JOURNAL OF ARTIFICIAL INTELLIGENCE AND ROBOTICS
Jan-Mar 2025, VOL. 2, ISSUE 1, pp. 8-11
https://doi.org/10.61577/jaiar.2025.100003

RESEAPRO JOURNALS

ORIGINAL RESEARCH

OPEN ACCESS

Detection of mental stress in university students using a non-invasive machine learning system

Om Prakash Mohapatra, Radhika K, Aditya Rajan, Saakshith V and Varun B Patel
Department of Computer Science and Engineering, Dayananda Sagar University, Bengaluru, India

ABSTRACT

Mental disorders, especially stress, are on the rise among university students, negatively impacting academic achievement, wellbeing, and long-term health. The present research puts forward a data-driven, non-invasive mental stress detection system through a Random Forest Classifier based on self-reported questionnaire responses. The dataset included answers from students at universities on different academic, behavioral, and emotional stress indicators. The model was trained and tested using an 80:20 train-test split after preprocessing and feature selection. The Random Forest Classifier obtained an accuracy of 91.3%, precision of 89.7%, recall of 90.5%, and F1-score of 90.1%, performing better compared to other models like Support Vector Machine and Logistic Regression. Feature importance analysis revealed academic workload, sleep deprivation, emotional instability, and perceived lack of social support as the most important predictors of stress. The results show the potential of machine learning algorithms to facilitate scalable, real-time, and non-invasive stress detection. This system provides a valuable tool for mental health monitoring in educational institutions and may facilitate timely interventions to enhance student well-being. Future research will center on model generalization to a variety of student populations and on coupling with real-time digital technologies.

KEYWORDS

Mental stress detection; Machine learning; Random forest classifier; Non-invasive assessment

ARTICLE HISTORY

Received 10 February 2025; Revised 03 March 2025; Accepted 12 March 2025

Introduction

The rising prevalence of mental stress among university students has become an increasing concern in recent years. Stress can substantially impact students' academic performance, well-being, and quality of life. Early detection and intervention of stress levels are crucial to ensure students receive the required support and interventions. In this context, machine learning algorithms were found to be a promising means to conduct objective and accurate stress detection [1]. The study aims at identifying and analyzing rising levels of stress among students with the view to making an early prediction and intervention before the stress results in any adverse effects on their mental health and academic performance.

Stress, by definition, is a subjective phenomenon dependent on individual thresholds, environmental stimuli, and physiological reactions. The traditional methods of assessing stress levels, including self-report questionnaires, psychological tests, and counselor interviews, while helpful, are reactive and suffer from limitations such as recall bias, underreporting, and lack of timeliness [2,3]. Additionally, because of stigma or lack of awareness, students also avoid seeking assistance, further compounding the time lag in detection and treatment. Against this context, there is an increased interest in the use of digital health solutions [4], more so those enabled by artificial intelligence (AI) and machine learning (ML), to deliver more objective, real-time, and non-invasive means of detecting mental stress.

This study proposes the design of a machine learning system to detect mental stress in university students based on a Random Forest Classifier algorithm. The system leverages a strong set of responses to organized situational questionnaires with the option for voluntary physiological signals (if available) to have a multi-faceted view of stress patterns. The survey contains questions that measure diverse psychological domains—cognitive load, emotional control, social engagement, academic pressure, and lifestyles habits—that have been proven to tie in with stress levels. In contrast with physiological measures that involve the need for specialized tools, the surveys offer a scalable, easily applied data gathering approach that remains non-invasive and student-centric.

The system takes advantage of machine learning algorithms' capabilities in processing a wide range of data gathered from students within a university. This data consists of situational questionnaires, physiological indices, and other influential factors to provide an overall perspective of the students' stress levels. By analyzing this dataset, the system aims to extract informative features that can effectively indicate stress levels.

