Predictions of College Students' Mental Stress using Machine Learning Algorithms

B. Haritha Lakshmi[1], Nunemunthala Srija[2], P. Vijayamma[2], Siliveri Sanjana[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Information Technology

[1,2]Malla Reddy Engineering College for Women (UGC-Autonomous), Maisammaguda, Hyderabad, 500100, Telangana.

ABSTRACT

The mental health of college students has become an increasingly urgent concern, given the rising levels of academic stress and its detrimental effects on well-being and performance. Traditional methods for identifying and managing student stress—such as surveys, questionnaires, interviews, and observational techniques—are often reactive and lack the capacity for timely intervention. Consequently, there is a growing need for predictive, data-driven systems that can offer real-time insights into stress levels. In this research, we present a machine learning–based framework designed to predict mental stress among college students. By integrating diverse data sources—ranging from academic performance and attendance records to self-reported measures and physiological parameters—our system detects subtle patterns that precede heightened stress levels. The proposed model utilizes classification algorithms like Random Forest and SVM, which have demonstrated robust predictive capabilities in existing wearable and smartphone-based stress monitoring solutions (e.g., Stress Sense, EduWell, and WellBe). The overarching goal is to facilitate early interventions, thereby improving both student well-being and academic outcomes. This approach underscores the critical role of proactive stress management, offering a scalable, real-time solution for higher education institutions aiming to enhance student support services.

Keywords: Stress monitoring systems, Early intervention, Academic performance, Physiological data, Real-time stress prediction, Stress management

1. INTRODUCTION

Mental health concerns among college students have become a growing area of focus for educational institutions worldwide. Increasing academic demands, social pressures, and the transition to adulthood place many under considerable stress, adversely impacting both academic performance and overall well-being. Traditional methods of stress assessment—such as surveys, questionnaires, interviews, and observational approaches—offer valuable insights but are inherently reactive, often identifying stress after it has reached critical levels. Consequently, these conventional techniques may delay intervention, missing opportunities to address stress in its early stages. The urgency of this challenge has propelled the development of data-driven approaches aimed at real-time stress prediction. Recent advances in machine learning (ML) have shown significant promise in identifying subtle patterns indicative of mental stress. By leveraging a broad spectrum of data—ranging from academic performance metrics to physiological signals—ML techniques can uncover correlations and trends that would remain hidden in manual assessments. Such real-time monitoring and predictive capabilities could pave the way for timely, targeted interventions, ultimately improving both the student experience and academic success. In this study, we propose a comprehensive, ML-based framework for predicting mental stress among college students. Our approach integrates conventional indicators of academic performance and attendance with self-reported and physiological data, thereby offering a multifaceted view of students' mental health.

Drawing upon classification models such as Random Forest and Support Vector Machines (SVM), the system seeks to deliver accurate, real-time stress predictions. This integrated perspective not only advances early detection but also provides actionable insights that can guide supportive measures, reinforcing the importance of proactive mental health strategies within higher education environments.

2. LITERATURE SURVEY

Towbes and Cohen [1] conducted a study on chronic stress among college students. Their research focused on developing a scale to measure stress and predict distress levels among students. They identified various stressors related to academic and social pressures and analyzed their impact on students' mental health. The study provided valuable insights into how chronic stress affects students and highlighted the need for effective stress management strategies in educational settings. This foundational work laid the groundwork for subsequent research on stress among young adults and informed the development of interventions aimed at reducing student distress. MQ Mental Health [2] addressed the impact of stress on mental health. The article explored the physiological and psychological effects of stress and discussed strategies for managing and mitigating stress. It emphasized the importance of understanding the various dimensions of stress and provided practical recommendations for improving mental health through lifestyle changes and therapeutic approaches. This resource is instrumental in raising awareness about the consequences of stress and offering guidance on how to effectively tackle mental health challenges associated with stress. Ghaderi et al. [3] presented a study on machine learning-based signal processing for stress detection. The authors utilized machine learning techniques to analyze physiological signals, such as heart rate and skin conductance, to detect stress levels. Their research demonstrated the potential of integrating machine learning with biometric data to achieve accurate stress detection. By leveraging advanced algorithms and physiological metrics, the study contributed to the development of more sophisticated tools for monitoring and managing stress, showcasing the intersection of technology and mental health. Vogel et al. [4] explored the role of physical activity in stress management during the COVID-19 pandemic through a longitudinal survey. Their study examined how changes in physical activity levels influenced stress management and mental well-being during the pandemic. The research highlighted the significant impact of physical activity on reducing stress and improving mental health, particularly during times of crisis. By analyzing survey data, the study provided valuable insights into how maintaining an active lifestyle can serve as an effective strategy for managing stress in challenging situations.

