enhanced Privacy Measures**: Implement advanced encryption techniques to address ethical concerns

XI. CONCLUSION

This paper presents a novel approach to detecting student mental health issues using machine learning. By integrating diverse data sources and employing scalable technologies, the system addresses significant gaps in current frameworks. Future enhancements aim to further improve the system's accuracy, scalability, and privacy.

REFERENCES

[1]. Thejaswini M, Harsha Vardhan K M, E Vigneshwar, Harsha B, Darshan S. "Stress Detection in IT Professionals by Image Processing and Artificial Intelligence." *International Journal of Research Publication and Reviews*, June 2023.
[2]. Satvik Gurjar, Chetna Patil, Ritesh Suryawanshi, Madhura Adadande, Ashwin Khore, Noshir Tarapore.
[3]. "Mental Health Prediction Using Machine Learning." *International Research Journal of Engineering and Technology (IRJET)*, December 2022.
[4]. Jetli Chung, Jason Teo. "Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges." *Applied Computational Intelligence and Soft Computing*, 2022.
[5]. Teng Guo, Wenhong Zhao, Mubarak Alrashoud, Amr Tolba, Selena Firmin, Feng Xia. "Multimodal Educational Data Fusion for Students' Mental Health Detection." *IEEE Access*, July 2022.
[6]. K. Vembandasamy, R. Sasipriya, E. Deepa. "Heart Disease Detection Using Naive Bayes Algorithm."
[7]. *International Journal of Innovative Science Engineering & Technology (IJISET)*, September 2015.