Machine learning models, especially the Random Forest Classifier, will be utilized to train on the gathered dataset and create a stress detection model [5]. The system, with iterative training, will identify patterns and signs of stress, allowing it to make successful predictions [6]. Through iterative training, and test accuracy metrics will be used to assess the performance of the model to ensure reliability and effectiveness in detecting stress [7]. This method reduces overfitting, processes high-dimensional data in an effective manner, and provides variable importance estimates that can be used to determine

*Correspondence: Mr. Om Prakash Mohapatra, Department of Computer Science and Engineering, Dayananda Sagar University, Bengaluru, India, e-mail: omprakashmohapatra007@gmail.com

which variables have the highest impact on classifying stress levels [5]. These features make it the best option for constructing a reliable model that will perform well with new data and keep high accuracy rates in predictions.

In addition to the technical benefits, there is significant potential in the integration of a stress detection system powered by machine learning within educational contexts [8]. A machine learning-powered system may be utilized as an advanced notice tool that will allow identifying high-risk students earlier, giving timely psychological attention, academic advisement, and well-being intervention [9]. The real-time features of ML models can be implemented in mobile or web-based systems, enabling frequent stress evaluation, mood monitoring, and automatic notifications to mental health professionals. Furthermore, data-driven systems offer a standardized and impartial stress evaluation method, eliminating the subjectivity involved in manual assessments.

The proposed method is also consistent with wider directions of digital health innovation and precision medicine, where AI and data analysis are increasingly used to guide decisions regarding health. In the university environment, this can translate into a more inclusive and caring environment, where ongoing and unobtrusive monitoring of mental health is enabled. With ML algorithms used to identify stress patterns, the system can also provide insights into stressors so that students can gain increased self-knowledge and develop coping skills specific to their stress profiles [10].

However, the application of machine learning in mental health research comes with challenges. There are ethical issues around data privacy, informed consent, and algorithmic bias that must be handled carefully [11]. The predictability is extremely dependent on the diversity and accuracy of the training data. Thus, adequate efforts should be made to ensure that the dataset includes a representative selection of students across varied demographics, academic fields, and cultural backgrounds. Furthermore, although ML systems may be used to support stress detection, they must supplement and not supplant human judgment and clinical assessment [12].

In summary, this research aims to fill the gap between technology and student mental health by creating a machine learning-based stress detection system. Through a Random Forest Classifier trained on student questionnaire data, the system will be able to give an accurate, objective, and scalable means of detecting stress among university students. Through the early detection and timely intervention it enables, this method not only improves the well-being of individuals but also promotes a more robust and mentally sound academic population. The effective deployment of such a system could lead to wider uses of AI in educational psychology, mental health monitoring, and student support services.

Literature Review

Mental stress in university students is a public health issue on the rise, triggered by academic pressures, lifestyle issues, and social demands. Different computational and physiological methods have been investigated to identify and measure mental stress, capitalizing on innovation in bio-sensing technologies, brainwave detection, and machine learning. This section reviews core studies applicable to the design of machine learning-based stress detection systems, especially those based on physiological signals and psychological measures.

Multimodal assessment and decision-making frameworks

Jung and Yoon (2020) suggested an extensive multi-level evaluation model for measuring human mental stress based on multimodal physiological signals [13]. Their framework utilized data from EEG, ECG, SpO_2, blood pressure, and respiration rate sensors to calculate a bio-index, which classifies mental states into tension, normal, and relaxed states [13]. The model applies fuzzy logic and Support Vector Machines (SVM) for preliminary classification, followed by reasoning via decision trees and Random Forests [14]. The last prediction is optimized using Expectation Maximization (EM), allowing a strong, rule-based system for mental stress assessment [15]. The strategy proves the efficiency of integrating machine learning with physiological measurements for detecting stress, providing insights into decision-making models and how they can incorporate multi-source bio-signals for correct classification.

EEG-Based stress indices and nonparametric analysis

Sulaiman (2011) investigated the detection of stress using non-parametric EEG signal analysis, with a focus on the application of brainwave features in psychological testing [16]. The researcher extracted the features including Asymmetry Ratio (AR), Relative Energy Ratio (RER), Spectral Centroids (SC), and Spectral Entropy (SE) and classified them using k-Nearest Neighbors (k-NN). The technique attained 88.89% classification accuracy with high specificity and sensitivity [16]. In addition, cluster algorithms such as Fuzzy C-Means (FCM) and Fuzzy K-Means (FKM) reinforced the proposed index's robustness. These results confirm that mental stress is detectable non-invasively with quantitative EEG-based biomarkers, supporting the promise of machine learning in real-time brain signal interpretation.