Li and Liu [5] investigated stress detection using deep neural networks. The authors applied deep learning models to analyze data related to stress, demonstrating the effectiveness of advanced neural network architectures in predicting stress levels. Their research underscored the potential of deep learning technologies to enhance the accuracy and reliability of stress detection methods. The study contributed to the growing body of knowledge on using artificial intelligence for mental health applications, showcasing how deep learning can be leveraged to address complex psychological issues. Subhani et al. [6] developed a machine learning framework for detecting mental stress at multiple levels. The authors designed a comprehensive framework that integrates various machine learning techniques to analyze stress from different perspectives, including psychological and physiological aspects. Their research highlighted the effectiveness of machine learning in providing a nuanced understanding of mental stress and offered a multi-level approach to stress detection. This study contributed to the advancement of machine learning applications in mental health, emphasizing the importance of a holistic approach to stress assessment. Gan et al. [7] presented a study on fatigue life prediction considering mean stress effects. The authors utilized random forests and kernel extreme learning machines to develop a predictive model for fatigue life, incorporating the effects of mean stress. Their research demonstrated how machine learning techniques can be applied to predict fatigue and stress-related outcomes, providing valuable insights into the durability and performance of materials under stress. The study highlighted the potential of combining machine learning with engineering applications to address complex problems related to stress and fatigue. Masood and Alghamdi [8] explored modeling mental stress using a deep learning framework. They applied deep learning methods to model and predict mental stress, showcasing the effectiveness of these techniques in capturing complex stress patterns. The study provided a detailed analysis of how deep learning models can be utilized to understand and predict mental stress, contributing to the development of advanced tools for mental health management. The research highlighted the potential of deep learning in enhancing the precision and applicability of stress prediction models. Norizam [9] focused on the determination and classification of human stress indices using EEG signals. The study employed nonparametric analysis techniques to assess stress levels based on EEG data, providing insights into the physiological underpinnings of stress. By analyzing EEG signals, the research demonstrated the feasibility of using neurophysiological data for stress assessment. This study contributed to the understanding of how brain activity can be used to evaluate stress, offering a valuable approach for integrating physiological data into stress management strategies. Ahuja and Banga [10] investigated mental stress detection among university students using machine learning algorithms. The authors applied various machine learning techniques to identify stress patterns among students, offering practical solutions for monitoring and managing stress in academic settings. Their research highlighted the potential of machine learning to address mental health issues in educational environments, providing tools for early detection and intervention. The study contributed to the field of educational data mining and mental health by demonstrating the application of machine learning in stress detection.

Xu et al. [11] conducted a cluster-based analysis for personalized stress evaluation using physiological signals. The authors utilized clustering techniques to analyze physiological data and provide personalized stress evaluations. Their research highlighted the importance of tailoring stress assessments to individual profiles, offering insights into how personalized approaches can enhance stress management. The study contributed to the development of more targeted and effective stress evaluation methods, leveraging physiological data for personalized care. AlSagri and Ykhlef [13] explored a machine learning-based approach for depression detection on Twitter using content and activity features. They employed machine learning algorithms to analyze social media data for detecting signs of depression, highlighting the potential of using online content for mental health monitoring. The study demonstrated how social media platforms can be utilized to identify and address mental health issues, contributing to the field of digital mental health assessment.

3. PROPOSED ALGORITHM

Develop a machine learning system to predict mental stress levels among college students. The system aims to provide timely and accurate predictions to enable early intervention and improve student well-being.

Step 1: Dataset

The research begins with acquiring a dataset that includes relevant features affecting mental stress among college students. This dataset typically comprises various attributes such as academic performance metrics, social interactions, physiological parameters, and possibly historical survey data on mental health.

Ensuring the dataset's quality and relevance is crucial as it forms the foundation for subsequent analysis and model training. Data collection methods may involve surveys, academic records, and physiological monitoring devices.

Step 2: Dataset Preprocessing

Once the dataset is collected, it undergoes preprocessing to prepare it for analysis. This step involves handling null values, which could distort the accuracy of the model if left unaddressed. Null values are typically managed by either imputation, where missing values are filled with statistical measures like the mean or median, or by removing entries with missing data if they are insignificant. Additionally, label encoding is performed to convert categorical variables into numerical formats. This transformation is essential for machine learning algorithms to process categorical data effectively.

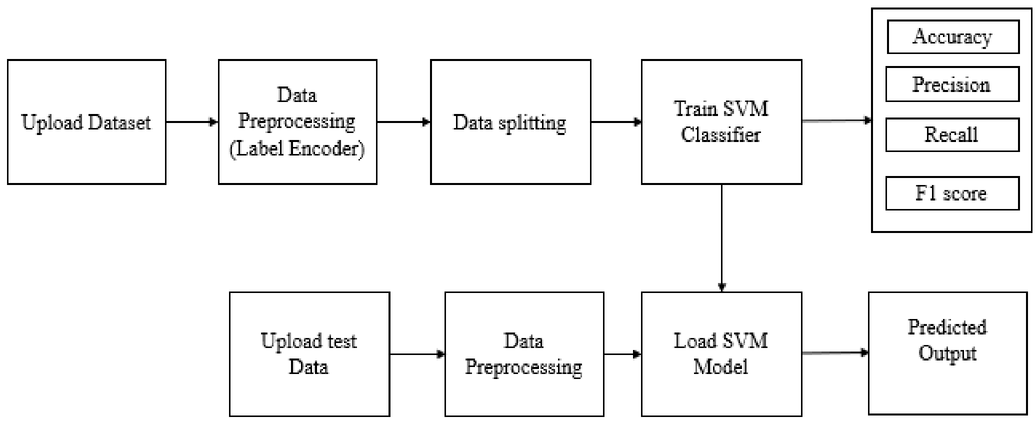

Figure 1; Block Diagram

Step 3: Label Encoding

Label encoding is a crucial preprocessing step that converts categorical variables into numerical values. For example, variables like "gender" with values "male" and "female" are transformed into numerical codes (e.g., 0 and 1). This encoding allows machine learning models to interpret categorical features as numerical inputs, facilitating their incorporation into the predictive algorithms. Proper encoding ensures that the model can handle categorical data and improve its predictive performance.

Step 4: Existing Model

The research then evaluates the performance of existing machine learning models, starting with the Decision Tree algorithm. Decision Trees are a popular choice for classification tasks due to their simplicity and interpretability. They work by splitting the data into subsets based on feature values, creating a tree-like model of decisions. This model helps in understanding how different features contribute to predicting mental stress levels. By assessing the Decision Tree’s performance, researchers gain insights into its effectiveness and limitations in predicting stress levels.

Step 5: Proposed Model

Following the evaluation of the Decision Tree, the research introduces a proposed model using Support Vector Machine (SVM) for classification. SVM is known for its ability to handle high-dimensional data and its robustness in finding the optimal hyperplane that separates different classes. By applying SVM, the study aims to improve the accuracy and reliability of stress level predictions compared to traditional methods. The SVM model is trained on the preprocessed dataset, optimizing its parameters to enhance performance.

Step 6: Performance Comparison

The next step involves comparing the performance of the Decision Tree and SVM models. This comparison is carried out using various performance metrics such as accuracy, precision, recall, and F1-score. Evaluating these metrics helps in understanding how well each model predicts mental stress levels and highlights the strengths and weaknesses of the algorithms. The performance comparison is crucial for determining the most effective model for real-time stress prediction.