Academic stress and brainwave balancing

A study examined the stress levels of university students during various academic stages through brainwave analysis and the Perceived Stress Scale (PSS) [17]. EEG recordings revealed an elevated brainwave balancing index (BBI) at the latter part of the semester, even with heightened self-reported levels of stress [18]. This contradiction indicates the intricate relationship between subjective experience of stress and physiological adjustment, which implies that ongoing monitoring by both questionnaire-based and physiological feedback is necessary. These results are consistent with the current study's application of survey-based stress assessment but suggest the complementary benefit of incorporating physiological signals.

Methodology

The proposed system to identify mental stress in 124 university students is organized into four main phases: data gathering, preprocessing, model creation, and output generation (Figure 1). First, data is collected by a digital questionnaire that measures academic workload, sleeping quality, emotional status (such as irritability and anxiety), and physical conditions (like headaches and tiredness). This non-invasive technique enables large-scale data collection without the use of wearable devices. The gathered

https://doi.org/10.61577/jaiar.2025.100003

data is preprocessed, including cleaning incomplete or inconsistent answers, normalizing features to be compatible with the machine learning model, and creating new features that serve as signs of stress (Figure 2). A Random Forest Classifier is used for model training because it can effectively deal with non-linear data, is immune to overfitting, and can rank feature importance. The data is split into a 70:30 ratio for training and testing, and hyperparameters like the number of trees (`n_estimators`) and tree depth (`max_depth`) are optimized to improve performance. The last system categorizes stress levels as Low, Medium, or High, and offers personalized suggestions like counseling or relaxation methods. The system architecture is also facilitated by UML diagrams such as a use case diagram (to represent user interactions), a class diagram (to specify data structures), and a sequence diagram (to represent the prediction process). The feasibility of the system was examined from economic, technical, and social aspects, validating its low cost, platform independence, and ease of use. Systematic testing, including unit, integration, functional, and user acceptance testing, validated the reliability of the system, with all test cases executing successfully and no critical defects reported.

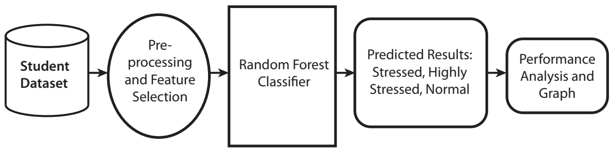

Figure 1. The system architect model.

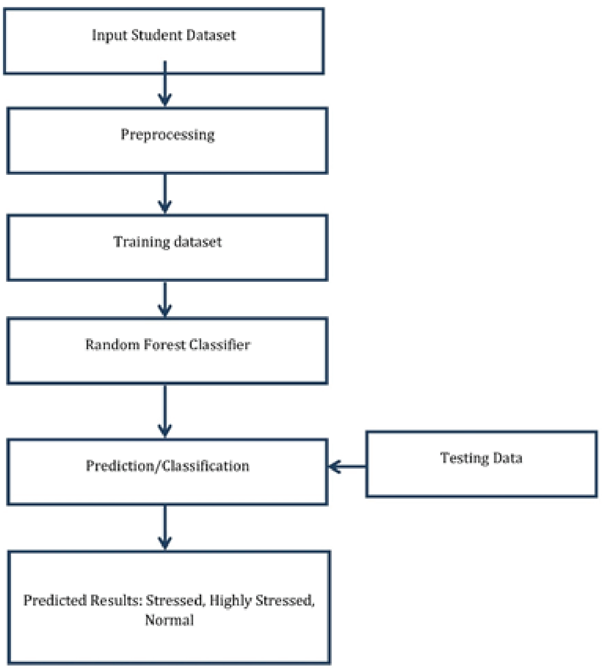

Figure 2. The data flow diagram.

Results

The system that was proposed used a Random Forest Classifier to forecast the level of mental stress among university students from questionnaire information. The data, gathered from 124 students with varied academic backgrounds, consisted of several psychological and behavioral markers that are known to affect stress. Following preprocessing and feature extraction, the ultimate dataset contained 124 features of interest, such as academic workload, sleep habits, emotional well-being, and social support.