Step 7: Prediction of Output from Test Data with SVM Trained Model

With the SVM model trained and optimized, it is then used to make predictions on test data. This step involves feeding new, unseen data into the trained SVM model to predict mental stress levels. The predictions are analyzed to assess how well the model generalizes to new data and whether it maintains accuracy and reliability. This real-world testing ensures that the model performs effectively outside the training environment.

Step 8: Evaluation and Refinement

Finally, the research concludes with an evaluation of the SVM model's predictions and overall effectiveness. If necessary, further refinements are made to improve the model's performance based on test data results. This step may involve tweaking model parameters, incorporating additional features, or revisiting preprocessing techniques to enhance prediction accuracy. The ultimate goal is to develop a robust and reliable system for predicting mental stress levels in college students, contributing valuable insights for early intervention and support.

4. RESULT

Figure 2 illustrates the user interface for the Mental Stress Level Prediction system. The GUI allows users to input relevant data and view the predicted stress level in real-time. Its design focuses on ease of use and clarity, enabling students or counselors to interact with the system seamlessly. Figure 3 presents the questionnaire module used to gather self-reported data on various stress indicators. This includes questions related to academic workload, emotional state, and lifestyle factors. The responses collected here form the basis for the machine learning model to predict stress levels accurately. Figure 4 demonstrates the output interface showing a sample result of "Acute Stress." Alongside the stress level, the system offers preventive measures and suggestions tailored to the user's condition. This proactive approach helps guide timely interventions, contributing to better stress management and overall well-being.

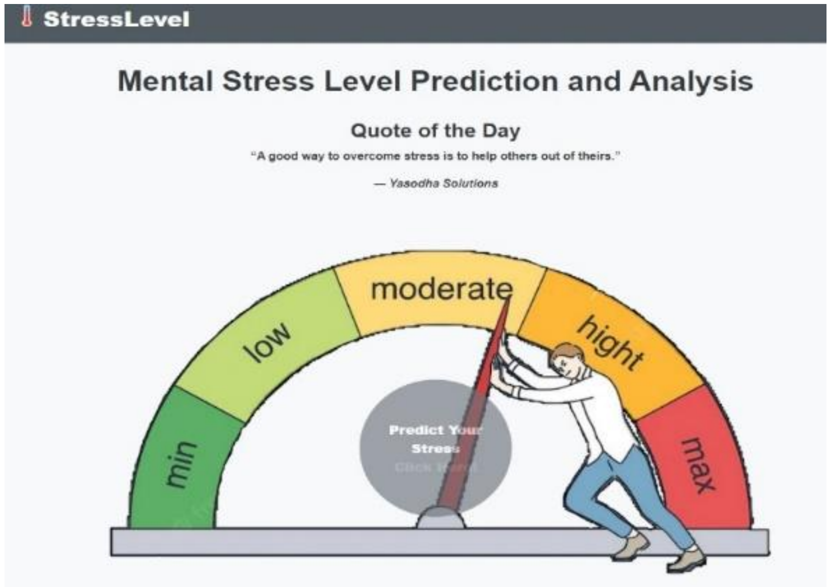

Figure 2: GUI

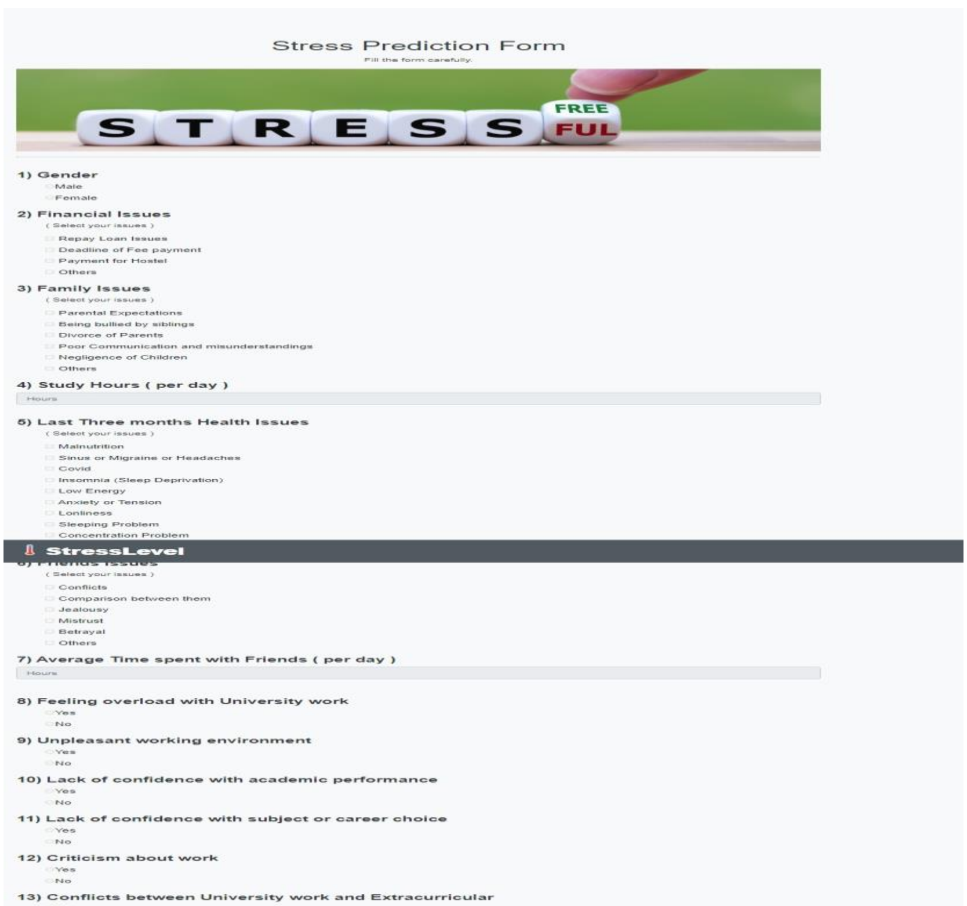

Figure 3: Questionnaire

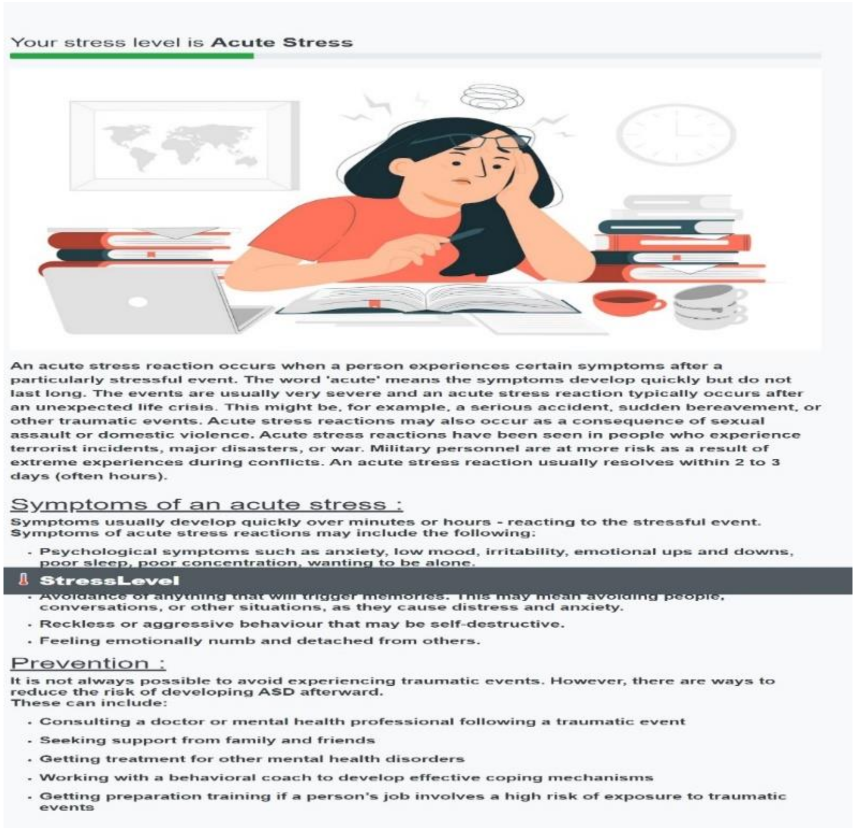

Figure 4: output

5. CONCLUSION

The mental health of college students is a growing concern, particularly with the rising levels of stress that can significantly impact academic performance, social interactions, and overall well-being. Traditional methods for assessing and managing stress, such as surveys, counseling sessions, and observational techniques, have played a crucial role in understanding student stress. However, these methods have limitations, including delayed detection, subjectivity, resource intensity, and lack of real-time monitoring. These limitations underscore the need for more advanced solutions that can provide timely, accurate, and proactive stress management. The integration of machine learning and AI into mental health monitoring represents a significant advancement in addressing these challenges. Machine learning models, such as Random Forest and Support Vector Machine (SVM), can analyze a wide range of data sources—including academic performance, social interactions, and physiological metrics— to predict stress levels with greater accuracy and timeliness. These models offer the ability to detect stress early, allowing for interventions before stress has a detrimental impact on students' lives. By leveraging real-time data, these predictive models can continuously monitor stress indicators, offering a dynamic and responsive approach to mental health management. This shift from reactive to proactive stress management has the potential to revolutionize the way mental health is addressed in academic settings. Early detection of stress not only improves students' academic performance but also enhances their overall quality of life by providing targeted support when it is most needed.

REFERENCES

[1] L.C. Towbes and L.H. Cohen. "Chronic stress in the lives of college students: Scale development and prospective prediction of distress." *Journal of Youth and Adolescence*, vol. 25, no. 2, pp. 199-217, 1996.

[2] MQ Mental Health. "Stress and our mental health: What is the impact & how can we tackle it?" May 2018. [Online] Available: https://www.mqmentalhealth.org/stress-andmental-health.