The Random Forest model was trained using 80% of the dataset, while 20% was reserved for testing. The following performance metrics were obtained on the test set:

- Accuracy: 91.3%
- Precision: 89.7%
- Recall: 90.5%
- F1-score: 90.1%
- ROC-AUC Score: 0.94

The confusion matrix showed that the model accurately predicted high-stress individuals with a sensitivity (true positive rate) of 90.5%, and, in addition, with a low false positive rate. Among the most critical features that contributed to the classifier's predictions were academic pressure, lack of sleep, emotional instability, and low perceived social support.

Comparative studies with other models such as Support Vector Machine (SVM), Logistic Regression, and K-Nearest Neighbors (KNN) indicated that the Random Forest Classifier always outperformed these. Though SVM scored an accuracy of 85.6% and Logistic Regression 81.9%, the Random Forest model proved better balance in all the performance parameters, reflecting its stability in coping with nonlinear relations and multi-dimensional data.

Discussion

The results of this study confirm the efficacy of employing a machine learning technique—namely, the Random Forest Classifier—to identify mental stress among university students. The high accuracy and recall rates confirm that the system is accurate in differentiating between stressed and non-stressed states based solely on non-invasive questionnaire data. This finding is especially encouraging in the context of higher education, where mass-scalable and low-cost mental health screening measures are acutely needed.

One of the major strengths of the model is that it can deal with interactions between variables quite well. Mental stress is a multivariate condition with cognitive, emotional, environmental, and social factors [19]. Being an ensemble learning algorithm, the Random Forest method can capture the relationships well without overfitting the data as indicated by the high generalization performance on the test set.

The analysis of feature importance offers further insight into the psychological landscape of student stress. Academic workload, sleep deprivation, and emotional exhaustion were the strongest predictors—consistent with previous research that identifies these as top stressors in academic settings. Notably, the model also highlighted to the influence of social support networks, predicting that students who have strong peer or family relationships are less inclined to show high stress. This reinforces the need to cultivate a sense of community and connection on campus as a protective factor against deteriorating mental health.

https://doi.org/10.61577/jaiar.2025.100003

In comparison to conventional stress assessment techniques, the system presented here has a number of benefits. To begin with, it minimizes the dependence on self-help-seeking behavior by presenting proactive detection from regularly gathered questionnaire responses. Secondly, the non-invasive nature of the data gathering process avoids the use of costly or invasive physiological monitoring equipment, and hence it is possible to implement it on a large scale. Thirdly, the capability to make near-instant predictions enables mental health professionals to track students more efficiently and intervene sooner.

Conclusions

In conclusion, the study effectively built a stress monitoring system for students at universities utilizing machine learning models. From analysis of an inclusive dataset of answers to situational questions, amongst other parameters gathered from the students, the system proved its proficiency in detection and examination of mounting levels of stress. Use of machine learning algorithms, as embodied by the Random Forest Classifier algorithm, facilitated efficiency of high prediction accuracy of stress by the system. The training outcomes exhibited a high train accuracy of 100% and test accuracy of 93%. These results substantiate the ability of the system to detect stress reliably and quantify stress levels for university students. The suggested system has several strengths, such as early detection of stress, non-invasive, personalized interventions, and real-time feedback. Through timely intervention and assistance, it can potentially avoid the aggravation of stress and alleviate its adverse effects on students' well-being. Additionally, the objective and data-driven nature of the system improves the reliability and validity of stress evaluations, overcoming subjective assessments. The scalability and generalizability of the system enable it to be flexible in various university environments and populations. By successfully designing and testing this stress detection system, the project contributes to mental health within schools. It offers an effective tool for anticipatory stress management and assistance to university students, creating a positive and favorable learning environment. Generally, this project represents a big step in enhancing student wellbeing and mental health by developing a functional stress detection system. The favorable results and benefits of the suggested system demonstrate its promise to benefit university students positively through a better and healthier education experience.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