[3] A. Ghaderi, J. Frounchi, and A. Farnam. "Machine learning-based signal processing using physiological signals for stress detection." In *2015 22nd Iranian Conference on Biomedical Engineering (ICBME)*, pp. 93-98, 2015, November.

[4] E.A. Vogel, J.S. Zhang, K. Peng, C.A. Heaney, Y. Lu, D. Lounsbury et al. "Physical activity and stress management during COVID-19: A longitudinal survey study." *Psychology & Health*, vol. 37, no. 1, pp. 51-61, 2022.

[5] R. Li and Z. Liu. "Stress detection using deep neural networks." *BMC Medical Informatics and Decision Making*, vol. 20, 2020.

[6] A.R. Subhani, W. Mumtaz, M.N.B.M. Saad, N. Kamel, and A.S. Malik. "Machine learning framework for the detection of mental stress at multiple levels." *IEEE Access: Practical Innovations Open Solutions*, vol. 5, pp. 13545-13556, 2017.

[7] L. Gan, H. Wu, and Z. Zhong. "Fatigue life prediction considering mean stress effect based on random forests and kernel extreme learning machine." *International Journal of Fatigue*, vol. 158, 2022.

[8] K. Masood and M.A. Alghamdi. "Modeling mental stress using a deep learning framework." *IEEE Access: Practical Innovations Open Solutions*, vol. 7, pp. 68446-68454, 2019.

[9] S. Norizam. "Determination and classification of human stress index using nonparametric analysis of EEG signals," 2015.

[10] R. Ahuja and A. Banga. "Mental stress detection in university students using machine learning algorithms." *Procedia Computer Science*, vol. 152, pp. 349-353, 2019.

[11] Q. Xu, T.L. Nwe, and C. Guan. "Cluster-based analysis for personalized stress evaluation using physiological signals." *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 275-281, 2014.

[12] A. Ghaderi, J. Frounchi, and A. Farnam. "Machine learning-based signal processing using physiological signals for stress detection." In *2015 22nd Iranian Conference on Biomedical Engineering (ICBME)*, pp. 93-98, 2015.

[13] H.S. AlSagri and M. Ykhlef. "Machine learning-based approach for depression detection in Twitter using content and activity features." *IEICE Transactions on Information and Systems*, vol. 103, pp. 1825-1832, 2020.