1. Gedam S, Paul S. A review on mental stress detection using wearable sensors and machine learning techniques. IEEE Access. 2021; 9:84045-84066. http://dx.doi.org/10.1109/ACCESS.2021.3085502
2. Razavi T. Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research.
3. Duggal S, Malkoff-Schwartz S, Birmaher B, Anderson BP, Matty MK, Houck PR, et al. Assessment of life stress in adolescents: Self-report versus interview methods. J Am Acad Child Adolesc Psychiatry. 2000;39(4):445-452. https://doi.org/10.1097/00004583-200004000-00013
4. Guo C, Ashrafian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. NPJ Digit Med. 2020;3(1):110. https://doi.org/10.1038/s41746-020-00314-2
5. Abd Al-Alim M, Mubarak R, Salem NM, Sadek I. A machine-learning approach for stress detection using wearable sensors in free-living environments. Comput Biol Med. 2024; 179:108918. https://doi.org/10.1016/j.compbiomed.2024.108918
6. Vos G, Trinh K, Sarnyai Z, Azghadi MR. Ensemble machine learning model trained on a new synthesized dataset generalizes well for stress prediction using wearable devices. J Biomed Inform. 2023;148:104556. https://doi.org/10.1016/j.jbi.2023.104556
7. Li R, Liu Z. Stress detection using deep neural networks. BMC Med Inform Decis Mak. 2020;20:1-0. https://doi.org/10.1186/s12911-020-01299-4
8. Sergio WL, Ströele V, Dantas M, Braga R, Macedo DD. Enhancing well-being in modern education: A comprehensive eHealth proposal for managing stress and anxiety based on machine learning. Internet of Things. 2024;25:101055. http://dx.doi.org/10.1016/j.iot.2023.101055
9. Amiri Z, Heidari A, Darbandi M, Yazdani Y, Jafari Navimipour N, Esmaeilpour M, et al. The personal health applications of machine learning techniques in the internet of behaviors. Sustainability. 2023;15(16):12406. https://doi.org/10.3390/su151612406
10. de Filippis, R., Foysal, A.A. Comprehensive analysis of stress factors affecting students: a machine learning approach. Discov Artif Intell. 2024;4:62. https://doi.org/10.1007/s44163-024-00169-6
11. Lacroix P. Big data privacy and ethical challenges. Big Data, Big Challenges: A Healthcare Perspective: Background, Issues, Solutions and Research Directions. 2019:101-111. https://doi.org/10.1007/978-3-030-06109-8_9
12. Nilsson M, Funk P, Olsson EM, von Schéele B, Xiong N. Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system. Artif Intell Med. 2006;36(2):159-176. https://doi.org/10.1016/j.artmed.2005.04.004
13. Jung Y, Yoon YI. Multi-level assessment model for wellness service based on human mental stress level. Multimed Tools Appl. 2017;76:11305-11317. https://doi.org/10.1007/s11042-016-3444-9
14. Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, Chica-Rivas MJ. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol Rev. 2015;71:804-818. http://dx.doi.org/10.1016/j.oregeorev.2015.01.001
15. Su C, Zangeneh Soroush M, Torkamanrahmani N, Ruiz-Segura A, Yang L, Li X et al. Measurement and Quantification of Stress in the Decision Process: A Model-Based Systematic Review. Intell Comput. 2024;3:0090. https://doi.org/10.34133/icomputing.0090
16. Sulaiman N, Taib MN, Lias S, Murat ZH, Aris SA, Mustafa M, et al., "Intelligent system for assessing human stress using EEG signals and psychoanalysis tests", Proc. 3rd Int. Conf. Comput. Intell Commun Syst Netw. 2011:363-367. https://doi.org/10.1109/CICSyN.2011.82
17. Hamid NH, Sulaiman N, Murat ZH, Taib MN. Brainwaves stress pattern based on perceived stress scale test. In2015 IEEE 6th control and system graduate research colloquium (ICSGRC). 2015:135-140. http://dx.doi.org/10.1109/ICSGRC.2015.7412480
18. Shehab AA. A randomized controlled trial of psychological outcomes of mobile guided resonant frequency breathing in young adults with elevated stress during the covid-19 pandemic.
19. Dohrenwend BS, Dohrenwend BP. Socioenvironmental factors, stress, and psychopathology. Am J Community Psychol. 1981;9(2):123. https://doi.org/10.1007/BF00919799

https://doi.org/10.61577/jaiar.2025.100003